

LAURENT BRAQUET ET DAVID MOUREY

Économie

Principes fondamentaux

Préface de Patrick Artus
Avant-propos d'Agnès Bénassy-Quéré

+ Classes prépas -
IEP - Concours

Ouvertures
économiques



deboeck **B**
SUPÉRIEUR

LAURENT BRAQUET ET DAVID MOUREY

Économie Principes fondamentaux

Sommaire

[Préface](#)

[Avant-propos](#)

[Préambule](#)

[Remerciements](#)

[Chapitre 1 - Principes de base de l'analyse économique](#)

[Chapitre 2 - Une brève histoire de la pensée économique](#)

[Chapitre 3 - Principes fondamentaux de microéconomie](#)

[Chapitre 4 - Les interdépendances entre agents et opérations économiques](#)

[Chapitre 5 - L'entreprise dans l'économie de marché](#)

[Chapitre 6 - La monnaie et le financement de l'économie](#)

[Chapitre 7 - Les rouages fondamentaux du marché et ses limites](#)

[Chapitre 8 - Les sources de la croissance économique à long terme](#)

[Chapitre 9 - L'instabilité de la croissance économique à court terme](#)

[Chapitre 10 - La croissance économique et la dynamique démographique](#)

[Chapitre 11 - La soutenabilité de la croissance économique](#)

[Chapitre 12 - Marché du travail, emploi, chômage](#)

[Chapitre 13 - Inégalités et justice sociale](#)

[Chapitre 14 - L'économie ouverte : commerce international et mondialisation](#)

[Chapitre 15 - L'économie ouverte : firmes multinationales et fragmentation du processus productif](#)

Chapitre 16 - L'économie ouverte : l'intégration économique européenne

Chapitre 17 - Les déséquilibres macroéconomiques et financiers internationaux

Chapitre 18 - Principes de politique économique

Chapitre 19 - Globalisation financière, instabilité financière et régulation des marchés financiers

Chapitre 20 - Crises financières, crises économiques et interventions des pouvoirs publics

Préface

Le livre de Laurent Braquet et David Mourey me paraît être un exemple de la manière dont on doit enseigner l'économie. Il réalise en effet une confrontation permanente entre l'histoire de la pensée économique et les différentes controverses associées, l'histoire des faits économiques et des institutions, la définition des concepts économiques, de comptabilité privée, de comptabilité nationale et les modèles usuels de la microéconomie et de la macroéconomie.

Les étudiants qui s'intéressent par exemple au comportement des entreprises trouvent dans ce livre la définition d'une entreprise et d'un entrepreneur chez les grands auteurs, l'analyse microéconomique du comportement de l'entreprise, la description de la nature des entreprises en France, l'analyse de la comptabilité d'entreprise, les débats sur la gouvernance des entreprises et la nature du capitalisme, y compris les évolutions les plus récentes : rôle des actionnaires, Responsabilité Sociale de l'Entreprise.

Les lecteurs qui s'intéressent par exemple aussi à la croissance trouvent des faits stylisés sur l'évolution de la croissance potentielle, les grands modèles explicatifs de la croissance, les débats modernes sur la « stagnation séculaire », une vision de très long terme de la croissance, les débats sur les liens entre innovation, institution et croissance, l'analyse des cycles économiques et des politiques contracycliques, l'examen du concept de croissance durable et des politiques climatiques.

Je pourrais ainsi multiplier les exemples avec le chômage et le fonctionnement du marché du travail, les inégalités et la protection sociale, le commerce mondial, l'intégration européenne, les déséquilibres globaux et les crises financières, les politiques monétaires et budgétaires.

La seule manière d'apprendre utilement l'économie est bien celle proposée dans ce livre : l'interaction entre l'analyse historique des faits et des institutions, la réflexion théorique, l'utilisation des modèles, la réflexion sur les controverses entre économistes.

Nous sommes aujourd'hui dans une phase de remise en cause des certitudes qu'on croyait acquises en économie. Les bienfaits de la globalisation économique et financière et du libéralisme semblent maintenant dominés par leurs effets négatifs sur l'emploi, la pauvreté, la stabilité financière. L'innovation, les Nouvelles Technologies ne semblent pas conduire à un supplément de croissance, mais à l'ouverture des inégalités. Les dérèglementations financières, la libre circulation des capitaux ont abouti à l'instabilité de la circulation internationale des capitaux et à la déstabilisation des économies des émergents. La rigueur dans les politiques économiques semble être remplacée par l'idée qu'on peut utiliser de manière plus active les politiques monétaires et budgétaires.

Cette remise en cause présente des certitudes, l'apparition de nouvelles théories, qui en réalité font appel à des réflexions théoriques anciennes et sont liées à une approche interdisciplinaire (historique, sociologique, politique...) de l'économie. C'est dans ce cadre que le livre remarquable de Laurent Braquet et David Mourey trouve toute son utilité.

Patrick Artus

Chef économiste et membre du comité exécutif de Natixis
Professeur d'économie à l'École d'Économie de Paris

Avant-propos

L'économie a envahi nos vies. Nos vies personnelles d'abord, avec les problèmes d'emploi, de logement, de pouvoir d'achat. Nos vies au travail, ensuite, avec la concurrence internationale, les réorganisations, les exigences de productivité. Nos vies citoyennes, enfin, lorsque l'économie devient omniprésente dans le débat politique.

Pourtant, l'économie relève souvent du mystère, voire de la magie dangereuse. Comment les difficultés de ménages pauvres américains pour rembourser leurs emprunts immobiliers ont-elles pu mettre en péril des banques prétendument puissantes ? Pourquoi la Banque centrale européenne ne nous envoie-t-elle pas tout simplement les billets de 50 euros qui, dépensés en produits fabriqués dans la zone euro, pourraient faire repartir la croissance ? Pourquoi ne met-on pas à la retraite anticipée les travailleurs seniors pour faire de la place aux jeunes ? Et pourquoi ne pas annuler la dette des États ? Autant de questions que chacun se pose sans forcément trouver une réponse intelligible.

Car les experts abondent, mais ils parlent une langue étrange dont les mots clés sont croissance potentielle, effet de levier, solde structurel, stabilisateurs automatiques, contrainte budgétaire intertemporelle, etc. Que faire ? Abandonner la partie ? Tirer au sort son bulletin de vote ? S'en remettre à un gourou qui, lui au moins, parle distinctement – tant pis si ses raisonnements ne respectent pas la loi de l'offre et de la demande ni les contraintes de budget des États ou des familles ?

Non, il faut exiger de comprendre. Refuser le jugement de valeur sans en avoir saisi le soubassement ; refuser le charabia dans lequel il est si facile de se draper ; refuser d'opposer l'offre et la demande, les flux et les stocks, le présent et le futur, etc. Pour cela, nous avons besoin de

passeurs, qui prendront le temps de nous mener pas à pas, de nous frayer un chemin dans la jungle de l'économie. Sans jamais céder à la facilité de l'à-peu-près, le passeur nous explique les enchaînements logiques, repère ce qui fait consensus et discute les points de divergence.

L'économie n'est pas une science de salon. Elle exige un apprentissage, une maîtrise des outils de base. On ne peut pas affirmer tout et son contraire. Certaines questions, comme l'impact du coût du travail sur l'emploi, ou bien l'effet de la politique monétaire sur le taux de change, sont aujourd'hui relativement bien connues. D'autres, comme les effets de la politique budgétaire sur l'activité et l'emploi, font encore l'objet de discussions passionnées. Il est important de savoir distinguer ces deux configurations.

Pour découvrir un territoire, rien de mieux que la marche à pied. Le rythme lent de la marche permet de s'imprégner de la géologie, de la flore, de l'habitat, des hommes et des femmes qui peuplent le lieu. Il en est de même de l'apprentissage de l'économie. Surtout ne pas aller trop vite. Forts de leur expérience d'enseignants et de leur engagement citoyen, David Mourey et Laurent Braquet nous offrent un merveilleux guide de randonnée, à suivre pas à pas pour ne pas se perdre sur les flancs des pentes escarpées.

Agnès Bénassy-Quéré
17 avril 2015

Préambule

« *It's the economy, stupid.* »

« *Pense à l'économie, idiot !* »

Telle était la phrase, devenue célèbre et désormais gravée dans la mémoire collective, que le Président des États-Unis Bill Clinton avait affichée au-dessus de son bureau afin de ne jamais oublier l'économie et son omniprésence dans nos vies quotidiennes, dans nos démocraties modernes. En tant que citoyens, une fois que nous avons intériorisé cette réalité incontournable, nous devons faire feu de tout bois pour mieux comprendre les mécanismes fondamentaux et les logiques implacables qui gouvernent nos choix économiques, individuels et collectifs, aux effets sociaux, démographiques, écologiques et sociétaux. Sans prétendre aucunement à l'exhaustivité, ce livre vise à aider tous ceux qui s'intéressent sérieusement au fonctionnement de nos économies, à mieux en appréhender les fondamentaux.

Les principaux concepts, du langage économique de base, y sont abordés dans le cadre des problématiques et questionnements qui traversent les débats de notre temps. L'objectif est de fournir à chacun des clés pour comprendre l'économie réelle et non de compiler des connaissances déconnectées de toute réalité, en dehors de toute interrogation légitime et de prendre des distances avec nos avis et opinions qui n'expliquent rien ! Nous nous sommes imposés une autre contrainte majeure mais cruciale : éviter toute formalisation inutile à la compréhension de l'essentiel. Cependant, par souci de rigueur, de cohérence et de pertinence, nous avons fondé notre texte sur une démarche scientifique, au sens où la complémentarité des approches théoriques et empiriques reste intangible. En effet, théories et faits ne s'opposent pas comme on l'entend trop souvent, mais se combinent pour nous permettre de lire, de décrypter ces

réalités qui ne se produisent pas spontanément, mais qui sont construites par les croisements de regards théoriques confrontés aux faits statistiquement construits et nécessaires à la procédure de réfutation. Dans ce cadre, les principales questions économiques sont abordées avec un langage qui se veut clair et le plus simple possible. Pour autant, comme il n'était nullement question de sacrifier la qualité et la rigueur pour éviter de passer de la simplification et de la clarification pédagogique au simplisme fallacieux, le lecteur devra produire quelques efforts intellectuels minimums.

Les vingt chapitres qui composent ce livre peuvent être lus et étudiés selon l'ordre qui convient au lecteur et à ses besoins.

Ce livre vise un public large. Il s'adresse aussi bien aux lycéens des séries ES et STMG qui étudient l'économie qu'aux étudiants de Licence de sciences économiques et de gestion, des diverses classes préparatoires B/L et ECE, des étudiants des IEP, des candidats au CAPES, au CAPET et à l'agrégation de SES ou d'Économie-gestion... Il peut être utile au citoyen également.

*

Laurent Braquet est professeur de SES au lycée Gustave Flaubert de Rouen, formateur en SES dans l'académie de Rouen, et enseigne en Classes préparatoires au concours commun des IEP.

David Mourey est professeur de Sciences économiques et sociales au lycée Charles Le Chauve à Roissy-en-Brie, formateur en SES dans l'académie de Créteil, chargé de TD de Politique économique à l'UFR d'économie-gestion de Marne-la-Vallée, professeur d'économie à l'ISMaPP, pour un cours d'« Introduction à l'économie » et Fondateur, organisateur et modérateur depuis 2005 des « Rencontres économiques » de Pontault-Combault et de Paris.

Ils sont les auteurs des ouvrages suivants, parus chez De Boeck Supérieur :

Comprendre les fondamentaux de l'économie : Introduction approfondie à l'économie, 2015

Politiques économiques, Sup en poche, 2017

Économie monétaire et financière, Sup en poche, 2018

Marché du travail, emploi, chômage, Sup en poche, à paraître
(juin 2019), David Mourey et Jean-Paul Brun

Remerciements

Nous tenons à remercier les éditions De Boeck Supérieur pour nous avoir accordé leur confiance et nous avoir donné la chance et l'opportunité d'écrire une nouvelle édition de ce manuel. En particulier, nous adressons nos sincères remerciements à Lucie Verlinden pour l'attention qu'elle porte à la conception de l'ouvrage, pour ses précieux conseils et sa patience, à Dominique De Raedt pour sa confiance, sa compréhension et son soutien depuis de longues années, mais également toutes les autres personnes de l'équipe sollicitées pendant les différentes étapes de la conception de ce manuel.

Nous adressons nos sincères remerciements à Patrick Artus pour avoir accepté de rédiger la préface de cette nouvelle édition et à Agnès Bénassy-Quéré pour avoir accepté de rédiger la préface de l'édition de 2015. Nous sommes conscients de notre chance et des encouragements qui nous sont témoignés.

Nous remercions également nos collègues et co-auteurs dans la collection Sup en Poche Économie (De Boeck Supérieur), Jean-Paul Brun, Patrice Canas, Jérôme Lecoindre et Bernard Schwengler, pour leur relecture de différents chapitres de l'ouvrage. Nous restons, bien entendu, seuls responsables des erreurs ou omissions restantes.

Nous remercions aussi nos collègues qui, par leurs débats et questionnements récurrents, ont alimenté nos réflexions scientifiques, didactiques et pédagogiques. Nous remercions tous nos élèves et étudiants qui depuis des années nous obligent par leurs interrogations, incompréhensions, suggestions... à repenser la structure de nos cours, de nos dossiers, l'enchaînement de nos explications... Sans eux, ce livre n'aurait aucune raison d'être. Ils sont la principale source de nos inspirations et de notre motivation et nous sommes

heureux chaque fois que nous parvenons à les faire avancer sur le sentier sinueux et semé d'obstacles de la compréhension de l'économie.

David Mourey tient à remercier sincèrement tous les économistes et autres grandes personnalités de l'économie qui ont accepté de participer à des conférences, débats ou colloques pédagogiques pour lycéens et étudiants essentiellement depuis 2005 à Pontault-Combault et depuis 2008 à Paris, au Palais du Luxembourg (Sénat), à la Banque de France, à la Fédération bancaire française... Ces remerciements s'adressent aussi aux très nombreux partenaires qui ont permis à ces rencontres d'avoir lieu. Vous pourrez retrouver l'historique de ce Cycle de conférences, débats et colloques pédagogiques sur le blog *Démocratie, Économie et Société* créé en janvier 2007 (<http://www.davidmourey.com/>) et une grande partie des vidéos déjà en ligne sur la *page YouTube des « Rencontres économiques »* : <http://www.youtube.com/user/DavidMourey?feature=watch>.

PRINCIPES DE BASE DE L'ANALYSE ÉCONOMIQUE

« Le plus grand ennemi de la connaissance n'est pas l'ignorance, c'est l'illusion de la connaissance. » - Stephen Hawking, 1942-2018

« La théorie ne fournit aucun ensemble constitué de conclusions immédiatement utilisables pour définir les politiques. Il s'agit d'une méthode plutôt que d'une doctrine, d'une série d'outils intellectuels qui aident leurs détenteurs à tirer des conclusions correctes. » - John Maynard Keynes, 1883-1946

« Face au réel, ce qu'on croit savoir offusque ce qu'on devrait savoir. Quand il se présente à la culture scientifique, l'esprit n'est jamais jeune. Il est même très vieux, car il a l'âge de ses préjugés. Accéder à la science, c'est, spirituellement rajeunir, c'est accepter une mutation brusque qui doit contredire un passé. » - Gaston Bachelard, 1938

« Comme pour la plupart des disciplines scientifiques, la recherche en économie requiert une combinaison de théorie et d'empirique. La théorie fournit le cadre de pensée. Elle est aussi la clé pour la compréhension des données : sans théorie, c'est-à-dire sans grille de lecture, les données ne sont que des observations intéressantes, qui ne disent rien sur ce que l'on doit en déduire pour la politique économique. Inversement, la théorie se nourrit de l'empirique, qui peut invalider ses hypothèses ou ses conclusions, et ainsi conduire soit à une amélioration, soit à un abandon de la théorie en question. » - Jean Tirole, 2018

SOMMAIRE

[1.1 Science et démarche scientifique](#)

[1.2 Principes scientifiques de l'économie](#)

[1.3 Quels sont les différents niveaux d'analyse en économie ?](#)

[1.4 Problèmes et principes économiques fondamentaux](#)

[1.5 De quelques pièges du raisonnement en science économique](#)

1.1 SCIENCE ET DÉMARCHE SCIENTIFIQUE

L'époque moderne tient la science en haute estime. La science envahit tous les domaines de la connaissance. Simultanément, il existe de plus en plus de scientifiques dans diverses disciplines qui se réclament de la méthode empirique (méthode expérimentale) consistant à recueillir des faits par de soigneuses observations et expériences puis à en tirer des lois et des théories par une procédure logique. « *Sans la possibilité de mesurer, le savoir n'est qu'une peau de chagrin* » peut-on lire sur la façade du bâtiment de sciences sociales de l'Université de Chicago. Mais c'est une vision trompeuse de la science. En effet, il n'existe aucune méthode permettant de prouver que les théories scientifiques sont vraies ou probablement vraies. La mesure, en tant que preuve, n'est pas indispensable.

Une science est constituée d'un ensemble de théories.

1.1.1 *Une théorie n'est pas une doctrine*

Le terme **théorie** vient du grec *theôria*, qui désigne l'action de voir, d'observer, ou encore du mot *theôrein* qui correspond à une contemplation de l'esprit. Une théorie est donc une représentation, nécessairement simplifiée, de la réalité que l'on ne peut saisir, mais que l'on cherche à expliquer. Dans ces conditions, affirmer qu'une théorie n'est pas réaliste n'a aucun sens ! Le plus souvent, une théorie prend la forme d'un modèle mathématique, c'est-à-dire d'un ensemble d'équations. On ne peut pas toujours procéder à des expérimentations et à des vérifications (donner des preuves). Dans ces conditions, on procède par propositions, réfutées ou non réfutées.

Une **doctrine**, terme qui trouve son origine dans le latin *doctrina*, correspond à un ensemble de propositions théoriques considérées comme vraies.

De fait, par cette simple considération, on s'éloigne de la théorie et donc de la science. Il y a souvent un abîme entre la doctrine et la théorie qui lui sert de support. La doctrine ne vise pas à expliquer les phénomènes, mais cherche plutôt à dire comment les choses devraient être. La doctrine se présente comme un objectif incontournable à atteindre.

1.1.2 Science et démarche scientifique

1.1.2.1 De l'inductivisme au raisonnement hypothético-déductif

Selon l'inductivisme naïf, la science commence par l'observation. Sans préjugés, l'observateur scientifique doit produire des énoncés d'observation qui seront la base sur laquelle reposeront les lois et les théories. Les énoncés singuliers sont des énoncés d'observation qui dépendent des sens de l'observateur. Ils portent sur des événements particuliers. Les énoncés généraux portent sur des propriétés. Les énoncés universels correspondent aux lois et théories qui constituent le savoir scientifique.

Le passage d'un énoncé singulier à un énoncé universel s'opère par légitimation, sous conditions de généralisation.

La théorie guide l'observation, laquelle présuppose l'existence d'une théorie. Les théories sont des conjectures ou des suppositions librement pensées qui s'efforcent de résoudre les problèmes posés aux théories précédentes. Les théories sont confrontées rigoureusement à l'observation et à l'expérience par le biais de tests. Seules celles qui résistent aux tests durent, survivent. **Selon l'analyse du philosophe Karl Popper (1902-1994), la science progresse ainsi par conjectures et réfutations.** Aucune théorie n'est absolument vraie, elle est seulement la meilleure disponible pour un état donné de la connaissance. La question n'est pas de savoir si une théorie est vraie, mais seulement de déterminer la plus efficace d'entre elles pour comprendre le réel.

1.1.2.2 La réfutation pour légitimer la science : les fondements poppériens de la science

La logique est en faveur de la réfutation, du falsificationnisme

Si on ne peut montrer qu'une théorie est vraie, on peut, en revanche, montrer que certaines théories sont fausses à partir des résultats d'observation et d'expérience. Ainsi, des déductions logiques qui prennent comme prémisses des énoncés singuliers d'observation, supposés vrais, peuvent nous conduire à rejeter des lois et/ou théories. Autrement dit, la fausseté d'énoncés universels peut être logiquement déduite d'énoncés singuliers appropriés. Le falsificationniste exploite complètement cette propriété logique. Par exemple, tous les corbeaux ne sont pas noirs. Il suffit de voir un seul corbeau blanc...

La réfutabilité, falsifiabilité, devient le critère de délimitation pour les théories

Toute hypothèse, ou tout système d'hypothèses, doit satisfaire une condition fondamentale pour acquérir le statut de loi ou de théorie scientifique. Pour faire partie de la science, une théorie (une hypothèse) doit être falsifiable. Une hypothèse est falsifiable si la logique autorise l'existence potentielle d'un énoncé d'observation, au moins, qui lui est contradictoire. En effet, les énoncés infalsifiables ne nous apprennent rien sur le monde (par exemple, l'énoncé : soit il pleut, soit il ne pleut pas). Une loi ou une théorie doit nous fournir des informations sur le comportement du monde... et être falsifiable, sinon elle doit être rejetée ! Si une théorie a un contenu informatif, c'est-à-dire qu'elle nous apprend quelque chose de nouveau, elle doit nécessairement courir le risque d'être falsifiable.

Le degré de falsifiabilité : clarté et précision des hypothèses

Plus une théorie est falsifiable, meilleure elle est. Une très bonne théorie énonce un grand nombre d'observations de portée très générale. Elle est donc hautement falsifiable. La qualité d'une théorie est d'autant plus grande qu'elle fournit davantage d'énoncés susceptibles d'être falsifiés. Dans ce

cas, elle devient plus falsifiable et sa résistance à toute falsification lui confère une supériorité par rapport aux autres théories. Toute théorie falsifiée doit être rejetée. C'est une règle de la méthode scientifique. En fait, la science progresse par essais et erreurs, par conjectures et réfutations. Mais cette importance qu'accordent les falsificationnistes au rôle de la falsification, de la réfutation, est contraire à l'intuition (et pourtant...) et sera donc critiquée.

1.1.2.3 Le falsificationisme comme condition des progrès de la science

Dans une perspective falsificationniste, on ne peut jamais dire qu'une théorie est absolument vraie, même si elle a surmonté de multiples tests très rigoureux, mais on peut dire qu'elle est supérieure à celles qui l'ont précédée et qui ont été falsifiées.

On déduit de ce qui précède que la science commence par des problèmes, ces problèmes sont plus ou moins le résultat direct d'observations, mais des observations qui ne posent des problèmes qu'à l'aune d'une théorie donnée. Le falsificationisme est sauf, mais non l'inductivisme naif ! **La science ne commence pas par l'observation pure !** Le savoir scientifique est un savoir qui a été soumis à l'épreuve de la réfutation, falsification, par la confrontation aux observations. La science est donc objective. Il n'y a pas de place pour les opinions personnelles, les goûts et les spéculations de l'imagination.

1.1.2.4 La diversité et la complémentarité des disciplines scientifiques

Chaque discipline scientifique cherche à expliquer une dimension de la « réalité ». Parfois, plusieurs disciplines peuvent analyser le même objet sous des angles différents. La physique et la chimie analysent la matière dans des cadres différents et avec des objectifs distincts.

De manière analogue, l'économie et la sociologie analysent des comportements individuels et collectifs dans des cadres différents et avec des objectifs distincts. Il existe donc une

grande complémentarité entre les enseignements, savoirs, proposés par les différentes sciences.

1.2 PRINCIPES SCIENTIFIQUES DE L'ÉCONOMIE

Il existe de nombreuses définitions de l'économie en tant que discipline scientifique. On pourra observer de nombreuses convergences entre elles en ce qui concerne les domaines fondamentaux d'analyse : rareté, choix, production, répartition, échange...

1.2.1 Définitions

Selon Jean-Baptiste **Say**, « *l'économie politique enseigne comment se forment, se distribuent et se consomment les richesses qui satisfont aux besoins des sociétés* » (*Traité d'économie politique*, 1803).

Pour Lionel **Robbins** (1898-1984), la science économique est la « *science qui étudie le comportement humain comme une relation entre des fins et moyens rares qui ont des usages alternatifs* » (*La nature et la signification de la science économique*, 1932).

Paul-Anthony **Samuelson** (1915-2009) définit la science économique comme « *la manière dont les sociétés utilisent les ressources rares pour produire des marchandises ayant une valeur et pour les répartir entre une multitude d'individus* » (*Micro-économie*, Les éditions d'organisation, 1995).

Selon Edmond **Malinvaud** (1923-2015), « *l'économie est la science qui étudie comment des ressources rares sont employées pour la satisfaction des besoins des hommes vivant en société ; elle s'intéresse, d'une part, aux opérations essentielles que sont la production, la distribution et la consommation des biens, d'autre part, aux institutions et aux activités ayant pour objet de faciliter ces opérations* » (*Leçons de théorie microéconomiques*, 1986).

Enfin, pour Joseph **Stiglitz**, « *la science économique étudie comment les individus, les entreprises, les pouvoirs publics et d'autres organisations sociales font des choix, et comment ces choix déterminent la façon dont sont utilisées les ressources de la société* » (*Principes d'économie moderne*, Stiglitz, Walsh et Lafay, De Boeck Supérieur, 2007).

1.2.2 La querelle des méthodes en économie : méthode déductive ou méthode inductive

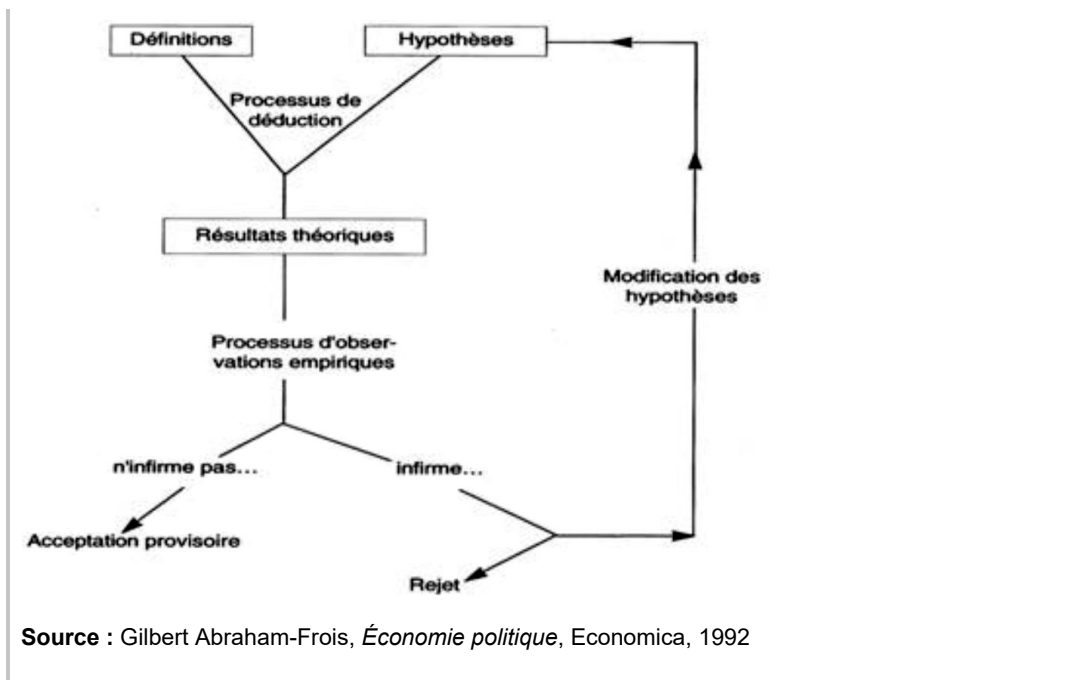
Les économistes classiques, néoclassiques... s'appuyaient déjà sur la méthode déductive : sur la base d'hypothèses simplificatrices comme la rationalité, la recherche de l'intérêt personnel, la flexibilité des prix, l'immobilité internationale des facteurs de production..., ils proposent différents modèles stylisés pour rendre le réel compréhensible : l'homo œconomicus comme agent représentatif, la main invisible, l'état stationnaire, l'économie capitaliste, le marché, le circuit économique, la décroissance de la courbe de demande...

Selon l'École historique allemande, en économie, il convient de privilégier une démarche historique et inductive. Cela consiste à partir de l'observation de faits dans le but de formuler des lois fondées sur le principe de la généralisation (cf. inductivisme).

On retrouve en économie, et dans les autres sciences sociales, la controverse entre l'inductivisme et le falsificationnisme. Pourtant, l'expérience scientifique montre que dans le cadre de la méthode déductive, l'observation ne peut se faire sans se fonder sur un concept théorique préalable. La méthode hypothético-déductive s'est donc imposée et généralisée.

1.2.3 La démarche scientifique appliquée à l'économie

On peut représenter simplement cette méthode hypothético-déductive appliquée à l'économie.



GRAPHIQUE 1.1.

La méthode hypothético-déductive appliquée à l'économie

1.2.4 Les caractéristiques d'un modèle économique

Les modèles économiques sont des représentations simplifiées de la réalité. Un modèle ne prend donc pas en compte toutes les informations. Un modèle ne se fonde que sur les éléments indispensables à une bonne compréhension de la réalité que l'on cherche à expliquer. L'exemple de la carte routière (papier ou GPS) permet de comprendre ce point : une carte routière est un modèle, car c'est une représentation simplifiée de la réalité, qui ignore de nombreux détails comme les forêts, les maisons, les panneaux de signalisation, mais qui représente précisément les routes et les chemins pour que nous puissions nous repérer au mieux. Les modèles sont donc incontournables dans le cadre scientifique.

Les modèles encouragent le débat scientifique, car un modèle étant une représentation simplifiée de la réalité, il repose sur des hypothèses par nature simplificatrices. Ces

hypothèses peuvent donc être contestées, modifiées. Un modèle suscite aussi des débats scientifiques sur ses conclusions. Un modèle est censé être réfuté ou non réfuté par les « faits », c'est-à-dire les observations empiriques. La non-réfutation n'est donc que temporaire. Cette confrontation aux « faits » entraîne aussi des débats, car on recherche toujours un meilleur modèle. Un modèle peut donc être modifié, amélioré, voire rejeté et abandonné.

1.2.5 La diversité et la complémentarité des modèles en sciences sociales

Cette complémentarité peut s'expliquer simplement par la multiplicité des angles sous lesquels l'objet d'étude peut être observé et analysé pour améliorer notre compréhension des phénomènes qu'ils véhiculent, ceux dont il est la cause ou la conséquence.

1.2.6 Analyse positive et analyse normative, jugement de valeur et jugement de fait

En économie et en sciences sociales, compte tenu des sujets analysés, il convient de ne pas confondre un énoncé objectif et un énoncé subjectif, un jugement de valeur et un jugement de fait pour reprendre la distinction opérée par Max Weber (1864-1920), une analyse positive et une analyse normative, théorie et fait.

Nous avons tous des préférences et opinions politiques, nous adhérons tous à des valeurs morales. Dans nos échanges avec autrui, nous mélangeons des données objectives, des préférences personnelles, des raisonnements plus ou moins rigoureux, des principes moraux. Il est clairement difficile d'être vraiment objectif, de ne pas confondre son opinion avec l'explication neutre et rigoureuse, de distinguer jugement de valeur et jugement de fait.

Adopter une **démarche positive** est le seul moyen d'avoir une attitude scientifique en économie. Dans un tel cadre, on peut élaborer des théories que l'on pourra confronter aux données empiriques, à distance de toute subjectivité. Face à des problèmes économiques comme le chômage ou l'inflation, l'analyse positive nous indiquera les principales causes et les principaux effets de ces phénomènes. L'**analyse normative** nous conduira à définir une hiérarchie des objectifs, des priorités pour la société, fondée sur des jugements de valeur qui ne peuvent donc être réfutables. En tant que discipline scientifique, la science économique poursuit une démarche positive. Pour autant, les économistes étant aussi des citoyens, ayant des préférences, ils peuvent (volontairement ou involontairement) utiliser leurs théories à des fins normatives pour influencer les choix qui devront être collectivement faits. Or, aucun résultat scientifique ne peut être fondé sur jugement de valeur.

1.2.7 *Les limites de l'analyse économique*

Fondamentalement, l'économie étudie la manière dont les sociétés s'organisent pour subvenir aux besoins de la population et assurer son bien-être matériel. Mais elle est confrontée, en tant que science, à des contraintes particulières. Pour le comprendre, il faut faire un détour par l'épistémologie : selon Thomas Kuhn, les révolutions scientifiques sont des changements de « paradigmes », c'est-à-dire des changements de systèmes de références ou de croyances, des ensembles cohérents de concepts et de méthodes qui organisent la recherche scientifique. Mais, à la différence de la science physique, où généralement, et malgré la persistance de controverses, un nouveau paradigme chasse l'ancien et le remplace, l'une des caractéristiques de l'économie est la *coexistence* des paradigmes et l'existence d'un pluralisme théorique. Des travaux de recherche peuvent aussi se poursuivre dans le cadre des différents courants de pensée qui existent en

économie (néoclassique, keynésien, marxiste, régulationniste, institutionnaliste, etc.).

Les désaccords entre les économistes ne viennent pas forcément des imperfections de leur science : ils **sont liés à la spécificité de leur objet d'étude**, à leur angle de vue, ou au cadre d'analyse qu'ils ont choisi. Ainsi, le paradigme keynésien a fondé l'analyse macroéconomique en se plaçant du point de vue d'un ministre des Finances désireux de réguler l'économie et de mettre en œuvre la politique économique en traitant des phénomènes économiques globaux comme l'inflation, la croissance, le chômage, le commerce extérieur, etc. ; le paradigme néoclassique analyse plutôt l'économie du point de vue des choix microéconomiques, ceux de l'individu et de l'entreprise face à des choix et des ressources limitées, en cherchant à justifier la liberté d'entreprendre et la performance de l'économie de marché si elle n'est pas perturbée par l'intervention intempestive de l'État ; tandis que le paradigme marxiste établit, quant à lui, une vision de l'économie cohérente avec les objectifs d'un courant révolutionnaire critique du capitalisme, qu'il vise à renverser. **Le travail de l'économiste est rendu plus difficile** par une autre différence fondamentale entre les sciences de la nature (physique, biologie) et **la science économique** : les premières étudient des objets extérieurs au chercheur, tandis que la seconde **étudie un objet dans lequel l'économiste est lui-même immergé**, ce qui est d'ailleurs le cas dans les autres sciences sociales (sociologie, anthropologie, ethnologie). Il peut avoir des préjugés, des opinions sur l'objet qu'il étudie, qui brouillent son jugement sur les faits observés. En raison de la difficulté à réaliser des expériences (malgré le développement de l'économie expérimentale, on ne peut déclencher une crise économique pour analyser ses effets !), de la proximité des économistes avec la décision politique (ils sont parfois les conseillers des hommes et des femmes politiques), l'ambition explicative de la science économique est alors perpétuellement menacée de devenir non scientifique, voire

de servir purement et simplement des idéologies et des doctrines. Si les sciences de la nature, qui traitent de la matière inanimée et des phénomènes naturels, sont soumises à des lois naturelles, les sciences sociales, qui, elles, décryptent les comportements de l'homme, plus ou moins rationnel, vivant en société, ne peuvent de fait être soumises à des lois intangibles. Malgré son haut niveau de reconnaissance académique (on décerne chaque année à des économistes le Prix Nobel d'économie), la science économique fait face à des critiques récurrentes et parfois violentes, d'ailleurs réactivées par l'enchaînement des crises depuis 2008 : optimisme excessif sur les capacités autorégulatrices du système économique (finance de marché), spécialisation trop poussée des travaux de recherche en économie (finance, économie industrielle, économie internationale, etc.), attachement trop grand à la modélisation et à l'usage intensif des mathématiques, etc. **La science économique est toutefois parvenue à produire un savoir robuste et un langage commun**, et à s'élever à un haut niveau d'intégration et de reconnaissance en tant que science, même si ces crises et leurs effets pourraient susciter de nouvelles pistes de recherche dans le champ de l'économie.

1.3 QUELS SONT LES DIFFÉRENTS NIVEAUX D'ANALYSE EN ÉCONOMIE ?

L'analyse économique nous invite à bien distinguer l'approche microéconomique de l'approche macroéconomique.

1.3.1 *La perspective microéconomique*

L'**individualisme méthodologique** est la démarche selon laquelle un phénomène social doit, pour être expliqué, être envisagé comme le produit d'actions individuelles. Le tout social n'est, dans cette perspective, que la somme des parties individuelles. En économie, cela se traduit par l'approche microéconomique. **La microéconomie** est la

branche de la science économique qui étudie les **comportements** des unités économiques (consommateurs et producteurs) et leurs interactions. La microéconomie est donc la science des **arbitrages** et des **choix** (de production, de consommation...) des acteurs de l'économie qui sont supposés rationnels (« *homo œconomicus* »), sous contraintes. Selon cette approche, ce sont les choix individuels d'agents supposés rationnels qui fondent les grands principes de l'échange, de la production, de la consommation, de l'investissement... L'économie dans son ensemble n'est que l'agrégation des choix et comportements d'individus isolés, mais en interaction.

La microéconomie conduit donc à adopter quelques hypothèses et principes généraux que sont : l'**individualisme méthodologique** (elle développe son raisonnement en partant des comportements individuels et des choix des agents) ; l'**analyse marginale** (elle analyse l'impact d'une évolution infinitésimale d'une variable explicative X sur une autre variable expliquée Y) ; l'**hypothèse de divisibilité des biens et des facteurs de production** (ces derniers sont divisibles en petites quantités pour pouvoir effectuer des calculs) ; l'**utilisation de fonctions dérivées** (elle utilise l'analyse des fonctions mathématiques) ; le **raisonnement « toutes choses égales par ailleurs »** (elle considère par hypothèse qu'une variable évolue alors que l'autre est supposée constante) ; le principe de **rationalité** individuelle (elle considère que l'agent a la capacité de faire le meilleur choix pour lui en vertu d'un calcul coûts/bénéfices) ; enfin, la microéconomie retient comme cadre de référence le modèle stylisé de concurrence pure et parfaite (elle considère par hypothèse qu'il y a sur le marché atomisticité des acteurs, homogénéité des produits, libre entrée et libre sortie du marché, transparence du marché et mobilité des facteurs de production).

Cette approche microéconomique sera développée dans le chapitre 3 de cet ouvrage.

1.3.2 *L'analyse macroéconomique*

Le **holisme méthodologique** est une démarche selon laquelle un fait social doit être considéré comme une chose indépendante des parties qui le constituent. Le corollaire est qu'un fait social ne peut s'expliquer qu'à partir d'un autre fait social. Autrement dit, le social s'explique par le social, le tout n'est pas la somme des parties, la société (le tout) est une réalité *sui generis*. En économie, la **macroéconomie** est la branche de la science économique qui étudie la structure, le fonctionnement et les résultats de l'économie globale. La science économique cherche à établir des lois macroéconomiques indépendantes des comportements individuels. L'économie globale n'est plus la simple agrégation d'actions individuelles en interaction. L'économie globale, tout comme la société dans son ensemble, est une réalité en elle-même, elle a ses lois propres. La macroéconomie relève d'une **approche systémique** : l'économie globale est un système économique dont les éléments sont interdépendants.

La macroéconomie cherche à expliquer des phénomènes globaux tels que l'évolution du produit intérieur brut, le taux de chômage ou le taux d'inflation. Elle analyse principalement l'économie en termes de **circuit** reliant les variables principales de l'économie : ces phénomènes sont dits « macroéconomiques » parce qu'ils sont mesurés à l'échelle de l'économie nationale, voire de plusieurs pays (comme dans le cadre de la zone euro ou de l'Union européenne).

La macroéconomie se distingue donc de la microéconomie par son domaine et son approche :

- Son **domaine d'analyse** concerne la régulation globale d'un système incluant une multitude d'agents économiques reliés par diverses opérations : principalement la détermination du revenu, la production, l'emploi et le niveau général des prix. Il inclut donc l'analyse des problèmes majeurs de l'économie nationale : le chômage, l'inflation, le manque de croissance...

– En termes d'**approche**, tandis que la microéconomie étudie les questions d'affectation des ressources rares entre diverses utilisations, la macroéconomie appréhende l'économie au niveau des fonctions globales reliant des agrégats, comme la production, le revenu, la consommation et l'investissement, ce qui n'exclut pas de prendre en compte certains prix (niveau général des prix, taux d'intérêt, taux de salaires).

1.3.3 *Les fondements microéconomiques de la macroéconomie*

Depuis la fin des années 1960, on observe l'essor des approches en termes de fondements microéconomiques de la macroéconomie. Ce sont ces approches, à travers la construction de ponts (*bridges*), qui permettent de comprendre qu'une séparation stricte semble devenue obsolète et qu'au contraire une approche « intégrée » est plus riche d'enseignements. Bien entendu, le maintien de la séparation des approches dans le cadre de travaux universitaires, de travaux de chercheurs, ne saurait être abandonné. De très nombreux auteurs ont travaillé dans ce sens : R. Clower, A. Leijonhufvud, E. Phelps, E. Malinvaud, J.P. Fitoussi, J. Stiglitz, G. Mankiw... En effet, les comportements en situation d'imperfection de l'information, d'incertitude, d'anticipations non complètement rationnelles... changent les conditions dans lesquelles les variables macroéconomiques sont déterminées et évoluent.

Les approches microéconomiques et macroéconomiques sont interdépendantes

Les millions, voire milliards, d'agents économiques font quotidiennement des choix, prennent des décisions afin de satisfaire leurs besoins. Les besoins des uns sont satisfaits grâce aux activités des autres et réciproquement. En effet, les spécialisations productives se traduisent logiquement par des échanges. Ces échanges ont pour contrepartie obligée des interdépendances multiples. Il en découle que la combinaison

de ces choix individuels décentralisés et des échanges-interdépendances va entraîner des mouvements des variables globales de l'économie et donc avoir des effets sur l'économie dans son ensemble. En retour, le contexte économique aura des conséquences sur le cadre des choix, décisions et opérations diverses des agents économiques. On ne peut séparer, au-delà des exigences de l'analyse rigoureuse, les aspects microéconomiques et macroéconomiques du fonctionnement de l'économie.

1.4 PROBLÈMES ET PRINCIPES ÉCONOMIQUES FONDAMENTAUX

1.4.1 *Le vaste champ des questions posées à la science économique*

Chacun d'entre nous est un acteur de l'économie et peut avoir une idée intuitive de ce qu'est l'économie : la vie quotidienne est constituée de faits et de décisions divers qui sont des actes et des faits économiques : changer son forfait de téléphonie mobile, payer une place de concert, créer son entreprise et remplir les formalités nécessaires, embaucher un travailleur et le rémunérer pour son travail par un salaire, décider de poursuivre ses études pour une année supplémentaire plutôt que de travailler dans un restaurant à plein temps, placer son épargne sur un livret A, installer un site de production dans un pays où le coût du travail est relativement faible ou affecter le produit des impôts prélevés au financement d'infrastructures publiques, donner des cours d'économie, écrire des livres d'économie... sont des exemples de décisions qui intéressent l'économiste.

De même, le ralentissement de la productivité globale des facteurs (PGF) des pays développés depuis la fin des années 1960, le ralentissement actuel de la croissance du PIB français, l'évolution du taux d'endettement des ménages américains, le niveau des réserves de change de la banque centrale de Chine, la crise des dettes souveraines dans la

zone euro depuis 2010 constituent des phénomènes complexes qui nécessitent de mobiliser l'analyse économique pour comprendre les mécanismes à l'œuvre.

Tous ces faits sont essentiels pour décrypter le monde contemporain : chaque jour, l'économie est partout présente, et les sujets économiques occupent une place importante dans les médias et les colloques ; des revues y sont consacrées, et des chaînes de télévision scrutent les événements marquants de la vie des entreprises et des décisions de l'État.

Quel sera le niveau de la croissance en 2019 ? Les hypothèses retenues pour construire le budget de la Nation seront-elles tenues ? Comment les fruits de cette croissance seront-ils répartis entre le facteur travail et le facteur capital ? Les impôts vont-ils baisser ? La crise de la zone euro est-elle terminée ? La déflation constitue-t-elle une menace à court terme ? La mondialisation de l'économie est-elle responsable du chômage de masse ?

Non seulement ces questions clés suscitent de nombreux analyses et commentaires dans les grands médias, mais elles occupent une partie du temps des milliers d'élèves et étudiants qui ont choisi des cursus scolaires et universitaires au sein desquels l'enseignement de l'économie est présent. Étant donné que les activités économiques sont omniprésentes, étudier l'économie est alors un impératif et devient... fondamental.

La compréhension du langage de l'économie permet d'éclairer le débat démocratique, tant pour les citoyens confrontés à des choix cruciaux face à l'offre politique et aux programmes proposés par les candidats aux élections que pour les décideurs politiques eux-mêmes, conseillés par des économistes qui leur apportent une expertise afin d'éclairer leurs arbitrages. De nombreuses enquêtes montrent que les enjeux économiques ne sont pas seulement des débats d'experts, mais des préoccupations citoyennes du quotidien largement partagées, à propos de la crise, du chômage, des

délocalisations, de la dette, de l'avenir de l'euro, etc. A *contrario*, la méconnaissance des mécanismes économiques peut laisser le champ libre à des explications idéologiques simplistes et des utopies démagogiques qui conduisent au fatalisme. Une meilleure maîtrise du langage de l'économie est alors utile pour démocratiser un débat qui semble abstrait, mais dont les conséquences ont un impact bien réel sur le devenir de nos sociétés et sur nos vies quotidiennes.

Même s'il est important de savoir répondre aux nombreuses questions économiques évoquées ci-dessus, et à d'autres encore, les économistes considèrent néanmoins qu'il existe des concepts clés, incontournables, des questions centrales, des problèmes économiques fondamentaux. Ces manières progressives et diverses d'approcher l'économie nous montrent que, d'une part, il existe différentes voies et chemins à suivre et que d'autre part, de multiples progressions sont possibles dans la mesure où nous pouvons classer les concepts clés et questions de base par degré de priorité.

1.4.2 Les trois problèmes économiques fondamentaux

Paul-Anthony Samuelson considère que les économies et la science économique doivent résoudre trois problèmes économiques fondamentaux : **que produire, comment et pour qui ?**

1.4.2.1 Que produire ?

Faut-il produire du beurre ou des canons, des biens privés ou des biens publics, des smartphones ou des écoles, des routes ou des hôpitaux... ? En amont de la question « que produire ? » se pose celle des besoins à satisfaire. Quels sont ses besoins à satisfaire et comment faut-il faire ? Quels biens et services doit-on produire pour satisfaire les besoins de la population ? Doit-on hiérarchiser les priorités ? Par exemple, doit-on produire du beurre ou des canons ? Doit-on produire des armes ou de la nourriture ?

Par exemple, dans un pays pauvre, logiquement la priorité devrait être la production de nourriture. Si on fait le choix de fabriquer des canons, il restera moins de temps et de ressources pour fabriquer du beurre. C'est le problème du **coût d'opportunité**. Il y a coût d'opportunité quand on renonce à faire un choix pour en faire un autre. Le coût économique, *lato sensu* (au sens large), représente le coût de production (en unités monétaires) auquel s'ajoute le coût d'opportunité. Chaque choix économique, chaque décision économique de produire ou de dépenser se traduit par un coût d'opportunité. Par exemple, si dans l'économie on choisit de produire des armes, on utilise des facteurs de production pour les produire. Or, les facteurs sont disponibles en quantités limitées, ils sont rares. De facto, il y a moins de quantités de facteurs de production disponibles et moins de temps pour produire du beurre, de la nourriture...

En résumé, quand on choisit de produire ceci, on renonce automatiquement à l'opportunité de produire cela. On choisit également pour qui on produit, pour satisfaire quel besoin... Le coût d'opportunité est omniprésent et inévitable.

« Tout canon fabriqué, tout lancement de bateau de guerre, tout tir de fusée signifie, en dernière analyse, un vol au détriment de ceux qui ont faim et ne sont pas nourris », rappelait Dwight D. Eisenhower.

Le choix du « que produire ? » dépend notamment de la distinction entre les besoins élémentaires (primaires) et les besoins dérivés (secondaires). La **rareté** permet d'expliquer la **nécessité de faire des choix**... La planète est un espace limité où les ressources sont rares. On ne peut pas tout produire, car les ressources productives sont rares. Les ressources naturelles sont disponibles en quantités limitées. Les autres ressources productives (facteurs de production) comme le facteur travail et le facteur capital sont en quantité limitée.

Une fois que les agents économiques ont choisi ce qui doit être produit, il convient de déterminer comment cela sera fait. La question qui en découle est : comment produire ?

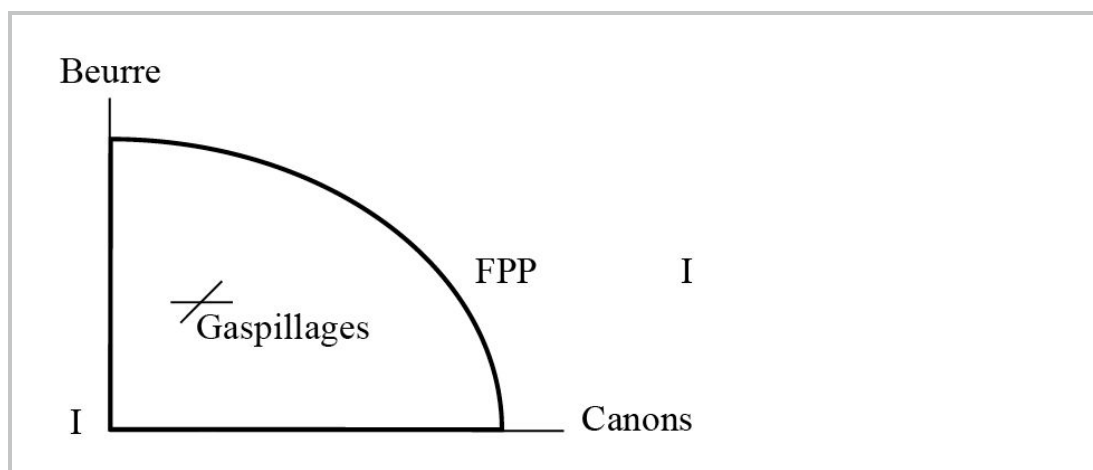
1.4.2.2 Comment produire ?

Quelle combinaison des ressources productives disponibles ? Quelle combinaison des facteurs travail et capital sera la plus efficace pour produire ? Étant donné que ces facteurs de production sont en quantité limitée, ils sont relativement rares et donc plus ou moins chers. Économiquement, il faut choisir la combinaison productive qui est la moins coûteuse et qui rapporte le plus. C'est la combinaison optimale.

Le choix d'une combinaison productive pose implicitement la question de l'efficacité de celle qui sera choisie, la question de l'existence éventuelle de gaspillages des ressources productives rares et celle de situations impossibles, car inaccessibles.

Le concept de **frontière des possibilités de production** (FPP) désigne l'ensemble des possibilités alternatives de production efficaces de deux biens et/ou services. La FPP illustre le problème du choix de la combinaison productive optimale (la plus efficace) compte tenu du choix concernant le « que produire ? ».

Le point A se situe sur la FPP. Il n'y a donc aucun gaspillage de ressources en A. C'est une combinaison efficace. Le point G se situe à l'intérieur de la FPP, zone de gaspillage. C'est une combinaison inefficace. Le point I, hors FPP correspond à une combinaison impossible : il y a insuffisance de ressources productives.



GRAPHIQUE 1.2.

Beurre ou canons

Quand, dans une économie, on choisit une combinaison productive, on choisit d'utiliser plus ou moins le facteur travail ou plus ou moins le facteur capital. Ainsi, le choix d'une combinaison se traduit, *in fine*, par un choix de répartition directe de revenus entre facteur de production et donc de répartition indirecte des biens et services produits.

Ce qui nous conduit à la troisième question économique fondamentale : pour qui produire ?

1.4.2.3 Pour qui produire ?

Les biens et services produits seront répartis par le biais d'une répartition des revenus, qui est la contrepartie de la participation des facteurs de production à l'activité productive. Compte tenu de l'existence du coût d'opportunité, quand on répond à la question « comment produire ? », on influence indirectement une répartition des revenus, et donc on influence indirectement la répartition des biens et services qui seront achetés grâce à ces revenus.

On retrouve la **logique du circuit économique** selon laquelle la production crée les revenus, les revenus permettent les dépenses et les dépenses donnent vie à la production.

Tous les agents économiques ne font pas les mêmes choix de consommation, d'investissement, d'épargne... Si la répartition du revenu global (de la valeur ajoutée) d'une économie se réalise en faveur des salaires (des salariés), ce sera au détriment de la part des profits (des entreprises). Le salaire étant surtout utilisé pour l'achat de biens de consommation alors que les profits sont davantage destinés à l'achat de biens de production. Les entreprises qui produisent des biens de consommation verront leur activité augmenter, alors que les entreprises qui produisent des biens de production verront leur activité ralentir, voire diminuer¹.

De plus, quand on fait un choix de combinaison productive, on fait donc implicitement le choix d'utiliser davantage certains facteurs que d'autres. Par exemple, on peut choisir d'utiliser davantage des machines que des humains pour produire. En conséquence, les différentes parts respectives de revenus des facteurs dans le revenu global vont changer.

Les revenus vont s'orienter vers le capital productif si dans une économie, on fait le choix de plus investir dans les machines que d'embaucher des humains. On versera plus de revenus au facteur de capital qu'au facteur de travail.

Aucune économie ne peut échapper à l'obligation de répondre à ces trois questions fondamentales, ni à la nécessité de trouver des solutions aux trois PEF. Quand on recherche la réponse à ces questions, on ne peut pas se soustraire au problème du coût d'opportunité. Dans ces conditions, chaque économie pour atteindre ces objectifs va chercher à s'organiser. Elle va mettre en place des moyens permettant aux agents économiques de résoudre les grands problèmes auxquels ils sont confrontés.

1.4.3 *Trois formes typiques d'organisation économique*

De manière générale, pour résoudre les trois problèmes économiques fondamentaux, une économie peut être régulée selon trois formes typiques d'organisation économique.

- **L'économie de marché**, économie dans laquelle un système de prix, d'offre et de demande permet la coordination de décisions décentralisées de production, consommation, investissement... en fonction du mécanisme de la main invisible.
- **L'économie planifiée**, économie dirigée par une autorité centrale (État) (qui se substitue à la main invisible du marché pour coordonner les décisions économiques centralisées [main visible]).
- **L'économie mixte**, combinaison des principes de l'économie de marché et des principes de l'économie

planifiée, c'est le cas de la plupart des économies contemporaines.

L'objectif final de toute organisation économique est de résoudre les trois problèmes économiques fondamentaux et les problèmes économiques dérivés.

1.4.4 *Cinq concepts microéconomiques centraux pour raisonner en économiste*

Dans leur ouvrage *Principes d'économie moderne*, Joseph Stiglitz et C. Walsh insistent sur cinq concepts microéconomiques centraux : **arbitrages**, **incitations**, **échange**, **information** et **distribution**. Pour eux, ces cinq concepts font référence à des comportements et mécanismes que nous devons maîtriser pour comprendre le fonctionnement de nos économies.

On retrouve indirectement les questions relatives aux problèmes économiques fondamentaux. Pour pouvoir expliquer comment les individus font des choix et les conséquences de ces choix sur les ressources de la société, il est indispensable d'examiner attentivement le rôle des arbitrages, des incitations, de l'échange, de l'information et de la distribution.

Tout choix (que produire, comment...) implique des arbitrages, assortis de coûts d'opportunité. Les choix des agents économiques sont guidés par des incitations (prix, salaires...). Les choix entraînent différents types d'échanges (pour qui...). Les décisions sont fondées sur un éventail d'information que les agents vont essayer d'obtenir et de traiter le plus rationnellement possible. Ces choix permettant de répondre aux deux premiers PEF vont se traduire par une répartition des revenus, une distribution des biens et services produits.

1.5 DE QUELQUES PIÈGES DU RAISONNEMENT EN SCIENCE ÉCONOMIQUE

Quand on souhaite réfléchir avec rigueur sur le fonctionnement de l'économie, sur les mécanismes de causalité pour mieux en comprendre le fonctionnement et les enchaînements, il convient de :

- ne pas oublier que, souvent, on raisonne en faisant implicitement l'hypothèse « ceteris paribus » (toutes choses égales par ailleurs) ;
- ne pas verser dans le sophisme « Post hoc, ergo propter hoc » (après ceci, et donc nécessairement à cause de ceci) ;
- ne pas tomber dans le sophisme de composition (le tout n'est pas la somme des parties).

1.5.1 *L'hypothèse « toutes choses égales par ailleurs »*

Il convient de ne jamais oublier que, souvent, on raisonne sous l'hypothèse « ceteris paribus », c'est à dire « toutes choses égales par ailleurs. C'est pratique et indispensable pour pouvoir tenter d'isoler l'effet d'une variable sur une autre (en faisant l'hypothèse que les autres ne varient pas), mais cela réduit la portée des causalités mises en évidence. En effet, dans la vie économique quotidienne, quand un paramètre change (un prix par exemple), rien n'est stable par ailleurs. Ce point sera explicité dans le Chapitre 7 : Les rouages fondamentaux du marché et ses limites.

1.5.2 *Ne pas confondre corrélation et causalité*

Il est tout aussi nécessaire de ne pas se laisser piéger par le sophisme « Post hoc, ergo propter hoc », autrement dit, « après ceci, et donc nécessairement à cause de ceci ». En effet, si l'événement A survient après l'événement B, A n'est

pas nécessairement la cause de B. A n'est pas nécessairement la cause de B. Le fait que par ailleurs tout ne soit pas égal implique déjà que d'autres causes sont possibles. Par exemple, une hausse du pouvoir d'achat peut s'expliquer par une baisse des prix, une hausse des revenus, etc.

1.5.3 *Se méfier du sophisme de composition*

Le sophisme de composition nous rappelle que « le tout n'est pas la somme des parties ». En économie, les mouvements macroéconomiques ne sont pas que le résultat de la simple agrégation des mouvements microéconomiques.

¹ C'est toute la complexité des relations causales économiques qui est évoquée ici : pour accroître la production de biens de consommation, les entreprises de ce type vont devoir acheter des biens de production dont la demande peut augmenter même si le taux de marge diminue...

UNE BRÈVE HISTOIRE DE LA PENSÉE ÉCONOMIQUE

« [L'économie politique] ne traite pas de l'ensemble de la nature humaine telle qu'elle est modifiée par l'état social ni de tout le comportement de l'homme en société. Elle ne s'intéresse à lui qu'en tant qu'être animé par le désir de posséder de la richesse et qui est capable de juger de l'efficacité comparée des moyens visant à atteindre cette fin. » - John Stuart Mill, 1831

« Le chasseur et le pêcheur isolés, ces exemplaires uniques d'où partent Smith et Ricardo, font partie des fictions pauvrement imaginées du XVIII^e siècle, de ces robinsonnades qui, n'en déplaie à tels historiens de la civilisation, n'expriment nullement une simple réaction contre des excès de raffinement et un retour à ce qu'on se figure bien à tort comme l'état de nature. » - Karl Marx, 1857

SOMMAIRE

[2.1 Les pères fondateurs de l'économie « classique » \(1776-1871\)](#)

[2.2 La théorie néoclassique \(1871-1936\)](#)

[2.3 La révolution keynésienne ou le capitalisme sous tente à oxygène](#)

[2.4 La rupture monétariste](#)

[2.5 La controverse entre la Nouvelle École classique \(NEC\) et la Nouvelle École keynésienne \(NEK\)](#)

[2.6 La tradition marxiste : la critique du capitalisme](#)

[2.7 Les courants hétérodoxes](#)

[2.8 Jean Tirole : l'économie du bien commun](#)

2.1 LES PÈRES FONDATEURS DE L'ÉCONOMIE « CLASSIQUE » (1776-1871)

2.1.1 *Les problématiques de l'école classique*

Cette première école de pensée regroupe de nombreuses figures qu'on appelle parfois les « pères fondateurs » de l'économie (Adam Smith, David Ricardo, Thomas Robert Malthus, Jean-Baptiste Say, John Stuart Mill). Son ambition est d'affranchir l'économie des jugements de valeur et des questions de philosophie morale et politique. Avec ces économistes, qui se situent entre la fin du XVIII^e siècle et la première moitié du XIX^e siècle, l'économie acquiert réellement un statut de science autonome. Si les théories de l'école classique sont très diverses, elles ont en commun de défendre les principes du libéralisme économique : 1) les mécanismes de marché doivent occuper une place centrale dans l'économie, 2) l'État doit intervenir le moins possible dans les affaires économiques et 3) les processus de production sont inséparables des conditions de répartition des richesses entre les classes sociales, et celles-ci répondent à des lois explicatives qu'il faut tenter de déchiffrer. Malgré les différences d'approche des questions économiques, les classiques ont des préoccupations communes : quelle est l'origine de la richesse ? Quelle est l'origine de la valeur ? Quelle place pour l'État ?

2.1.2 *Adam Smith, le père de l'économie politique classique*

Le fondateur de ce courant « classique » est Adam Smith (1723-1790), qui est également le père de la pensée libérale en économie. On peut résumer sa pensée par quelques grands principes : 1) la richesse des nations provient de la production matérielle permise par le travail, et plus spécialement de la **division du travail**, la spécialisation et

l'échange marchand , 2) il faut laisser les individus s'enrichir, car en recherchant leur intérêt personnel, ils enrichissent non intentionnellement la nation tout entière, et l'intérêt personnel permet d'atteindre l'intérêt général (c'est le principe de la « **main invisible** » et de l'harmonie naturelle des intérêts), 3) il défend résolument la concurrence, car elle permet de se procurer les produits aux meilleurs prix sur les marchés et elle profite aux consommateurs. Il étend cette idée de concurrence bénéfique à l'échelon international : il est possible d'acheter à l'extérieur ce qui y est moins cher. Adam Smith est donc un partisan du **libre-échange** et de la spécialisation internationale de chaque pays dans les productions où il est le plus efficace (où il est nécessaire d'utiliser le moins de travail possible). Le reste doit être importé. 4) Enfin, Adam Smith est aussi le théoricien de l'**ordre naturel** : il faut respecter la liberté individuelle et ne pas perturber l'échange marchand qui est un fait naturel, qui a toujours existé dans les sociétés. L'économie s'équilibre automatiquement et l'**État** ne doit surtout pas intervenir dans son fonctionnement. Son rôle se limite dès lors à trois **fonctions** essentielles, qu'on appelle « **régaliennes** » : l'armée, la justice, la police, mais aussi les travaux d'infrastructures nécessaires pour le développement économique et non rentables pour le secteur privé (creusement de canaux fluviaux, construction de ponts, etc.), c'est-à-dire les biens collectifs.

2.1.3 David Ricardo, le théoricien de l'avantage comparatif

David Ricardo (1772-1823) a prolongé les travaux de Smith et fondé le raisonnement économique sur des bases rigoureuses. On peut retenir de ses analyses plusieurs points importants :

1) **Une théorie de la valeur travail.** Dans son ouvrage *Principes de l'économie politique et de l'impôt*, publié en 1817, David Ricardo écrit que « *toutes les marchandises sont*

le produit du travail et n'auraient aucune valeur sans le travail dépensé à les produire ». Il considère que la valeur d'échange est liée au travail incorporé dans la marchandise par l'activité de production. Il introduit notamment une distinction entre les biens reproductibles dont la valeur dépend du travail et les biens non reproductibles (vin de qualité, œuvre d'art...) dont la valeur dépend de la rareté.

2) Une théorie de la répartition du revenu entre les classes sociales. L'analyse de la répartition du revenu issu de la production entre salaire (rémunération du travail), profit (rémunération du capital) et rente (rémunération de la terre) permet à Ricardo de formuler la théorie de la rente différentielle. Ricardo estime que les salariés sont toujours payés au niveau minimal de subsistance ; leur salaire dépend donc du prix des produits de première nécessité et, notamment, du prix du blé (qui permet de faire le pain). Il explique qu'à mesure que la population augmente, il est nécessaire d'utiliser des terres de moins en moins fertiles pour pouvoir nourrir la population. Comme le prix doit au moins couvrir le coût de production de la dernière terre mise en culture, le prix du blé augmente, ainsi que celui du pain. Et dès lors, le salaire de subsistance qu'il faut bien verser aux ouvriers augmente aussi. En conséquence, les profits baissent tendanciellement selon Ricardo, ce qui devrait conduire l'économie à un état stationnaire, puisqu'aucun entrepreneur ne souhaitera investir du fait de perspectives de profits trop faibles. L'accumulation du capital est stoppée. Ricardo envisage deux possibilités pour contrer cette tendance : le progrès technique, qui permettrait d'utiliser de façon plus efficace les terres disponibles (en utilisant par exemple des engrais pour un meilleur rendement) ou l'ouverture au commerce international, pour mettre en culture des terres fertiles jusque-là inexploitées (et importer du blé).

3) Une défense vigoureuse du libre-échange et une théorie des **avantages comparatifs**. Ricardo explique que, lorsque les facteurs de production sont immobiles entre pays (les travailleurs ne peuvent se déplacer d'un pays à l'autre et les

capitalistes investissent uniquement dans leur pays), chaque pays a intérêt à se spécialiser dans la production du bien pour lequel il est relativement le plus performant et à importer les autres biens. En faisant cela, plus de biens peuvent être produits au niveau mondial, avec une quantité constante de facteurs de production : le libre-échange entre pays est à l'origine de la création d'un gain à l'échange, à l'échelle mondiale.

4) Une réflexion sur les **questions monétaires**. Ricardo établit un lien étroit entre la quantité de monnaie en circulation et l'inflation. Si l'offre de monnaie est trop importante par rapport à la quantité de biens et services en circulation, la monnaie va perdre de sa valeur (il est un **précurseur** de ce que les monétaristes appelleront la « **théorie quantitative de la monnaie** »). Une unité monétaire ne permettra pas d'acheter autant de biens et services que par le passé, ce qui signifie qu'il y a inflation. Afin de limiter le pouvoir de création monétaire des banques et de contrôler l'inflation, Ricardo propose que la monnaie soit définie en référence à l'or.

2.1.4 Deux auteurs classiques marquants : *Jean-Baptiste Say et Thomas R. Malthus*

Jean-Baptiste **Say** (1767-1832) est un autre auteur majeur de l'école classique. Il insiste sur le rôle prépondérant de l'entrepreneur dans l'économie de marché et la primauté de la production *avant* sa répartition puis sa consommation, dans le cadre de sa célèbre « **loi des débouchés** », en vertu de laquelle « les produits s'échangent contre les produits » et l'offre, par les revenus qu'elle génère, crée sa propre demande. Toute production génère des revenus (profits, salaires, intérêts) susceptibles de racheter toute la production de biens et services à sa valeur. Il est donc le précurseur de ce que l'on appelle aujourd'hui, dans le débat macroéconomique, la « politique de l'offre ». Selon sa « loi des débouchés », c'est donc uniquement du niveau de la

production que les débouchés des entreprises découlent. Les crises industrielles sont donc de simples déséquilibres conjoncturels, qui ne remettent pas fondamentalement en cause les capacités autorégulatrices des marchés. Jean-Baptiste Say affirme donc l'impossibilité des crises générales de surproduction et d'engorgement général des marchés (manque de demande), même s'il peut survenir des déséquilibres temporaires et localisés à certains secteurs de l'économie, vite résorbés grâce à la loi de l'offre et de la demande. La **monnaie** est **neutre** dans le système économique et ne sert que d'intermédiaire dans les échanges : les variations de la quantité de monnaie n'affectent pas le niveau de la production réelle.

Au sein de l'école classique, Thomas R. **Malthus** a été un économiste influent de son temps. Il **s'oppose à la loi des débouchés** de J.-B. Say (acceptée par Ricardo) et considère pour sa part que la demande peut être un problème, car une épargne excessive constitue un danger pour l'économie par une insuffisance des débouchés. Dans son *Essai sur le principe de population* (1798), Thomas R. Malthus formule sa « **loi de population** » selon laquelle la population tend à s'accroître selon une progression géométrique alors que les subsistances s'accroissent selon une progression arithmétique : il y a donc une tendance permanente à la surpopulation. Thomas R. Malthus préconise notamment la « *contrainte morale* » (chasteté et mariage tardif) pour limiter le nombre des naissances. Selon lui, la pauvreté découle de l'excès de population, or l'aide aux pauvres leur permet d'avoir plus d'enfants (voire les y incite) et conduit donc à aggraver la situation. C'est la raison pour laquelle il propose l'abrogation des lois sur les pauvres et notamment de l'Acte de Speenhamland (1795).

Même s'il défend la propriété privée et l'économie concurrentielle, **John Stuart Mill (1806-1873)**, que l'on considère souvent comme le dernier des économistes classiques, est conscient des inégalités sociales et refuse d'assimiler le progrès économique au progrès social. Dans

son ouvrage *Principes d'économie politique* de 1848, l'intervention de l'État apparaît pour lui comme moyen de pallier les effets négatifs du marché concurrentiel. Pour J. S. Mill, il faut assurer une vie décente à chacun et donc « *attacher moins d'importance au simple accroissement de la production* ». Pour cela il faut une meilleure allocation des richesses, en mettant en place une législation sur les droits de succession et les donations (réduction des disparités en termes de patrimoine). Il est favorable à la participation des salariés aux résultats de l'entreprise, non seulement pour accroître leurs revenus, mais aussi pour modifier les relations entre maîtres et ouvriers. Le progrès social correspond pour J. S. Mill à un changement dans les rapports sociaux, plus égalitaires y compris pour les rapports hommes/femmes. Pour permettre ce progrès social, J. S. Mill propose une intervention de l'État pour concilier les bienfaits de la concurrence avec la justice sociale.

2.2 LA THÉORIE NÉOCLASSIQUE (1871-1936)

2.2.1 *Ruptures et continuité avec l'école classique*

À la fin du XIX^e siècle, l'école néoclassique, dont les chefs de file sont l'économiste autrichien Carl Menger, l'économiste britannique William Stanley Jevons et l'économiste français Léon Walras, conserve de l'école classique la loi de Say et l'idée centrale selon laquelle le libre jeu des marchés et les forces de la concurrence permettent de résorber les déséquilibres. Dans cette approche fondée sur l'existence d'un équilibre général de tous les marchés, la crise est toujours, en théorie, impossible : les déséquilibres constatés ne peuvent survenir qu'en raison de perturbations extérieures au système (exogènes). Ces perturbations sont le plus souvent imputées à l'intervention de l'État, qui introduit, par son action, des rigidités sur les marchés et qui, en contrôlant

la création de la monnaie, perturbe les signaux envoyés aux agents par les prix de marché. Les théoriciens de l'école néoclassique prolongent ainsi les travaux des premiers auteurs classiques en leur donnant des fondements plus robustes, grâce en particulier à l'usage de la formalisation mathématique.

2.2.2 *L'analyse marginaliste*

Les néoclassiques utilisent un **raisonnement microéconomique** fondé sur le comportement d'un agent représentatif **rationnel** qui cherche, par les arbitrages qu'il effectue, à maximiser sa satisfaction sous contrainte, et en expliquant le fonctionnement général de l'économie de marché par l'agrégation des comportements de tous les individus qui la composent. Ils s'intéressent plus particulièrement aux comportements du consommateur et du producteur en tant qu'agents théoriques et génériques : ainsi, le consommateur cherche à acheter telle ou telle marchandise en cherchant à **maximiser** sa satisfaction **sous contrainte** de son budget (théorie du consommateur), tandis que le producteur choisit entre fabriquer telle ou telle marchandise afin de maximiser son profit sous contrainte de son coût de production (théorie du producteur). Les théoriciens néoclassiques utilisent le raisonnement à la marge (on évoque parfois la « **révolution marginaliste** ») : ils ne raisonnent pas sur des quantités globales ou moyennes, mais sur des unités marginales, c'est-à-dire la dernière unité. La théorie du consommateur se base ainsi sur l'utilité (la satisfaction) apportée par la dernière unité consommée (utilité marginale), tandis que la théorie du producteur est fondée sur la production tirée par la dernière unité de facteur de production (productivité marginale). Ce raisonnement à la marge permet aux auteurs néoclassiques de formaliser mathématiquement leur raisonnement, grâce à l'utilisation des dérivées partielles. Comme les auteurs classiques, ils ont cherché à déterminer ce qui procure sa valeur à un bien, et considèrent que la valeur d'un bien

dépend de l'utilité que le bien offre : le prix d'un bien est alors fonction du désir du consommateur de l'acquérir. Cette utilité est liée à la rareté du bien et constitue une notion subjective, qui dépend des préférences de l'individu.

2.2.3 Une théorie des marchés

Les théoriciens néoclassiques ont également approfondi la **théorie des marchés** : le français Léon **Walras**, premier professeur de la chaire d'économie politique de l'université de Lausanne, a démontré mathématiquement, sous la condition du respect des conditions de la concurrence pure et parfaite, l'existence d'un équilibre sur tous les marchés en même temps (marché des biens et services, marché du capital, marché du travail et marché de la monnaie). Dans ce cas, tous les facteurs de production sont utilisés à leurs capacités maximales, le plein emploi est assuré et tous les consommateurs trouvent satisfaction à leurs demandes. Le modèle théorique de l'**équilibre général** de Walras, qui demeure aujourd'hui encore une référence, formalise la métaphore d'Adam Smith selon laquelle une « main invisible » (celle du marché) conduit à l'harmonie naturelle des intérêts.

À la différence de L. Walras, Alfred **Marshall** à l'université de Cambridge raisonne quant à lui en **équilibre partiel** : il étudie la détermination de l'équilibre sur un marché particulier. Cela le conduit à mettre en avant le raisonnement **ceteris paribus**, « toutes choses égales par ailleurs » : il étudie l'impact d'une modification d'une variable sur un marché sans tenir compte de son effet sur les autres marchés.

Le successeur de Walras à l'école de Lausanne, Vilfredo **Pareto**, complètera la théorie de l'équilibre général par la théorie de l'optimum : il renonce à la mesurabilité des satisfactions et construit des **courbes d'indifférence**. Le concept d'optimum découle du refus d'introduire toute mesure et, par conséquent, de l'impossible comparaison des utilités intertemporelles. Dès lors, une situation est optimale au sens

de Pareto lorsqu'il n'est plus possible d'améliorer la satisfaction d'un individu sans détériorer celle d'un autre.

Cependant, l'introduction d'un « **commissaire-priseur** » ou « secrétaire de marché » se révélera indispensable au bon fonctionnement des marchés, selon la théorie des économies de marché développée par Kenneth **Arrow** et Gerard **Debreu** en 1954.

2.2.4 *Les limites de l'intervention publique*

Tout en reconnaissant la présence possible d'imperfections de concurrence sur les marchés (existence de monopoles, d'oligopoles, problèmes d'information), les auteurs néoclassiques pointent surtout les effets néfastes de l'excès de réglementation, qui freine la création d'entreprises, alors que les prix administrés empêchent la loi de l'offre et de la demande d'opérer et de maximiser la satisfaction des agents en créant des situations de rationnement. Par ailleurs, **l'intervention de l'État perturbe le fonctionnement optimal des marchés et les signaux envoyés par les prix**, afin d'orienter les décisions de production vers les secteurs les plus rentables. Sur le marché du travail, le monopole des syndicats (souvent associés dans l'analyse à l'État) conduit à un niveau trop élevé des salaires réels et du coût du travail, ce qui décourage les entrepreneurs de créer des emplois et génère du chômage, tandis que le droit du travail trop protecteur agit de même, en empêchant les entreprises d'adapter leurs effectifs aux fluctuations de la demande (insuffisante flexibilité du travail). Pour les néoclassiques, le marché est « autorégulateur » et constitue le mode de coordination le plus efficace en matière d'allocation des ressources dans l'économie. La manipulation de la création monétaire est d'ailleurs souvent vue par les auteurs néoclassiques comme le point d'entrée des influences perturbatrices de l'État sur l'activité économique. Avec les théoriciens néoclassiques, l'échange devient l'objet d'étude principal de la science économique et le marché l'institution

centrale de l'économie, ce qui suppose une intervention minimale de l'État dans les mécanismes concurrentiels.

2.3 LA RÉVOLUTION KEYNÉSIENNE OU LE CAPITALISME SOUS TENTE À OXYGÈNE

Avec Keynes, nous assistons à l'avènement de la macroéconomie en tant que champ disciplinaire spécifique. Pour mieux comprendre les débats entre économistes depuis Keynes, on peut s'appuyer sur les grands domaines d'accords et de désaccords en macroéconomie.

Ceux-ci concernent :

- la vitesse d'ajustement des marchés
- le niveau de flexibilité des prix sur ces marchés
- la conception de l'équilibre (plein emploi, sous-emploi, stable...)
- la formation des anticipations (extrapolatives, adaptatives, rationnelles)
- le cadre temporel de référence (courte période, longue période)
- la source majeure d'instabilité (monétaire, réelle)
- le type de politique économique proposé.

Ces principaux domaines de désaccords, d'accords, ne sont pas exclusifs les uns des autres. Ils sont plutôt liés, et il est difficile d'établir une hiérarchie entre eux.

2.3.1 *La critique de la loi des débouchés*

Si la théorie néoclassique domine largement la fin du XIX^e siècle et le début du XX^e siècle, en édifiant une science économique moderne, la grande crise des années 1930, qui dégénère en une dépression économique d'une gravité et d'une durée jamais observées jusque-là, fragilise la légitimité du capitalisme libéral, à une époque où le modèle socialiste d'économie planifiée exerce une certaine attraction dans le monde. Face à cette dépression, et déflation, unique et

exemplaire, alors que les économistes néoclassiques misent toujours sur les mécanismes autorégulateurs du marché pour résorber les excès de l'offre (baisse des salaires pour réduire le chômage, baisse des prix pour relancer la consommation, baisse des taux d'intérêt pour relancer l'investissement), John Maynard **Keynes** (1883-1946) va contester l'idée d'une capacité naturelle du système économique à retrouver la prospérité à long terme après d'inévitables phases d'ajustement et critiquer l'idéologie libérale du « laisser-faire » qui en découle. Dans son ouvrage majeur, *Théorie générale de l'emploi, de l'intérêt et de la monnaie*, publié en 1936, il prouve que la crise du capitalisme est *possible*, car rien ne garantit que toute offre crée *ipso facto* et à tout moment sa propre demande (critique de la loi des débouchés de Say). Il insiste sur le caractère psychologique des déterminants principaux de l'activité économique (consommation, demande de monnaie, taux d'intérêt, décision d'investissement), qui rend le capitalisme libéral intrinsèquement instable.

2.3.2 L'équilibre de sous-emploi et le principe de la demande effective

Contrairement à la théorie néoclassique, qui postule le plein emploi, Keynes montre qu'il peut exister des **équilibres de sous-emploi** durables en raison d'une **demande effective** (demande globale anticipée par les entrepreneurs) insuffisante. En situation d'**incertitude** (au sens de Franck Knight, le risque caractérise une situation probabilisable, calculable. En revanche, en situation d'incertitude, rien ne peut être probabilisé. Tout simplement, on ne sait pas !) face à l'avenir, les entrepreneurs peuvent en effet mettre en œuvre un niveau d'investissement, de production et d'emploi insuffisant pour occuper toute la population active et utiliser toutes les capacités de production disponibles, même s'ils maximisent leur profit. La baisse des salaires pour résorber le chômage serait même pire que tout. En effet, *ceteris paribus*, si une entreprise peut attendre individuellement, par ce biais, une baisse de ses coûts de production, en raison des

interdépendances, la diminution de la demande effective globale deviendrait considérable si ce calcul se généralisait à l'ensemble des entreprises (de plus, la baisse des salaires entraînerait sûrement une baisse des prix, laissant inchangés les salaires réels). Le chômage n'est donc pas « volontaire » au sens des théoriciens néoclassiques (les travailleurs refusant le salaire d'équilibre jugé trop faible), mais il est bien de nature « involontaire » en raison d'une demande effective insuffisante. Par ailleurs, la dynamique du capitalisme permet une croissance du revenu réel, mais, en raison de la **loi psychologique fondamentale**, la consommation progresse moins vite que la hausse du revenu : la hausse tendancielle de l'épargne qui en résulte peut alors créer une contrainte de débouchés. C'est donc l'excès d'épargne (comme l'avait déjà analysé Thomas Robert Malthus) qui est à l'origine du risque de sous-consommation et constitue le talon d'Achille du capitalisme, et non son insuffisance. Si, dans la vision keynésienne, la crise du capitalisme est possible, elle n'est pas pour autant *utile*, car elle provoque une perte de capacité, de potentiel, de production (travailleurs au chômage, disparition d'entreprises) et une perte nette de revenus pour la société.

2.3.3 *L'indispensable action stabilisatrice de l'État*

Pour Keynes, l'intervention de l'État se justifie en permanence pour contrer les tendances perturbatrices de l'économie de marché. « *À long terme nous serons tous morts* », disait Keynes, donc nous n'avons pas le temps d'attendre que les marchés s'équilibrent... Si les pouvoirs publics peuvent agir pour « **réamorcer la pompe** » et **briser le cercle vicieux de la crise**, leur tâche est d'assurer en amont une régulation constante de l'activité pour éviter que la dépression ne survienne. Pour cela, l'État doit agir sur la variable motrice, mais très instable de l'investissement productif, dont les variations déterminent largement les fluctuations économiques. Celui-ci dépend de la rentabilité espérée par

les entrepreneurs en situation d'incertitude et du niveau des taux d'intérêt, variable liée à l'équilibre entre l'offre de monnaie (fixée par la banque centrale) et la demande de monnaie des agents économiques et non de l'offre et de la demande de capitaux, comme pour les néoclassiques. Ainsi, pour les néoclassiques, le taux d'intérêt (réel) détermine le partage entre consommation et épargne, alors que pour les keynésiens, le taux d'intérêt (nominal) détermine le partage de l'épargne entre monnaie et titres.

Les autorités monétaires peuvent donc réduire le taux d'intérêt pour stimuler l'investissement, à même de garantir une hausse plus que proportionnelle de la production (**effet multiplicateur**).

Mais c'est surtout la politique budgétaire qui apparaît la plus efficace lorsqu'elle utilise le déficit budgétaire de manière à soutenir l'investissement public (effet multiplicateur plus puissant), tandis que la redistribution des richesses en faveur des bas revenus (qui consomment la plus grande part de leur revenu disponible) peut éventuellement stimuler la propension à consommer. En conséquence, selon Keynes, le capitalisme doit être *organisé* pour survivre et parvenir à concilier croissance économique, plein emploi et équité dans la répartition des revenus.

2.4 LA RUPTURE MONÉTARISTE

2.4.1 *La crise de la pensée keynésienne*

Après 1945, la macroéconomie keynésienne centrée sur la régulation conjoncturelle à court terme s'impose dans le champ scientifique, mais aussi dans la pratique de la politique économique, pilotée par des gouvernements soucieux de poursuivre l'expansion économique (croissance régulière du PIB) dans la stabilité macroéconomique (contrôle de l'inflation, maintien du pouvoir d'achat de la monnaie). Cela semble bien fonctionner pendant plusieurs décennies et on parlera de *fine-tuning*, de « **réglage fin** », par les politiques

conjoncturelles. Mais les années 1970 marquent pourtant la crise de la pensée keynésienne, face à l'émergence concomitante d'une croissance faible, d'un chômage de masse et de l'accélération de l'inflation (que l'on a appelée à l'époque la « **stagflation** »). La relation/courbe de Philips et l'arbitrage inflation/chômage semblent révolus. Les politiques traditionnelles de relance à court terme butent sur les causes structurelles de la crise (hausse des coûts de production, ralentissement des gains de productivité, chômage persistant) et les effets de la mondialisation (contrainte extérieure).

2.4.2 *L'analyse néo-quantitativiste*

Milton **Friedman** (1912-2006), le chef de file de l'école monétariste, a entrepris dès les années 1950 à l'université de Chicago une critique des thèses keynésiennes dans le but de dénoncer les formes perverses de l'intervention de l'État dans l'économie. Dans ses travaux, Milton Friedman réhabilite la **théorie quantitative de la monnaie**, selon laquelle le niveau général des prix varie proportionnellement à la quantité de monnaie en circulation dans l'économie, et démontre, dans une vaste étude de l'histoire monétaire des États-Unis menée avec Anna Schwartz (1963), l'impact de la politique monétaire sur l'activité économique : entre 1929 et 1933, une diminution de la quantité de monnaie de plus de 25 % explique alors la sévérité de la Grande Dépression et montre le rôle déterminant des banques centrales dans la crise. À partir des années 1970, l'accélération de l'inflation dérègle le cercle vertueux expansionniste des « trente glorieuses » et invalide la théorie keynésienne de l'époque concernant le dilemme inflation/chômage, en vertu de laquelle une relance monétaire appropriée permet de réduire le chômage. Selon les monétaristes, l'inflation a essentiellement pour origine un taux d'expansion monétaire trop élevé par rapport au taux de croissance réel de l'économie, tandis qu'il existe un taux de chômage naturel qui ne dépend que des imperfections du marché du travail, et dont on ne peut s'écarter durablement par une injection de monnaie.

2.4.3 *L'inefficacité à long terme de la relance keynésienne*

Pour Milton Friedman, **les perturbations de l'économie de marché prennent naissance dans le secteur monétaire** et sont causées par les interventions intempestives des banques centrales à court terme, qui dérèglent le fonctionnement des marchés. Puisque la demande de monnaie des agents économiques est relativement stable et prévisible, les variations erratiques de l'offre de monnaie par la banque centrale sont responsables de l'érosion continue du pouvoir d'achat de la monnaie, au détriment des titulaires de revenus fixes. Selon Milton Friedman, la meilleure stratégie consiste, compte tenu des **anticipations adaptatives**, à neutraliser la politique monétaire en instaurant une norme de croissance de la masse monétaire (règle de taux constant de $k\%$) et l'indépendance des banques centrales.

En résumé, l'approche monétariste orthodoxe a permis un progrès à la fois théorique et empirique entre le milieu des années 1950 et le milieu des années 1970.

En effet, la théorie quantitative de la monnaie (TQM) fut reformulée par M. Friedman, les anticipations furent intégrées à la courbe de Philips par Friedman et Phelps, et une théorie monétaire de la balance des paiements fut élaborée. On admet que ces théories ont été longuement corroborées par les recherches empiriques.

2.5 LA CONTROVERSE ENTRE LA NOUVELLE ÉCOLE CLASSIQUE (NEC) ET LA NOUVELLE ÉCOLE KEYNÉSIENNE (NEK)

2.5.1 *La Nouvelle École classique (NEC)*

Dans les années 1970, le courant de la Nouvelle École classique (NEC) va prolonger et radicaliser les préceptes monétaristes : dans leurs travaux, ces auteurs (Robert Lucas, Thomas Sargent, Neil Wallace) critiquent les politiques

keynésiennes jugées totalement inefficaces, même à court terme. Pour eux, les marchés sont continuellement en équilibre et les agents économiques sont capables d'anticiper correctement les conséquences néfastes des politiques keynésiennes (en vertu de « l'hypothèse des **anticipations rationnelles** »). Par exemple, si l'État veut relancer l'économie par un déficit budgétaire, les ménages vont anticiper les hausses futures d'impôts qui ne vont pas tarder à survenir pour combler le déficit. En conséquence, la hausse des revenus distribués n'augmente pas la consommation, mais seulement l'épargne de précaution. La Nouvelle Économie classique préconise une politique économique neutre qui ne perturbe pas les marchés, censés s'équilibrer spontanément. Il n'existe pas de risque d'insuffisance de demande dans l'économie et il faut donc agir sur l'offre (et inciter les entreprises à produire, investir, innover, etc.).

Cette école doit être différenciée de l'école monétariste orthodoxe même si l'explication de l'inflation reste monétariste. Les contributions de Lucas, Sargent, Wallace... ont dominé les débats macroéconomiques des années 1970, surtout aux États-Unis. Mais dès la fin de cette décennie, plusieurs limites inhérentes à cette approche en termes d'équilibre sont devenues évidentes. La NEC ou NMEC (M pour macroéconomie) repose sur l'acceptation de trois hypothèses centrales, chacune d'elle pouvant être utilisée et évaluée pour elle-même. Ainsi, certains partisans de cette école n'adhèrent pas à chacune de ces trois hypothèses qui sont :

- l'hypothèse d'anticipations rationnelles (HAR)
- l'hypothèse d'ajustement continu des marchés
- l'hypothèse d'offre globale (HOG).

La NEC propose :

- une théorie de l'inefficacité de la politique monétaire²
- une théorie des coûts réels de la désinflation (il existe un ratio de sacrifice, volume de production perdu lorsqu'on cherche à diminuer l'inflation)

– une théorie du problème de la cohérence temporelle des politiques monétaires.

Dans les années 1980, le développement de la **NEC** a donné lieu à une réinterprétation des fluctuations économiques, fondée sur les conséquences de **chocs exogènes** aléatoires, par nature imprévisibles, sur les comportements d'agents optimisateurs et rationnels.

Selon la « **théorie des cycles réels** », les agents ont des anticipations rationnelles, ils ne peuvent donc pas être trompés par la politique monétaire et ils raisonnent toujours en termes réels. L'équilibre est alors toujours réalisé sur tous les marchés, il résulte des arbitrages intertemporels des agents entre travail et loisirs, consommation actuelle et future.

Selon les travaux de **C. Prescott** (né en 1940) et **F. Kydland** (né en 1943) (tous deux Prix Nobel 2004), les fluctuations proviennent de perturbations exogènes imprévisibles, causées par un **choc technologique**, c'est-à-dire une variation rapide de la technologie disponible, et donc de la productivité, due à l'introduction d'une innovation par exemple. Les agents modifient alors leurs arbitrages pour tenir compte des nouvelles conditions, ce qui entraîne une offre de travail différente, et un nouvel équilibre.

Plusieurs implications de cette analyse doivent être soulignées : d'une part, l'économie est fluctuante par nature, puisque les fluctuations viennent de l'adaptation rationnelle à des chocs extérieurs. D'autre part, plus un choc technologique est temporaire, plus les fluctuations qu'il entraîne sont importantes, car les agents doivent bouleverser leurs choix pour en tirer avantage (par exemple, une variation ponctuelle du prix du travail doit inciter un agent rationnel à modifier radicalement son offre de travail pour maximiser l'avantage qui en découle). Enfin, les cycles sont des cycles à l'équilibre ; les choix des agents sont à tout moment du cycle optimisateurs et assurent l'équilibre.

Ce modèle conduit les nouveaux classiques à souligner le caractère néfaste des politiques économiques contra-

cycliques, qui viennent perturber des choix rationnels (par exemple une hausse temporaire du chômage volontaire) en faussant les signaux-prix, ce qui conduit à un équilibre sous-optimal. Dans ce cadre, l'augmentation du chômage conjoncturel serait le résultat d'une adaptation rationnelle aux chocs technologiques.

2.5.2 *La Nouvelle Économie keynésienne (NEK)*

La réplique des économistes de la Nouvelle Économie keynésienne (NEK), comme Edmund Phelps, Joseph Stiglitz, Paul Krugman, Gregory Mankiw, durant les années 1980, va chercher à **donner à la macroéconomie keynésienne des fondements microéconomiques** issus de la théorie néoclassique. Les néokeynésiens considèrent que les forces du marché n'assurent pas l'optimum, mais leurs travaux s'appuient sur des modèles théoriques qui reprennent de nombreux outils de la théorie microéconomique néoclassique. Ce courant de pensée conserve deux conclusions fondamentales de la théorie keynésienne : l'imperfection des marchés, qui suppose l'intervention indispensable de l'État, et l'existence d'un chômage involontaire. Les travaux de la NEK reposent sur quelques grands principes : la **rigidité des prix** sur des marchés imparfaitement concurrentiels, qui implique qu'il n'y a pas d'ajustement instantané de ces derniers, contrairement aux enseignements de la théorie néoclassique ; la **rigidité des salaires** sur le marché du travail et le chômage volontaire, puisque les salaires sont rigides, car les entreprises rationnelles peuvent avoir intérêt à rémunérer les salariés au-delà du salaire d'équilibre (théorie du salaire d'efficience, modèles *insiders/outsiders*) ; l'efficacité de la politique monétaire : contrairement aux néoclassiques, les néokeynésiens estiment que la monnaie n'est pas neutre dans l'économie et qu'elle est un déterminant essentiel de la politique économique en cas d'insuffisance de la demande.

La NMEK (M pour macroéconomie) est un ensemble d'économistes très hétérogène. Pour autant, on peut définir cette école à partir des réponses qu'elle apporte aux questions sur la neutralité monétaire et sur les conséquences de l'imperfection réelle des marchés sur le cycle économique. Pour les tenants de cette école, la non-neutralité de la monnaie résulte de la lenteur de l'ajustement des prix et des salaires qui serait à l'origine des conséquences sur les fluctuations réelles de l'activité économique. Afin de justifier ces thèses, les keynésiens de la NMEK se sont particulièrement attachés à fonder théoriquement ces thèses en recherchant les fondements microéconomiques de ces rigidités et imperfections marché.

2.6 LA TRADITION MARXISTE : LA CRITIQUE DU CAPITALISME

2.6.1 *Les contradictions internes du capitalisme*

Selon Marx, le développement exceptionnel des forces productives du travail (innovation, division du travail, organisation de la production) est en effet le moteur de la croissance (depuis la révolution industrielle), mais **le système capitaliste est miné par une contradiction interne** qui rend la loi des débouchés inopérante. En effet, dans l'économie capitaliste, la production et la consommation constituent deux opérations disjointes et la propriété privée du capital contraste avec le caractère social de la production, à une époque où le salariat est en pleine expansion. Ainsi, les capitalistes ne sont pas assurés de rencontrer une demande solvable suffisante sur les marchés. Les **crises de surproduction** constituent donc une menace permanente, encore aggravée par la paupérisation relative de la classe ouvrière. En effet, la concurrence entre les capitalistes les incite à substituer sans cesse du capital technique (biens d'équipement, machines) à la force de travail, afin de réaliser des gains de productivité et de réduire le temps de rotation du capital.

Selon les descendants de Karl Marx (1818-1883), le capitalisme est un système économique historiquement condamné par ses contradictions internes : la crise est techniquement possible du fait des caractéristiques du système (dissociation de l'acte de vente et de l'acte d'achat, anarchie des décisions d'investissement, qui pousse à la suraccumulation du capital, manque de coordination générale de l'activité), mais elle est surtout une fatalité qui manifeste l'inefficacité croissante de ce système et l'épuisement de son rôle historique dans le développement des sociétés.

2.6.2 La théorie de l'exploitation et l'actualité du message marxiste

Selon Marx, le profit des capitalistes (la plus-value) résulte d'un prélèvement sur le seul produit du travail, et l'augmentation de l'intensité capitaliste réduit les moyens d'extraire la plus-value, pourtant indispensable au prolongement de la croissance capitaliste, et elle provoque une **baisse tendancielle du taux de profit**. Cette baisse expliquerait à la fois les crises périodiques du capitalisme et sa disparition finale, favorisée sur le plan politique et social par la lutte des classes et l'organisation des travailleurs face aux inégalités et à l'extension de la pauvreté. Si la crise des années 1930 a pu apparaître comme la crise finale du capitalisme chez les économistes marxistes qui en ont été les témoins, la phase de croissance exceptionnelle qui a suivi la Seconde Guerre mondiale a montré la grande résistance du système capitaliste. Cependant, la crise de la fin des années 2000 a relancé le débat, d'autant plus qu'elle présente des similitudes importantes avec la Grande Dépression de l'entre-deux-guerres. En outre, les institutions nées dans la lutte contre cette crise historique ont permis d'atténuer fortement, pour l'instant, la crise présente, ce qui suggère le rôle créateur de la crise.

2.7 LES COURANTS HÉTÉRODOXES

La pensée économique et la pensée macroéconomique moderne ne peuvent être limitées à cette succession d'écoles orthodoxes, dont les développements théoriques et les visions du monde, qui étaient sous-jacentes à ces derniers, ont dominé cette pensée et ont servi les fondements aux politiques macroéconomiques mises en œuvre depuis la Seconde Guerre mondiale. En effet, deux courants hétérodoxes, minoritaires, ont élaboré des théories qui, bien qu'étant sur certains points divergentes des points de vue orthodoxes, sont en accord avec les écoles orthodoxes sur d'autres aspects. De surcroît, ces développements qui émanent de l'école autrichienne et de l'école postkeynésienne permettent de poser de nouvelles questions et d'apporter de nouvelles réponses à des problèmes non encore définitivement résolus (mais est-ce possible ?). Il convient donc de présenter les principales caractéristiques de ces deux écoles, non négligeables en macroéconomie.

2.7.1 *L'école autrichienne*

– Le rôle de la monnaie, du temps, et la théorie du capital

À l'origine, la théorie autrichienne des fluctuations s'inspire des travaux de C. Menger, E. Von Böhm-Bawerk et du Suédois K. Wicksell. Le modèle de base proposé par L. Von Mises et approfondi par F. Von Hayek insiste sur **le rôle joué par la monnaie et le temps dans les fluctuations économiques**. Dans une économie monétaire, la prise en compte du facteur temporel permet d'insister sur la dynamique de l'activité économique et de mettre en lumière l'importance de l'incertitude. Dans ce cadre théorique où, donc, la monnaie et le temps sont à l'origine de l'**incertitude**, les Autrichiens ont pu démontrer l'absence de neutralité monétaire. Selon Von Mises et Hayek, toute politique monétaire active induit une modification des prix relatifs des biens de production par rapport aux biens de consommation et une variation de la structure temporelle de la production (variation du détour de production via l'investissement). Or,

cette évolution de la **structure temporelle de la production**, comme elle n'est pas le résultat des décisions individuelles des agents économiques (consommateurs et producteurs), conduit inéluctablement à la récession, puis à la dépression. À l'inflation doit succéder le chômage. Pour Hayek, **la monnaie n'est pas neutre** et une émission excessive de monnaie conduit à un allongement injustifié du détour de production (surinvestissement) et, à terme, un retour brutal à un stock de capital compatible avec les arbitrages inter-temporels des agents. Il s'oppose à toute intervention de l'État, c'est-à-dire à toute tentative d'un pilotage conscient de l'économie : l'ordre social est un ordre spontané qui ne doit pas être perturbé. En outre, il est préférable de laisser la dépression suivre son cours, de laisser l'économie se purger elle-même, car, dans le cas contraire, le risque est grand d'accentuer l'instabilité et d'allonger la période d'assainissement. L'analyse autrichienne permet de réconcilier théorie monétaire et théorie de la production en prenant en compte les effets des impulsions monétaires sur la structure des prix relatifs et la structure du capital productif. Ce modèle est considéré comme l'unique solution de recours, pendant l'entre-deux-guerres, à l'analyse keynésienne de la Grande Dépression, car il permet notamment de prévoir le retournement de l'économie américaine.

En fait, l'analyse autrichienne permet de réconcilier la théorie monétaire et la théorie de la production (cf. *Effets des impulsions monétaires sur la structure des prix relatifs et la structure du capital productif*). En d'autres termes, la théorie autrichienne propose implicitement de réconcilier théorie du cycle économique relative à la courte période, et théorie de la croissance économique relative à la longue période.

– **Le théoricien de l'innovation : Joseph Schumpeter (1883-1950)**

C'est à Joseph Schumpeter que l'analyse économique reconnaît traditionnellement d'avoir donné une consistance à la fonction de l'entrepreneur, en insistant sur son rôle dans l'innovation. Schumpeter partage avec Karl Marx un intérêt

pour **la dynamique du système capitaliste**. Cependant, au lieu d'en trouver la source dans les luttes de classes découlant des rapports de production, il l'identifie à une seule force : l'initiative individuelle de l'entrepreneur, dont l'étude est au cœur de son ouvrage *Théorie de l'évolution économique*, dès 1911. Cette évolution économique ne peut logiquement se produire que lorsque les routines sont brisées par l'action innovatrice qui définit ainsi l'**entrepreneur** schumpétérien. Cet entrepreneur, ou plutôt ces « troupes d'entrepreneurs », constitue ainsi la force et le moteur du changement économique. Les entrepreneurs affrontent des résistances, voire des oppositions aux changements, à la nouveauté et c'est à leurs qualités uniques (« *joie de créer* », « *désir de vaincre* », selon Schumpeter) qu'ils doivent d'en triompher. L'action innovatrice de l'entrepreneur reste cependant soumise, pour Schumpeter, à l'existence d'un système de crédit. Cela fait du banquier un autre acteur essentiel, puisque c'est lui qui maîtrise le processus de création monétaire. Schumpeter considère donc qu'avant d'être entrepreneur, l'individu est un débiteur et que c'est grâce à la création monétaire, c'est-à-dire la création d'un pouvoir d'achat à partir de rien, qui affranchit le financement de l'innovation de la contrainte d'une épargne préalable, que celui-ci peut bouleverser les conditions de la croissance en forçant l'économie à orienter ses ressources vers de « **nouvelles combinaisons** » (expression qu'utilise Schumpeter pour désigner les *innovations*).

Joseph Schumpeter défend l'idée que les **fluctuations économiques** sont **caractéristiques de la dynamique du capitalisme**, et que **la crise fait partie intégrante du processus de croissance**. Il développe une analyse des cycles économiques et de l'évolution économique impulsée par l'investissement, et surtout par l'innovation. Les innovations, portées par les entrepreneurs, apparaissent « en grappes » et se généralisent dans le système économique. Les entreprises innovantes en position de monopole dégagent ainsi temporairement un surprofit. Les secteurs de

l'économie liés aux entreprises incorporant le progrès technique sont incités à développer à leur tour de nouveaux produits, de nouveaux procédés, de nouvelles formes d'organisation des entreprises, à conquérir de nouveaux marchés ou de nouvelles sources de matières premières. Mais c'est la généralisation des innovations et l'imitation de celles-ci par d'autres entrepreneurs qui détériorent les profits et conduisent à l'essoufflement de l'activité économique. Joseph Schumpeter montre que l'apparition des innovations est un processus irrégulier dans le temps, ce qui explique les fluctuations de l'activité et la marche perpétuellement déséquilibrée de l'économie, tout au long des cycles économiques. Le **processus de croissance** dépend donc d'un **processus d'innovation** qui est un **processus de « destruction créatrice »**, d'entreprises, d'emplois, de revenus... *In fine*, à l'instar de la prophétie de Marx, Schumpeter pense que le capitalisme est condamné à sa perte, mais non du fait de ses contradictions internes et de ses échecs... mais de ses succès !

2.7.2 L'école post-keynésienne

Les origines de cette école se trouvent chez Marshall et Keynes pour la perspective monétaire et la prise en compte du temps, et chez Marx, Kalecki, Sraffa pour l'analyse en termes de répartition du produit.. Les analyses post-keynésiennes visent à généraliser le message contenu dans la *Théorie générale* aux concepts de croissance, accumulation et répartition. De plus les différentes origines de la tradition post-keynésienne ainsi que les multiples tentatives de généralisation de la théorie de Keynes à des domaines qu'il n'avait pas explorés explicitement ont contribué au développement de deux grands courants au sein de cette école, le fondamentalisme keynésien et l'école du surplus. Cependant, on peut mettre en évidence les thèmes communs étudiés de manière privilégiée par les auteurs post-keynésiens : le temps et la monnaie, l'insuffisance de la demande effective, les fondements microéconomiques de la

macroéconomie, l'incertitude (liée au temps et à la source) et une critique des anticipations rationnelles.

La monnaie, le temps et donc l'incertitude sont des caractéristiques majeures d'une économie monétaire de production. En distinguant risque et incertitude, selon Minsky, on met en exergue l'instabilité des comportements économiques, d'où l'adoption de conventions, contrats (Davidson) qui instaurent des rigidités, et on ne peut que rejeter l'HAR, hypothèse qui implique un certain déterminisme sur le futur, ce qui est incompatible avec une situation d'incertitude. Il en découle une explication de l'insuffisance de la demande effective qui est influencée par la détention de monnaie (liquidité, par excellence et préférence pour la liquidité) dans une économie monétaire de production frappée d'incertitude.

La monnaie est en outre considérée comme étant endogène au système. En d'autres termes, par le biais du crédit, l'offre de monnaie va s'adapter aux besoins de l'économie. Cette position concernant l'offre de monnaie a des conséquences fondamentales sur l'instabilité des fluctuations, sur l'inflation et sur les politiques économiques.

En effet, du fait de l'endogénéité de la monnaie, l'excès du crédit peut provoquer à terme un retournement brutal de l'activité conduisant à une dépression aux conséquences non négligeables (cf. Minsky).

De plus, en ce qui concerne l'inflation, l'offre de monnaie n'en est plus la cause, mais la conséquence. Les post-keynésiens l'expliquent par les conflits qui portent sur la répartition de revenus. Conflits qui sont le reflet d'un rapport de force qui peut être exacerbé par la politique monétaire. Dans ces conditions, la croissance endogène de l'offre de la monnaie permet de valider des hausses de salaires et de prix et une politique de revenus adaptée peut favoriser la stabilité des prix lorsqu'elle est indexée sur l'évolution des prix.

Il en résulte que selon les tenants de l'école post-keynésienne, dans une économie monétaire où l'incertitude

règne, les politiques de gestion de la demande (politique budgétaire, politique monétaire) accompagnées d'une politique des revenus adaptée permettront une stabilisation de l'activité économique.

Autrement dit, dans la perspective post-keynésienne, le cycle économique, s'il a des origines réelles (variations de la demande effective, conflits sur la répartition) a également des origines monétaires, mais les variations de l'offre de monnaie endogènes sont elles-mêmes fonction des variables réelles. Le cycle économique est donc fondamentalement lié aux évolutions de variables réelles. Pour autant, dans une économie monétaire, on ne peut négliger les changements monétaires et leurs conséquences. Tel est, selon les post-keynésiens, le message central de Keynes.

2.7.3 La théorie de la régulation

La théorie de la régulation (ou école de la régulation) s'est développée en France puis au niveau international à partir des années 1970. Elle articule les enseignements des traditions marxiste, schumpétérienne, keynésienne et institutionnaliste. Les contributions de ce courant, avec des économistes comme M. Aglietta, R. Boyer, A. Lipietz, portent sur la dynamique du capitalisme, la croissance et les crises, la monnaie et la finance, et les relations économiques internationales. Pour R. Boyer, la régulation est une « *conjonction de mécanismes concourant à la reproduction d'ensemble compte tenu des structures économiques et des formes sociales en vigueur* » (Boyer, 1986). **L'histoire du capitalisme est interprétée comme la succession de modes de régulation** (concurrentiel puis monopoliste) qui entrent tour à tour en crise. Un mode de régulation est la combinaison d'un régime d'accumulation et de formes institutionnelles qui codifient les rapports sociaux. Cette perspective permet d'articuler une réflexion sur la reproduction (continuité) et sur le changement (ruptures). Conçue au départ pour rendre compte de l'évolution à long terme des économies capitalistes industrialisées, cette

problématique a été appliquée aux pays en voie de développement et aux économies de type soviétique.

2.7.4 L'économie institutionnelle

Selon B. Chavance, dans son ouvrage *L'économie institutionnelle* (La Découverte, 2012), l'économie institutionnelle regroupe différents courants hétérodoxes de l'analyse économique qui, en dépit de leurs divergences, ont en commun de considérer que les institutions constituent un objet essentiel de l'analyse économique. Ce point de vue s'oppose à celui d'autres auteurs qui considèrent que l'analyse des institutions relève d'autres sciences sociales que l'économie (psychologie sociale, science politique, sociologie). Ces courants ont en commun de mettre l'accent sur la **diversité des institutions** et sur leur **importance dans la régulation des activités économiques**. À la différence du modèle standard, les approches institutionnalistes ne limitent donc pas la régulation au marché et au mécanisme des prix. Les définitions du concept d'institution sont variables au sein de ces divers courants, mais elles ont en commun de mettre l'accent sur le caractère durable des institutions et sur le fait qu'elles sont des règles (formelles ou informelles), des habitudes, des normes, des contraintes, au sein desquelles se déroulent les actions individuelles. Il existe donc un cadre institutionnel de la vie sociale (la famille, le marché, le droit, la morale, etc.) au sein duquel se mettent en place des organisations et se nouent des interactions sociales. Ainsi, selon D. North, récompensé par le prix Nobel d'économie en 1993, « *les institutions sont des contraintes établies par les hommes qui structurent les interactions humaines. Elles se composent de contraintes formelles (comme les règles, les lois, les constitutions), de contraintes informelles (comme des normes de comportement, des conventions, des codes de conduite auto-imposés) et des caractéristiques de leur application* ». Pour lui « *si les institutions sont les règles du jeu, les organisations et leurs entrepreneurs sont les joueurs* ». Les institutions

affectent les coûts de transaction et, selon leur plus ou moins grande efficacité, favorisent la croissance économique. Par l'évolution de la fiscalité, des droits de propriétés, les institutions peuvent modifier la structure des échanges économiques.

2.8 JEAN TIROLE : L'ÉCONOMIE DU BIEN COMMUN

L'économiste français Jean Tirole, président de l'École d'économie de Toulouse (TSE), directeur scientifique de l'Institut d'économie industrielle à Toulouse, professeur invité au MIT, médaille d'or du CNRS en 2007, s'est vu décerner en octobre 2014 le prix Nobel d'économie. Il est le troisième Français à l'obtenir, après Gérard Debreu en 1983 et Maurice Allais en 1988. Ses recherches ont été très largement consacrées aux théories des organisations industrielles et à la réglementation. Mais il a produit de nombreux travaux dans des domaines aussi divers que la liquidité, la réglementation bancaire, les conditions d'existence de bulles sur les prix des actifs, la gouvernance de l'entreprise et l'économie comportementale. En collaboration avec O. Blanchard (désormais chef économiste du FMI), il a rédigé en 2003 un rapport sur le fonctionnement du marché du travail français, dans lequel sont traitées des questions, comme celle du contrat unique, qui restent au centre des débats aujourd'hui (« Protection de l'emploi et procédures de licenciement », CAE, 2003). Lorsque le Prix Nobel d'économie lui a été décerné, de nombreux hommages ont afflué. Pour Philippe d'Arvisenet et Yves Nosbusch, « *J. Tirole est un économiste tout à fait exceptionnel [...] il se distingue par la **diversité de ses contributions scientifiques** [...] reflétée dans le nombre impressionnant de ses collaborateurs [...] Son œuvre est caractérisée par des théories élégantes motivées par des problèmes réels et ayant des implications pratiques importantes, notamment dans le domaine de la régulation. J. Tirole a également contribué très fortement à*

l'enseignement de l'économie. Ses manuels sur l'organisation industrielle, la théorie des jeux ou encore la finance d'entreprise font référence dans les universités du monde entier et nombreux sont les économistes qui ont découvert ces domaines à travers ses ouvrages. » (« Hommage à Jean Tirole », Conjoncture BNPP, janvier 2015, par Philippe d'Arvisenet et Yves Nosbusch).

Dans un ouvrage destiné au grand public, intitulé *Économie du bien commun* (PUF, 2016), Jean Tirole s'interroge sur le rôle de l'économie (et de l'économiste) dans la Cité, et sur la contribution que la science économique et le chercheur peuvent apporter à la recherche du bien commun, c'est-à-dire à l'organisation de la société que nous appellerions de nos vœux si nous étions placés sous le « voile d'ignorance » évoqué par le philosophe John Rawls. Pour y parvenir, Jean Tirole part d'un principe clé : **quelle que soit notre position sociale, nous réagissons tous aux incitations** auxquelles nous sommes confrontés, et en fonction de nos préférences, la quête du bien commun passe alors, dans une économie de marché comme la nôtre, par des institutions adaptées qui permettent de garantir que l'intérêt individuel coïncide avec l'intérêt collectif. Si la science économique, discipline exigeante et accessible, dans l'état actuel de ses connaissances, n'a pas vocation à se substituer à la société en définissant le bien commun selon J. Tirole, elle peut néanmoins alimenter le débat sur les objectifs que l'on souhaite viser, et développer les outils de politique économique les plus efficaces pour les atteindre. **L'économie a une grande utilité sociale** selon lui : **celle de pouvoir éclairer le décideur** sur la manière la plus efficace d'allouer les ressources rares sur les marchés et de gérer les problèmes d'information. Il insiste ainsi tout particulièrement dans cet ouvrage sur l'enjeu de la diffusion de la culture économique auprès des citoyens, qui est encore loin d'être optimale en France, pour les élites dirigeantes en responsabilité de la politique économique, mais aussi pour les

lycéens et étudiants qui ont besoin d'une initiation intuitive, pragmatique et rigoureuse à l'économie.

Pour J. Tirole, **le marché constitue un mode d'organisation de l'économie qui s'est imposé** à travers le monde, **mais on peut toutefois en analyser pleinement les limites**. S'il est un instrument souvent efficace, le marché ne doit pas être une fin en soi, et « bénéficier des vertus du marché requiert souvent d'ailleurs de s'écarter du laisser-faire. Le marché est à la fois un lieu de compétition et de collaboration, et l'équilibre entre les deux est toujours délicat : sans nier les motivations éthiques et morales légitimes, il s'agit de créer les bonnes institutions (règles juridiques et sociales) et les incitations pertinentes pour réduire les comportements nocifs pour le lien social et la recherche du bien commun », selon J. Tirole. Il y aurait donc une certaine complémentarité de l'État et du marché : « l'État ne peut faire vivre (correctement) ses citoyens sans marché ; et le marché a besoin de l'État, non seulement pour protéger la liberté d'entreprendre et sécuriser les contrats au travers du système juridique, mais aussi pour corriger les défaillances ».

2 La théorie de l'inefficacité de la politique monétaire est assortie de l'hypothèse de la surprise monétaire et de la nécessaire crédibilité des banques centrales.

PRINCIPES FONDAMENTAUX DE MICROÉCONOMIE

« Lorsque toutes les terres fertiles sont cultivées, si le capital et la population augmentent, il faut plus de nourriture et elle ne peut venir que de terres qui sont moins bien situées [...] Sur les terres cultivées en premier, le rendement est le même qu'auparavant, mais comme le taux général de profit du capital est déterminé par les profits faits sur les investissements les moins rentables réalisés dans l'agriculture, une nouvelle répartition du rendement s'instaure sur ces terres. » - David Ricardo, 1817

« J'appelle richesse sociale l'ensemble des choses matérielles ou immatérielles (car la matérialité ou l'immatérialité des choses n'importe ici en aucune manière) qui sont rares, c'est-à-dire qui, d'une part, nous sont utiles, et qui, d'autre part, n'existent à notre disposition qu'en quantité limitée. » - Léon Walras, 1874

« L'économie politique pure est essentiellement la théorie de la détermination des prix sous un régime hypothétique de libre concurrence absolue. L'ensemble de toutes les choses, matérielles ou immatérielles, qui sont susceptibles d'avoir un prix parce qu'elles sont rares, c'est-à-dire à la fois utiles et limitées en quantité, forme la richesse sociale. C'est pourquoi l'économie politique pure est aussi la théorie de la richesse sociale. » - Léon Walras, 1900

3.1 La théorie du consommateur : fonction de demande, surplus et optimum

3.2 La théorie du producteur : la maximisation du profit

3.1 LA THÉORIE DU CONSOMMATEUR : FONCTION DE DEMANDE, SURPLUS ET OPTIMUM

3.1.1 *L'utilité totale et l'utilité marginale*

L'économiste s'intéresse à la manière dont les consommateurs parviennent à opérer des choix, et ils prennent en compte trois dimensions :

- Ce que les consommateurs souhaitent acheter en fonction de leurs goûts et préférences ;
- Les prix des biens et services et leur évolution ;
- Le budget que les consommateurs sont prêts à mobiliser.

Le consommateur acquiert des biens et des services divers dans des quantités qu'il juge suffisantes pour atteindre une satisfaction maximale. Mais le consommateur est contraint par son revenu et ne peut acheter l'ensemble des biens et services dans des quantités illimitées.

Le choix du consommateur porte sur des paniers de biens. pour simplifier l'analyse, on raisonne généralement sur des paniers constitués de deux biens. Le panier X se note (x_1, x_2) , x_1 représentant la quantité de bien 1 que le consommateur décide d'acheter. C'est le concept abstrait de préférence qui permet d'ordonner les différents paniers de biens. Tout consommateur rationnel est ainsi capable de dire s'il préfère tel panier de bien à tel autre ou s'il lui est indifférent de consommer l'un ou l'autre des deux paniers. Du fait d'une hypothèse de non-satiété, un consommateur préfère toujours consommer plus. Tout panier de biens comportant une quantité supérieure de l'un des deux biens, à quantité égale du second, sera ainsi préféré au premier. Cependant, en vertu de l'hypothèse de désirabilité des biens, le consommateur ayant également une préférence pour la diversité, il consommera toujours des deux biens.

C'est le concept d'utilité qui permet de mesurer la satisfaction que retire le consommateur de la consommation d'un panier de biens. Le niveau de l'utilité dépend donc de la quantité consommée (Q) et peut être formalisé à l'aide d'une fonction d'utilité : $U = f(Q) = f(x_1, x_2)$. Cette fonction indique ainsi le niveau d'utilité d'un consommateur donné lorsqu'il acquiert les quantités x_1 et x_2 des biens 1 et 2.

Dans la théorie du consommateur, le raisonnement consiste toujours à faire augmenter la quantité de biens pour en analyser l'impact sur la fonction d'utilité.

La fonction d'utilité est croissante : plus les quantités de x et y augmentent, plus l'utilité retirée de la consommation du bien augmente. Si x_1 et x_2 progressent fortement, le niveau d'utilité atteindra un maximum voire décroîtra ensuite.

Dans la théorie microéconomique, ce n'est pas tant le niveau d'utilité qui importe que sa variation aux changements de consommation dans le cadre du **raisonnement à la marge** (marginalisme). On parle ainsi d'utilité marginale d'un bien pour désigner l'accroissement d'utilité procuré par la consommation d'une unité supplémentaire de ce bien, les quantités des autres biens restant constantes.

Le principe de « l'utilité marginale décroissante » stipule que le supplément de satisfaction obtenu par la consommation d'une unité supplémentaire de bien diminue au fur et à mesure de l'augmentation de la consommation de ce bien. Pour illustrer ce principe, les économistes prennent généralement l'exemple d'un individu qui divise la quantité d'un bien alimentaire, comme un plat de crêpes par exemple, en dix parts égales qui seront consommées en 24 heures : les deux premières portions du bien alimentaire sont indispensables à l'existence humaine et par conséquent, leur utilité est infinie. De la troisième à la dixième dose, l'utilité marginale décroît régulièrement, tout en restant positive. Ainsi, plus un individu consomme, plus il se rapproche de la satiété et moins la consommation supplémentaire apportera d'utilité.

Mais comment mesurer l'utilité ? On peut distinguer deux méthodes :

– La **mesure cardinale de l'utilité** : il s'agit de construire une échelle numérique d'utilité, l'individu pouvant dire que telle quantité du bien 1 lui confère une satisfaction qu'il peut évaluer numériquement. L'instauration d'une telle échelle est impossible lorsqu'il faut agréger les utilités individuelles, chacun ayant sa propre estimation de l'utilité. Selon la conception cardinale de l'utilité, le consommateur est capable d'exprimer par un nombre la satisfaction retirée de la consommation d'une quantité déterminée de biens. Le consommateur doit être ainsi capable de donner une « note » pour chaque quantité de bien consommée. Si un bien 1 reçoit la note 10 et un bien 2 la note 5, cela signifie donc que le bien 1 procure deux fois plus d'utilité que le bien 2. Cette hypothèse a été retenue par les premiers théoriciens marginalistes, tels **L. Walras**, **W. S. Jevons**, **C. Menger**. Selon ces auteurs, il fallait pouvoir mesurer l'utilité du consommateur afin de construire un modèle microéconomique, et il devenait donc absolument nécessaire de trouver un indicateur de mesure pertinent de l'utilité. Mais la construction d'une fonction d'utilité dans le cadre de la mesure cardinale de l'utilité s'avère impossible. En effet, le consommateur doit être capable de mesurer l'utilité de chaque bien de façon indépendante.

– La **mesure ordinale de l'utilité** : le problème de l'impossibilité de la mesure cardinale de l'utilité sera contourné par **V. Pareto** au début du ^{xx}^e siècle, qui propose de parler d'utilité ordinale, en substituant la classification à la quantification. Selon V. Pareto, considérer que les individus sont capables de classer les choix offerts selon un ordre de préférence, mais sans attribuer à chacun un indice quantitatif précis, est à la fois plus réaliste et surtout suffisant pour construire une analyse microéconomique du comportement du consommateur. Selon la conception ordinale de l'utilité, le consommateur peut ainsi établir un ordre de préférence

entre différents paniers de consommation sans pour autant attribuer à chacun d'eux une « note » précise. Ainsi, soit trois biens 1, 2 et 3. Un individu peut toujours dire qu'il préfère le bien 1 au bien 2 ou qu'il est indifférent au bien 1 et au bien 2. Dire qu'il préfère le bien 1 au bien 2 signifie que son niveau d'utilité est plus élevé pour le bien 1 que pour le bien 2 ; et dire qu'il est indifférent au bien 1 et au bien 2 indique que le niveau d'utilité est le même pour le bien 1 et pour le bien 2.

3.1.2 *L'optimum du consommateur*

C'est à partir de la fonction d'utilité que l'on construit les courbes d'indifférence du consommateur, grâce à une représentation graphique initiée par V. Pareto pour des biens substituables.

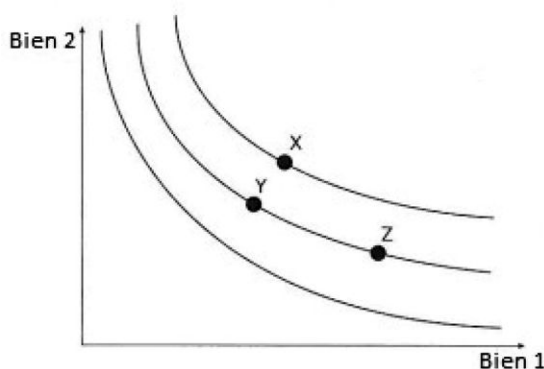
Une courbe d'indifférence représente l'ensemble des paniers de consommation qui procurent la même utilité au consommateur (on parle également de courbe d'iso-utilité). Chaque bien a un prix et le consommateur a un revenu limité ; sa contrainte budgétaire restreint donc ses dépenses de consommation. Du point de vue du raisonnement économique, étudier le comportement du consommateur, c'est évaluer l'incidence sur ses choix d'une variation de son revenu, du prix des biens, de la quantité de biens disponibles, etc.

Supposons par exemple que les deux biens essentiels à la satisfaction du consommateur soient le café et les crêpes. À la combinaison d'une tasse de café et d'une crêpe va correspondre une valeur mesurant la satisfaction. Il existe tout un ensemble de combinaisons possibles (deux tasses, une crêpe ; 1,5 tasse, deux crêpes, etc.) qui procurent au consommateur le même niveau de satisfaction et entre lesquelles il est indifférent : la courbe d'indifférence est le lieu de ces combinaisons. Le long de cette courbe, le renoncement à un peu de café est compensé par un supplément de crêpes. Il existe, pour un même individu, plusieurs courbes d'indifférence qui représentent chacune un

niveau de satisfaction différent. L'ensemble des courbes d'indifférence représente la **carte d'indifférence du consommateur**.

Les courbes d'indifférence ont trois caractéristiques :

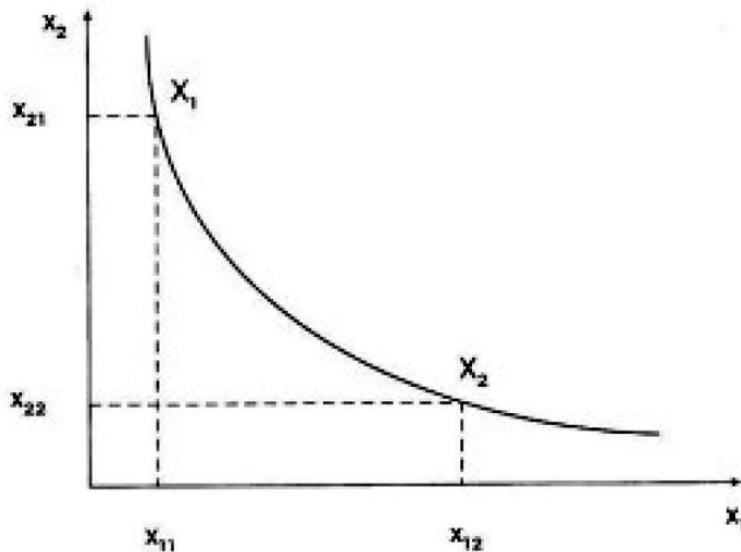
- **Elles sont décroissantes** : la courbe d'indifférence est décroissante, car la quantité de café qu'on est prêt à abandonner pour obtenir une quantité supplémentaire de crêpes décroît, à mesure que l'on a moins de café. Il existe en théorie une infinité de courbes d'indifférence possibles qui jamais ne se croisent : toute combinaison associant à la fois plus de café et plus de crêpes est toujours préférée tant que le seuil de satiété n'est pas atteint. Ainsi, plus on accède à une courbe d'indifférence éloignée de l'origine des axes et plus, selon la théorie microéconomique, la satisfaction augmente.
- **Elles sont convexes par rapport à l'origine des axes** : du fait de l'intensité décroissante des besoins, une augmentation du bien 1 doit être compensée par une baisse plus importante du bien 2.
- **Deux courbes d'indifférence ne peuvent pas se couper** : deux courbes d'indifférence n'ont, par définition, pas le même niveau de satisfaction. Dès lors, au point d'intersection entre ces deux courbes, ce point prend le niveau de satisfaction de l'une de ces deux courbes, ce qui est incohérent.



Dans cet exemple (courbe d'indifférence pour des biens substituables), le panier X est préféré aux paniers Y et Z, car il est situé sur une courbe d'indifférence représentant un niveau d'utilité plus élevé.

Les paniers Y et Z procurent le même degré de satisfaction au consommateur. On dit qu'il est indifférent à l'égard de ces deux paniers.

Le long d'une courbe d'indifférence, pour maintenir une utilité constante, le consommateur doit renoncer à une certaine quantité d'un des deux biens pour pouvoir consommer une quantité supplémentaire de l'autre bien.



Ainsi, dans le passage de X_1 à X_2 , il a dû accepter de renoncer à la quantité $(x_{21} - x_{22})$ de bien 2, en échange d'une augmentation $(x_{11} + x_{12})$ de la quantité de bien 1. Le taux auquel le consommateur est disposé à substituer une quantité de bien 2 au bien 1 est appelé **taux marginal de substitution (TMS)** entre les biens 1 et 2.

Ce TMS est égal au rapport des utilités marginales des deux biens.

3.1.3 De la contrainte de budget à l'équilibre du consommateur

Les préférences du consommateur ne vont pas pouvoir entièrement guider ses choix. En effet, ces choix sont limités par une contrainte imposée par son revenu et le prix des biens. Le problème du consommateur traite de ce qu'il lui est possible d'acheter. Le consommateur a un revenu R . Par hypothèse, il utilise l'intégralité de son revenu pour la consommation. Mais il a une contrainte de budget. L'ensemble budgétaire du consommateur est alors constitué

par les paniers accessibles pour des prix p_1 et p_2 (prix unitaires respectifs des deux biens, fixés par le marché) et un revenu R donné, intégralement dépensé.

La contrainte de budget marque la frontière de cet ensemble budgétaire et est formalisée par la relation :

$$p_1 x_1 + p_2 x_2 = R$$

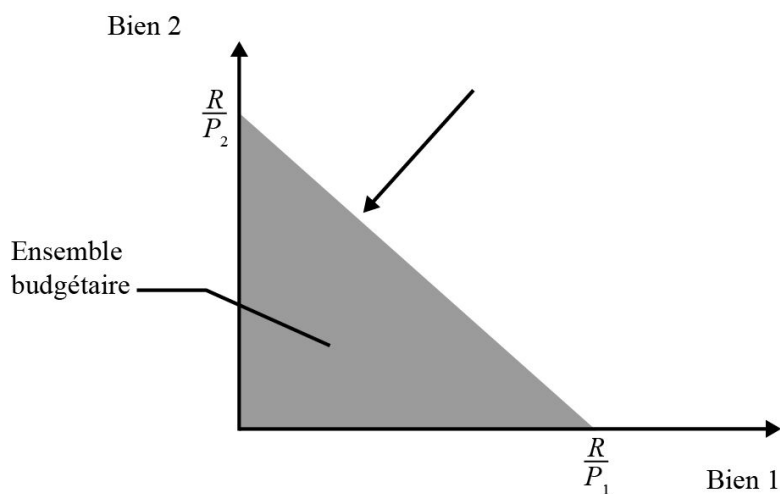
$p_1 x_1$ représente la dépense en bien 1. Elle est représentée par une droite, dont on peut écrire l'équation sous la forme :

$$x_2 = -\frac{p_1}{p_2} x_1 + \frac{R}{p_2}$$

L'ordonnée à l'origine (qui mesure la quantité maximale de bien 2 que l'on peut consommer si l'on dépense la totalité de son revenu dans l'achat de ce bien) est égale

à $\frac{R}{p_2}$ et la pente de la droite de budget est donnée par le rapport $-\frac{p_1}{p_2}$,

c'est-à-dire l'opposé du rapport des prix relatifs des deux biens : si le consommateur dépense une unité monétaire de plus pour consommer du bien 1, il devra réduire sa dépense en bien 2 (on parle également de coût d'opportunité de la consommation du bien 1).



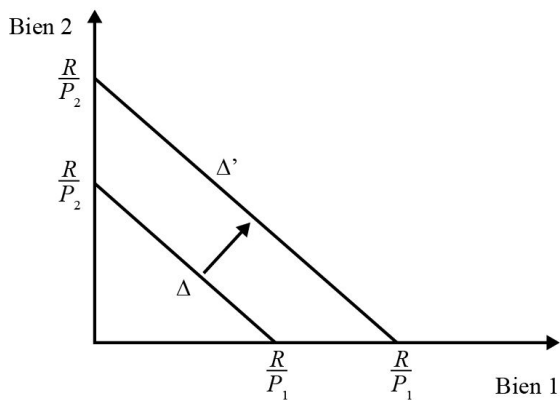
On peut analyser les déplacements de la droite de budget :

■ Dans le cas d'une variation de revenu

Un accroissement du revenu, toutes choses égales par ailleurs (à prix des biens inchangés)

accroît l'ordonnée à l'origine, $(\frac{R}{P_2})$

mais n'affecte pas la pente de la droite (elle est indépendante de R).



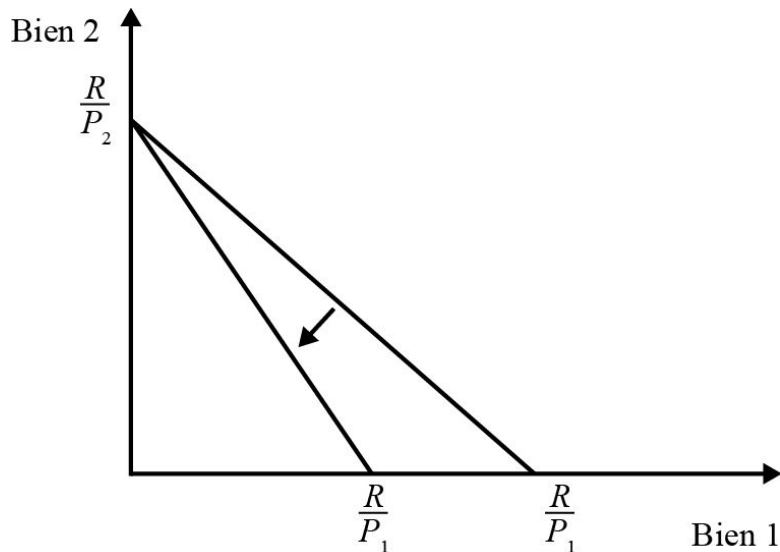
Une augmentation (baisse) du revenu se traduit donc par un déplacement parallèle et vers le haut (bas) de la droite de budget.

Les prix P_1 et P_2 étant fixés, le passage de Δ à Δ' traduit une augmentation du revenu et donc une augmentation du pouvoir d'achat du

consommateur.

■ Dans le cas d'une variation des prix

Un accroissement du prix p_1 , « toutes choses égales par ailleurs » (p_2 et R inchangés) ne modifie pas l'ordonnée à l'origine, mais accroît (en valeur absolue) la pente de la droite de budget. Celle-ci « pivote » alors autour de l'ordonnée à l'origine.



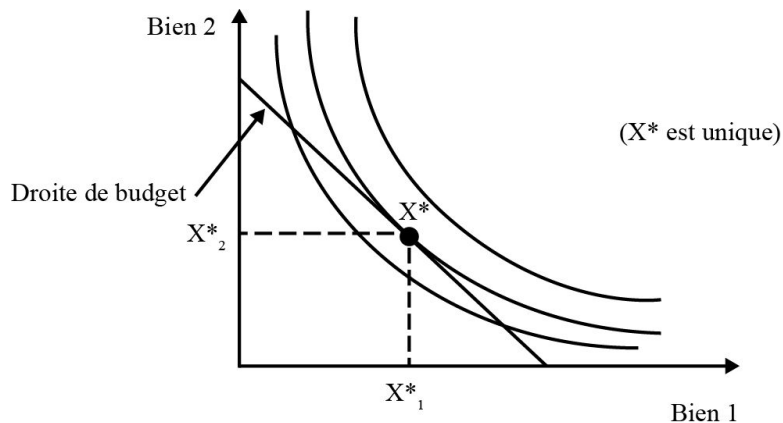
■ Effet substitution et effet revenu

L'effet global d'une variation du prix de l'un des deux biens donne lieu à deux effets : un effet substitution et un effet revenu.

- *L'effet de substitution* : la baisse du prix d'un bien va modifier les prix relatifs et le consommateur va avoir tendance, toutes choses égales par ailleurs, à augmenter la quantité demandée du bien devenu relativement moins cher et à diminuer la quantité demandée de l'autre bien.
- *L'effet revenu* : la baisse du prix de ce bien, toutes choses égales par ailleurs, provoque également un accroissement du pouvoir d'achat du consommateur. On peut donc s'attendre à ce que la demande de tous les biens (et non pas seulement de celui dont le prix a baissé) augmente.

On peut alors déterminer l'équilibre du consommateur :

L'équilibre du consommateur est le panier de consommation, traditionnellement nommé X^* , qui rend maximale la fonction d'utilité tout en étant compatible avec la contrainte de budget.



Le panier (x_1^*, x_2^*) constitue le choix optimal du consommateur. Il faut en effet choisir un point situé sur la droite de budget. Il en existe une infinité. Tant qu'il y a deux points d'intersection entre une courbe d'indifférence et la droite de budget, il existe une courbe d'indifférence située plus haut qui reste accessible au consommateur. Le point optimal (X^*) est le point de la droite de budget qui est tangent avec la courbe d'indifférence.

À l'équilibre du consommateur, le rapport des utilités marginales des deux biens est égal au rapport des prix des deux biens.

3.1.4 La fonction de demande

La fonction de demande de bien 1 (Dx_1) formulée par un consommateur exprime les quantités optimales consommées du bien x_1 en fonction des prix et du revenu auxquels ce consommateur est confronté :

$$Dx_1 = f(P, R).$$

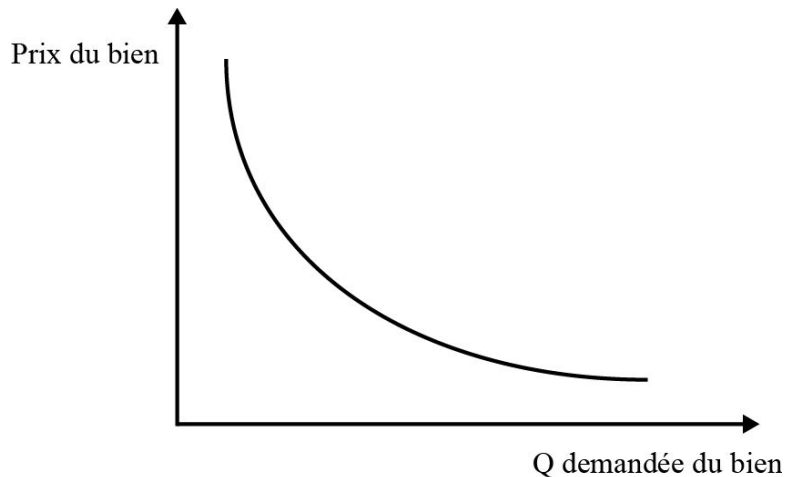
Il existe une infinité de courbes de demande pour les biens considérés : des places de concert, des bananes, des ordinateurs...

La courbe de demande fournit une indication sur la quantité de biens que le consommateur est prêt à acheter à chaque niveau de prix.

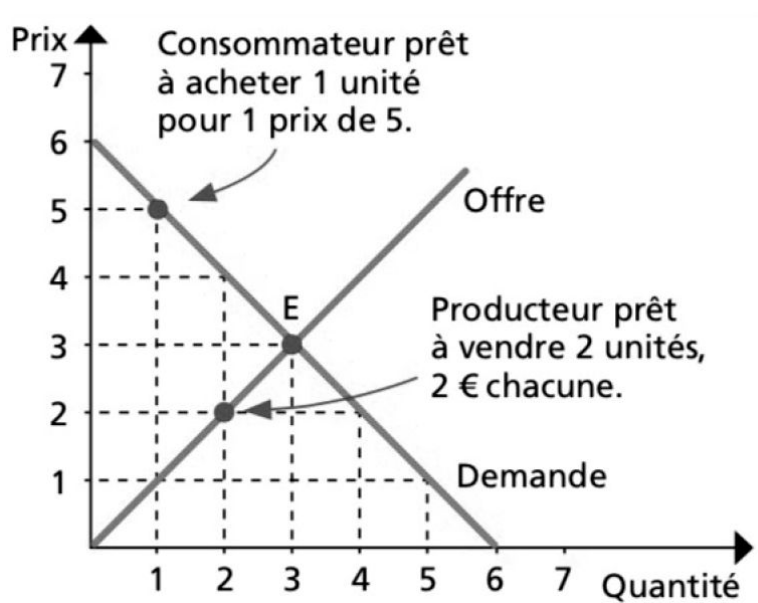
La loi de la demande exprime le fait que, dans la plupart des cas, la quantité demandée d'un bien décroît lorsque le prix de

ce bien augmente. La courbe de demande d'un bien est donc en général décroissante en fonction du prix de ce bien.

Les courbes de demande individuelle découlent des trois composantes du problème du consommateur : les préférences, les prix, et la quantité d'argent disponible.



Le surplus du consommateur mesure le montant (fictif) des gains qu'un consommateur retire de ses achats en faisant la différence entre le prix effectivement payé et le prix qu'il aurait été prêt à payer.



Sur cet exemple, on constate graphiquement que des consommateurs sont prêts à payer plus que 3 euros l'unité. Pour un prix de 5 euros, 1 unité du bien est demandée et pour un prix de 4 euros, les consommateurs sont prêts à

acheter 2 unités. Comme le prix de marché est de 3 euros, ils peuvent obtenir ces quantités à un prix inférieur à celui qu'ils auraient été prêts à payer. La notion de surplus est très importante pour pouvoir comparer les gains effectués par les agents dans différentes situations : plus le surplus est élevé, plus on peut considérer que les agents sont « gagnants » au prix fixé sur le marché, puisqu'ils obtiennent les biens considérés à un prix inférieur à ce qu'ils auraient été disposés à payer au maximum pour chaque unité de bien, c'est-à-dire aux prix indiqués le long de la courbe de demande.

La théorie microéconomique retient également une notion importante dans l'analyse, celle de l'élasticité de la demande :

■ L'élasticité-prix

L'élasticité de la demande par rapport au prix (ou élasticité-prix de la demande) peut être définie par le rapport entre la variation relative de la quantité demandée et la variation relative du prix. Elle mesure la variation en pourcentage de la quantité demandée d'un bien pour une variation de 1 % de son prix. Prenons le cas par exemple d'un consommateur de crêpes dont la demande diminue de 2 % quand le prix augmente de 1 %.

$$\text{On note alors : } e = \frac{\frac{\Delta q}{q}}{\frac{\Delta R}{R}} = \frac{-0,02}{0,01} = -2$$

L'élasticité-prix de la demande de crêpes est donc de -2 dans ce cas de figure. La crêpe constitue alors un bien ordinaire (la quantité demandée baisse quand le prix du bien augmente).

On peut dresser une typologie des différentes valeurs possibles de l'élasticité-prix.

$e < 0$	Biens ordinaires
$e > 0$	Biens Veblen (effet de snobisme) encore appelés biens de consommation ostentatoire ou biens Giffen : la quantité demandée augmente lorsque le prix du bien s'accroît.
$-1 < e < 0$	Biens faiblement élastiques (biens nécessaires) ; la quantité

demandée varie relativement moins vite que le prix.

- $e < -1$ Biens fortement élastiques ; la quantité demandée varie relativement plus vite que le prix du bien.
- $e = -1$ Élasticité unitaire : la quantité demandée varie proportionnellement au prix.
- $e = 0$ Bien parfaitement inélastique : le prix n'a aucun effet sur la quantité demandée.

■ L'élasticité-revenu

L'élasticité-revenu de la demande d'un bien mesure la sensibilité de la demande aux variations du revenu. Elle se mesure par le rapport entre la variation relative de la quantité demandée d'un bien par un consommateur et la variation relative du revenu de ce consommateur.

$$\varepsilon_{q/R} = \frac{\frac{\Delta q}{q}}{\frac{\Delta R}{R}}$$

On peut dresser une typologie des différentes valeurs possibles de l'élasticité-revenu.

- $e_{Q/R} > 0$ Biens normaux. La consommation augmente avec le revenu.
- $e_{Q/R} < 0$ Biens inférieurs. La consommation de ce type de bien diminue lorsque le revenu augmente.
- $0 < e_{Q/R} < 1$ Biens de nécessité. La consommation augmente avec le revenu, mais moins que proportionnellement.
- $e_{Q/R} = 1$ La dépense de consommation est strictement proportionnelle à la variation du revenu.
- $e_{Q/R} > 1$ Biens de luxe. La part des dépenses consacrées à ce type de biens augmente plus vite que le revenu.

■ L'élasticité-croisée de la demande

On peut également calculer une élasticité croisée de la demande du bien 1 par rapport au bien 2, qui mesure la variation relative de la quantité demandée du bien 1 quand le prix du bien 2 varie.

L'élasticité croisée mesure la sensibilité de la demande d'un bien à la variation du prix d'un autre bien. Mathématiquement, elle s'écrit :

$$E \text{ croisée} = \frac{\frac{\Delta q_{x1}}{q_{x1}}}{\frac{\Delta p_2}{p_2}}$$

L'élasticité croisée permet de préciser la substituabilité ou la complémentarité entre deux produits :

- 1) Si l'élasticité croisée est négative, cela signifie qu'une hausse du prix du bien 2 provoque une baisse de la quantité demandée du bien 1 : ce sont des biens *complémentaires*.
- 2) Si l'élasticité croisée est positive, cela signifie qu'une hausse du prix du bien 2 provoque une augmentation de la quantité demandée du bien 1 : ce sont alors des biens *substituables*.

■ La courbe de demande du marché

La courbe de demande du marché est construite par agrégation des courbes individuelles de demande (parfois appelée courbe de demande agrégée). À chaque niveau de prix, elle fait correspondre la somme des demandes de tous les demandeurs sur le marché.

3.2 LA THÉORIE DU PRODUCTEUR : LA MAXIMISATION DU PROFIT

3.2.1 *La fonction de production et la combinaison productive*

Une fois posés les fondements de la théorie du consommateur, essentiels dans une économie de marché où le succès (ou l'échec) des entrepreneurs dépend de leur

capacité à satisfaire la demande, il est possible d'analyser les rouages des décisions de production. Dans une économie de marché, c'est en effet le producteur qui fait le lien entre la demande et la production. Ce sont les producteurs, en quête du profit, qui perçoivent la demande et assument le risque et l'incertitude. Ils cherchent à répondre aux besoins des consommateurs en mobilisant des ressources (main-d'œuvre, machines, terre) et en réalisant des innovations. Le problème du producteur peut se résumer de la manière suivante : comment les producteurs décident-ils de *ce qu'ils doivent produire* et de *la quantité à produire* ?

On fait l'hypothèse que les producteurs prennent leur décision dans le cadre d'un marché parfaitement concurrentiel, où aucun participant n'est assez puissant pour influencer le prix du marché, où les producteurs offrent des biens homogènes et identiques sur le marché, et où il y a libre entrée et libre sortie sur le marché.

Les trois éléments pris en compte par le producteur dans ses calculs sont alors :

- La fabrication du bien
- Le coût de production
- Le gain commercial

Dans la théorie microéconomique, le producteur maximise sa production, son profit, et cherche à minimiser ses coûts de production. On appelle facteurs de production (ou bien *inputs*) tous les éléments nécessaires à la production d'un bien ou d'un service (ou bien *output*) et utilisés dans le processus de production. Ces facteurs de production sont combinés ensemble pour donner un produit.

La concurrence impose aux entreprises des contraintes techniques : l'ensemble de toutes les combinaisons d'inputs et d'outputs qui correspondent à un processus de production techniquement réalisable est appelé un ensemble de production.

Le passage des *inputs* aux *outputs* qui se fait par le plan de production s'étudie à travers la fonction de production

(expression analytique de la technologie). La fonction de production indique le niveau maximal de production qui peut être obtenu par les différentes combinaisons de facteurs de production, la technologie étant donnée.

La forme générale pour une firme quelconque est :

$$Q = f(X_1, X_2, \dots, X_j)$$

Dans cette expression, Q est la quantité du bien produite (*output*) par la firme considérée et $(X_1, X_2, \dots, X_j)$ sont les *inputs* utilisés par cette firme. La fonction de production totale est la relation fonctionnelle entre Q d'une part (l'*output*), et K et L d'autre part (les *inputs*). En effet, pour simplifier l'analyse de la fonction de production, on ne retient que deux inputs, K et L .

On désignera par L le facteur travail, par K le facteur capital et par Q la quantité produite à l'aide d'une combinaison productive de ces deux facteurs.

On aura donc : $Q = f(L, K)$

La notion d'homogénéité permet d'étudier la façon dont varie la production lorsque tous les facteurs augmentent dans la même proportion.

Une fonction de production sera homogène de degré k , si, lorsqu'on multiplie les quantités de facteurs utilisées par un réel λ , on multiplie la quantité produite par λ^k : $f(\lambda L, \lambda K) = \lambda^k f(L, K)$.

Les fonctions de production généralement étudiées en microéconomie sont les fonctions à élasticité de substitution constantes. L'élasticité de substitution est un indicateur qui mesure la variation relative du rapport entre le capital et le travail à la suite d'une variation du prix relatif du capital. Si les facteurs de production sont complémentaires, l'élasticité de substitution sera égale à zéro et si les facteurs sont substituables, elle sera proche de l'unité. Les fonctions à élasticité de substitution comprennent des cas particuliers, les plus couramment étudiés en microéconomie, notamment la fonction de type Cobb-Douglas.

Cette dernière est de la forme : $Q = A K^\alpha L^\beta$. Cette fonction est homogène de degré $(\alpha + \beta)$.

En effet : $A (I K)^\alpha (IL)^\beta = I^{(\alpha + \beta)} A K^\alpha L^\beta = I^{(\alpha + \beta)} Q$

Les rendements d'échelle seront donc constants si $\alpha + \beta = 1$.

La fonction de Cobb-Douglas homogène de degré 1 peut alors s'écrire sous la forme : $Q = A K^\alpha L^{1-\alpha}$, K et L étant des réels positifs, α et β étant des réels quelconques.

3.2.2 La mesure de l'efficacité de la production : la productivité

Pour mesurer l'efficacité de sa production, l'entrepreneur a besoin de calculer la productivité. La productivité est un indicateur d'efficacité de la combinaison productive : c'est le rapport entre le volume de production réalisé et la quantité de facteurs de production utilisés. On peut différencier le niveau de la productivité à un moment donné de l'accroissement de son niveau au cours d'une période, c'est-à-dire ce que l'on appelle les gains de productivité.

La productivité peut être évaluée en unité physique (par exemple le nombre de voitures produites par travailleur et par an), ou en valeur, lorsque la production et les facteurs sont évalués en monnaie.

On distingue généralement la productivité moyenne et la productivité marginale.

La productivité moyenne est le rapport entre l'output et l'un des inputs.

Si l'on raisonne à partir de chacun des deux facteurs, il en découle :

■ La productivité moyenne du travail

La productivité moyenne du travail est le rapport $PM(I) = Q / L$.

Q est alors exprimé en unités produites tandis que l'on choisit l'unité avec laquelle on cherche à mesurer (nombre d'heures travaillées, effectifs de salariés...). Si l'on prend en compte au dénominateur du rapport le nombre de salariés, on parle de

productivité par tête. Cette dernière s'interprète comme le nombre d'unités fabriquées par unité de travail. Autrement dit, si on fabrique 100 unités avec 10 personnes, la productivité par tête est de 10. Tout se passe comme si un travailleur fabriquait 10 unités.

■ La productivité moyenne du capital

La productivité moyenne du capital s'écrit : $PM(k) = Q / k$.

Elle s'interprète de façon analogue à la productivité moyenne du travail : c'est le nombre d'unités de biens et services fabriqués par unité de capital.

■ La productivité marginale

La productivité marginale se définit comme la variation de la production à la suite d'une variation infinitésimale du facteur de production. La productivité marginale apparaît donc comme la production supplémentaire résultant de la dernière unité de facteur utilisée.

On suppose qu'en partant d'une production nulle, la productivité marginale est d'abord croissante, décroissante, nulle puis négative. Selon la loi des rendements décroissants, la productivité marginale d'un facteur de production doit finir par décroître.

On peut en effet distinguer quatre phases :

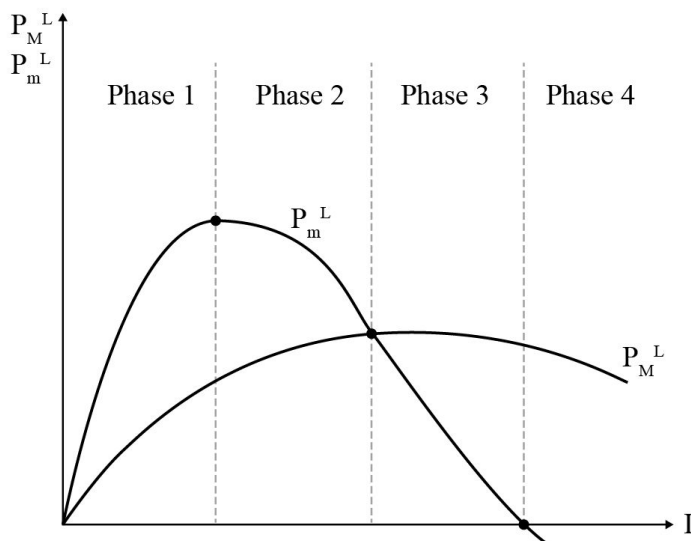
– **La période 1 : la productivité moyenne est croissante et la productivité marginale est elle aussi croissante.**

Dans ce cas, le producteur a intérêt à faire augmenter K et L, dans la mesure où la production augmente de plus en plus vite.

– **La période 2 : la productivité marginale reste positive, mais décroît.** En effet, la productivité moyenne continue d'augmenter, tandis que la productivité marginale reste positive, mais décroissante. Elle reste d'ailleurs supérieure à la productivité moyenne, ce qui explique que celle-ci continue d'augmenter. Le producteur a, dans ce cadre, toujours intérêt à augmenter la quantité d'input puisque sa productivité moyenne augmente.

– **La période 3 : la productivité marginale reste positive et poursuit son rythme décroissant.** La productivité moyenne atteint alors son maximum puis décroît tout en restant positive. La productivité marginale, quant à elle, continue de décroître : dans ce cas, le producteur doit se demander quand interrompre l'augmentation de l'utilisation de l'input.

– **La période 4 : la productivité marginale devient négative.** Dans ce cas, si le producteur accroît l'utilisation de l'input, comme la productivité marginale devient négative, cela devient irrationnel, car l'augmentation de la quantité d'input ferait baisser la production.



On remarque ainsi que les productivités moyenne et marginale croissent, atteignent un maximum, puis décroissent. La loi des rendements marginaux décroissants est précisément la zone dans laquelle l'usage d'une quantité plus grande de facteurs conduit à faire chuter la productivité.

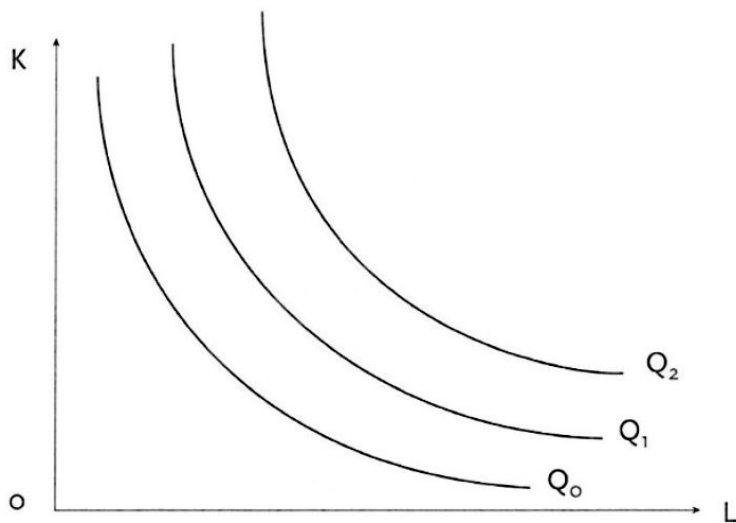
Un usage plus important du travail et/ou du capital fait augmenter la production dans un premier temps. Mais par la suite, il ne serait pas rationnel de pousser l'utilisation des facteurs de production au-delà d'un point maximum. En effet, quelle serait la logique d'un producteur qui, lorsqu'il accroît ses coûts de production, fait baisser la production ?

3.2.3 Des isoquantes au taux marginal de substitution technique (TMST)

La théorie microéconomique utilise pour la théorie du producteur un cadre analytique analogue à celui de la théorie du consommateur avec les courbes d'indifférence. Une isoquante (ou courbe d'isoproduit) est ainsi une courbe indiquant l'ensemble des combinaisons de capital (K) et de travail (L) qui, pour un état donné des techniques, permettent de produire une même quantité de biens (Q).

Il y a une infinité d'isoquantes, chacune correspondant à un niveau de production différent, qui va en s'accroissant lorsqu'on se déplace vers le haut.

Le **taux marginal de substitution technique** (TMST) est le taux auquel on peut échanger les facteurs de production tout en conservant le même niveau de production. Il est égal au rapport des productivités marginales des deux facteurs de production.



3.2.4 Les fonctions de coût

Les coûts de production représentent pour une firme l'ensemble des dépenses nécessaires à l'obtention d'une quantité donnée d'output. La fonction de coût exprime la

relation existant entre les coûts de production et les quantités produites dans les conditions optimales.

■ Coût total, coût moyen, coût marginal

Le **coût total** de production (C_T) est la dépense minimale qu'une entreprise doit engager pour atteindre un niveau de production donné. Il se calcule en faisant la somme des coûts fixes (C_F) et des coûts variables (C_V).

Les **coûts fixes** sont des coûts indépendants du volume de la production. Il s'agit par exemple des bâtiments, machines, véhicules... Les coûts fixes sont donc exprimés à l'aide d'une constante : $C_F = k$.

Les **coûts variables** au contraire dépendent des quantités produites : $C_V = C_V(Q)$.

On a donc finalement : $C_T(Q) = k + C_V(Q)$

Le **coût moyen** de production est obtenu en divisant le coût total par la quantité produite (non nulle). Il mesure le coût unitaire moyen.

$$C_M(Q) = \frac{C_T(Q)}{Q}$$

Le **coût marginal** est le supplément de coût entraîné par la production d'une unité supplémentaire d'*output*. Lorsqu'on raisonne sur des accroissements infinitésimaux de production (hypothèse de divisibilité), le coût marginal s'exprime par la dérivée de la **fonction de coût total** :

$$C_m(Q) = C'_T(Q)$$

■ Les relations entre les courbes de coût à court terme

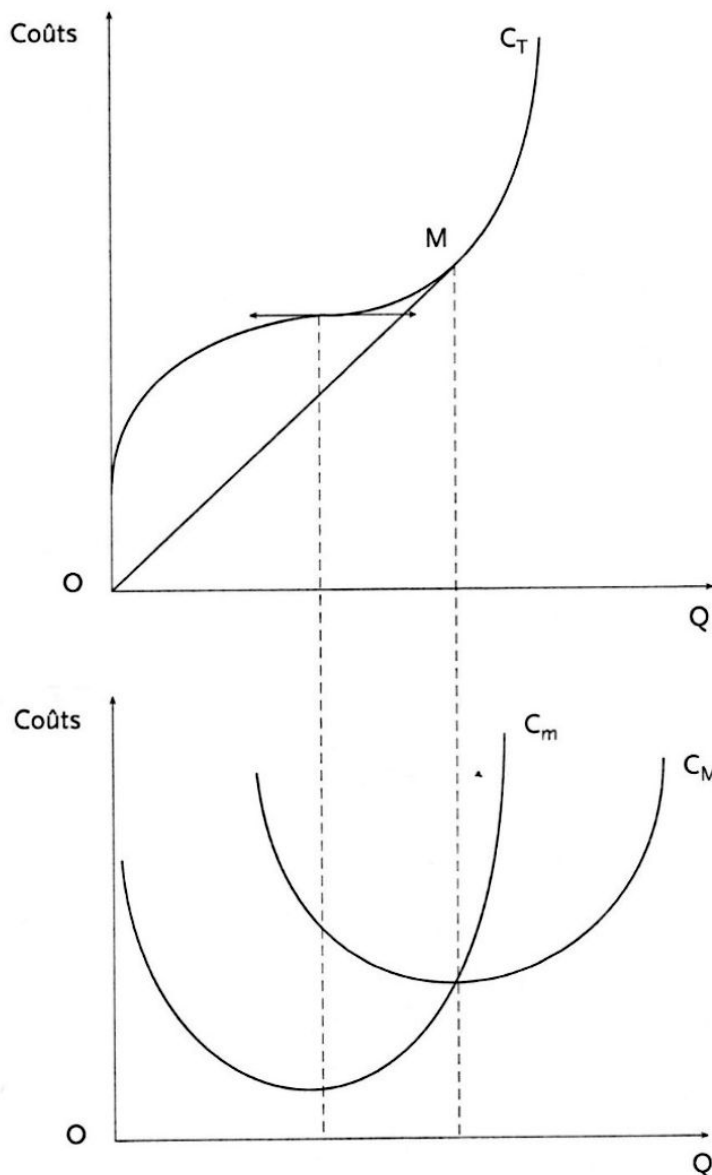
Soit M un point de la courbe de coût total.

La droite OM passe par l'origine. Elle est donc de la forme : $C_T(Q) = a \times Q$.

La pente de cette droite, a , est donc égale à : $\frac{C'_T(Q)}{Q}$

Le coût moyen correspondant à la quantité Q est donc donné par la pente de la droite (OM).

Le minimum du coût moyen est égal au minimum de la pente de la droite OM (M parcourant la courbe de coût total), qui est le coefficient directeur de la tangente à C_T passant par O. Le coût marginal est donné par la pente de la tangente en un point de la courbe de coût total. La courbe de coût marginal coupe la courbe de coût moyen en un point dont l'ordonnée est le minimum du coût moyen.



3.2.5 Minimisation des coûts et combinaison optimale de facteurs

Dans la théorie microéconomique, le producteur a pour objectif de maximiser son profit sous contrainte de ses coûts. Il est alors possible de mener l'analyse via la minimisation des coûts de production qui constitue alors l'objectif à atteindre, et ce pour un niveau de production donné.

■ La contrainte budgétaire du producteur : la droite d'isocoût

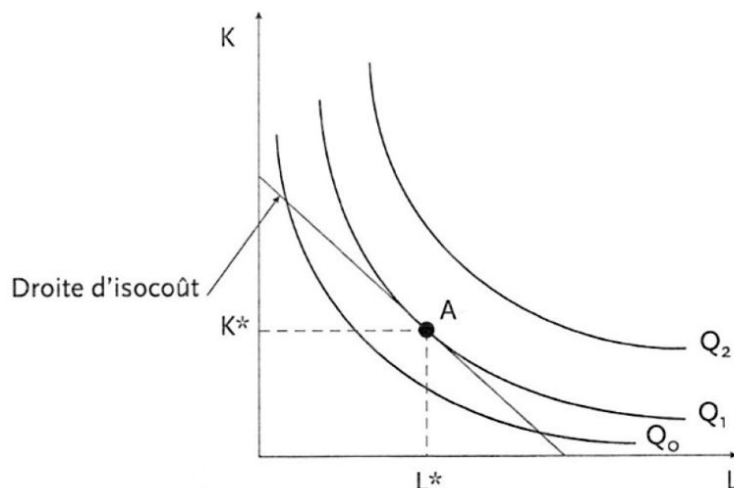
Si l'on appelle D la dépense que le producteur affecte à l'achat des facteurs de production (travail et capital), w le prix du facteur travail et r le prix du facteur capital, on peut écrire la contrainte de budget du producteur de la façon suivante : $D = wL + rK$.

Ou encore : $K = -w/r L + D/r$.

L'équation de la droite d'isocoût est donc celle d'une droite dont la pente est égale à l'opposé du rapport des prix des facteurs de production (inputs).

■ La combinaison optimale des facteurs

Le producteur qui veut atteindre un niveau de production donné minimise ses coûts lorsqu'il choisit la combinaison productive (combinaison de facteurs) qui se trouve au point de tangence entre l'isoquant considéré et une droite d'isocoût.



De la même façon que pour la théorie du consommateur, on peut montrer qu'à l'équilibre du producteur, le TMST, qui est égal au rapport des productivités marginales des facteurs, est égal au rapport des prix des facteurs de production.

La technique de production optimale est celle qui minimise les coûts de l'entreprise. En effet, pour un objectif de production donné (pour chaque isoquante), il existe un rapport K/L qui minimise le coût total et qui dépend du coût relatif des coûts de production.

■ Le profit

Le profit, que l'on peut noter (π), est égal à la différence entre les recettes totales et les coûts totaux. la recette totale se calcule en faisant le produit entre la quantité produite (Q) et le prix de vente (P) du bien.

$$\pi(Q) = (P \times Q) - C_T(Q) = Q \left(P - \frac{C_T(Q)}{Q} \right) = Q \times (P - C_M(Q))$$

La valeur du profit dépend donc de l'écart entre le prix du marché et la valeur du coût moyen. En concurrence pure et parfaite, le prix de vente d'un bien n'est en effet pas maîtrisé par le producteur ; il est fixé par le marché :

$$\pi(Q) = Q (P - C_M).$$

Ce profit est donc positif pour $P \geq C_M$.

■ L'équilibre du producteur en concurrence pure et parfaite

L'équilibre de court terme :

La recette moyenne du producteur est égale au prix de vente du produit.

En effet, la recette totale (R_T) est telle que $R_T(Q) = P \times Q$, et donc la recette moyenne est égale à :

$$R_M(Q) = \frac{R_T(Q)}{Q} = \frac{P \times Q}{Q} = P$$

Comme par ailleurs, en concurrence pure et parfaite, une unité supplémentaire de bien ne peut qu'être vendue au prix unitaire de ce bien, la recette marginale est aussi égale au prix.

On a donc : $R_M(Q) = R_m(Q) = P$.

Le profit (π) est tel que : $\pi(Q) = P \times Q - C_T(Q)$.

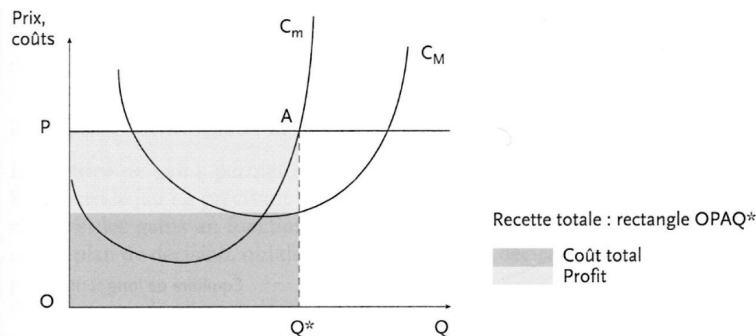
Le profit est maximal lorsque sa dérivée première s'annule (la dérivée seconde étant négative).

$$\pi(Q) = P - C'_T(Q) = 0 \leftrightarrow P = C'_T(Q) = C_m$$

La condition de second ordre exige que : $\pi''(Q) = -C''(Q) < 0$, soit $C''(Q) > 0$

Dans le cadre de la concurrence parfaite, la fonction de coût marginal doit donc être croissante pour que la valeur de la production maximise le profit.

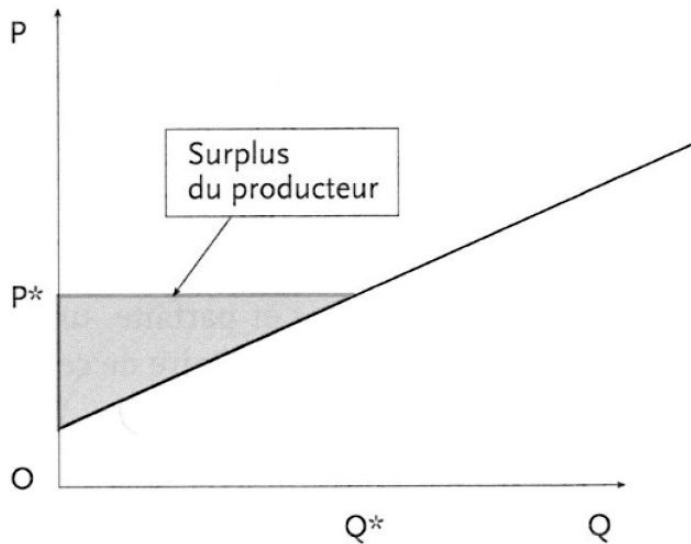
Ainsi, pour maximiser son profit, le producteur doit vendre une quantité telle que : $C_m = P$, et donc telle que : $C_m = R_m$.



À court terme, la courbe d'offre du producteur est donnée par la partie de la courbe de coût marginal située au-dessus de la courbe de coût moyen.

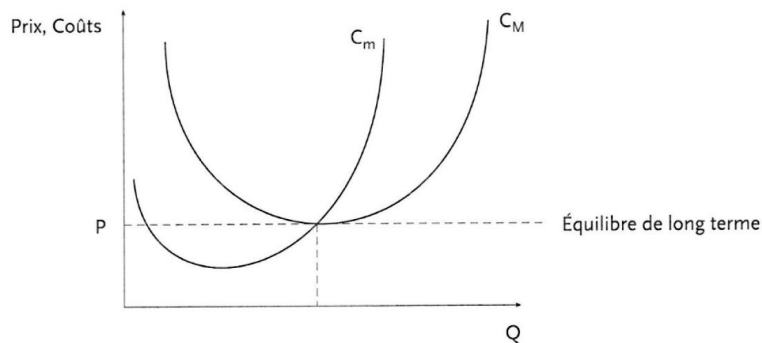
Le surplus du producteur :

Le concept de surplus du producteur permet de mesurer le montant des gains qu'un producteur retire de ses ventes en faisant la différence entre le prix auquel il vend effectivement les biens offerts et le prix qu'il aurait été prêt à accorder.



L'équilibre de long terme :

À long terme, les perspectives de profit attirent de nouvelles firmes sur le marché. Ces nouveaux arrivants utilisent la meilleure technique de production et les fonctions de coûts sont alors identiques pour toutes les entreprises. L'accroissement de l'offre provoque une baisse du prix de marché et donc du profit économique, jusqu'au niveau du seuil de rentabilité (minimum du coût moyen). À ce niveau, le profit économique (profit global) est nul et il n'y a donc plus d'incitation à entrer sur le marché.



Les rendements d'échelle :

À long terme, le producteur peut investir et modifier en profondeur son organisation. L'entreprise peut multiplier son niveau de production d'un certain coefficient : on peut raisonner en longue période lorsque les deux facteurs de production varient simultanément. On considère

généralement que la technique de production est la même, soit que la part du capital par rapport à celle du travail reste constante (le coefficient de capital K/L est constant).

Les rendements d'échelle désignent une relation entre la quantité produite d'un bien et la quantité de facteurs de production utilisés, et ils permettent de mesurer l'efficacité productive des facteurs.

On peut distinguer trois cas en relation avec le degré d'homogénéité de la fonction de production :

- Les rendements d'échelle sont croissants lorsque la quantité produite augmente plus vite que les quantités de facteurs mises en œuvre ;
- Les rendements d'échelle sont décroissants lorsque la quantité produite augmente moins vite que les quantités de facteurs mises en œuvre ;
- Les rendements d'échelle sont constants lorsque la quantité produite augmente au même rythme que les quantités de facteurs mises en œuvre.

Il ne faut toutefois pas confondre :

- La notion de *rendement d'échelle* et celle de *rendements factoriels* : les rendements factoriels désignent une relation entre la quantité produite d'un bien et la quantité d'un facteur de production supposé variable (par exemple le travail) et l'autre restant fixe (le capital). Les rendements factoriels permettent de mesurer l'efficacité productive d'un des facteurs et constituent un des éléments caractérisant une fonction de production. Les rendements factoriels affectent l'évolution de la productivité marginale du facteur variable. Ainsi, celle-ci progresse lorsque les rendements factoriels sont croissants, diminue lorsqu'ils sont décroissants, et ne change pas lorsque les rendements sont constants.
- La notion de *productivité globale des facteurs (PGF)* et celle de *rendement d'échelle* : la PGF mesure l'efficacité globale de la combinaison productive, et si en tant que notion macroéconomique, elle est souvent assimilée au

progrès technique, elle tient compte d'autres facteurs, comme les institutions et les infrastructures. La PGF représente la part de l'augmentation du volume de la production qui ne peut être expliquée par une augmentation de la quantité des facteurs de production utilisés.

Les entreprises peuvent également réaliser des *économies d'échelle* : elles apparaissent lorsque la baisse du coût moyen résulte de l'augmentation du volume de l'échelle de la production de l'entreprise. Cette diminution du coût unitaire provient essentiellement des rendements d'échelle croissants des facteurs de production au sein de l'entreprise. En effet, en présence de rendements croissants, la quantité de facteurs, dont découle l'accroissement du coût total, progresse moins vite que la quantité produite, ce qui conduit alors mécaniquement à une diminution du coût moyen.

3.2.6 La fonction d'offre

■ La fonction d'offre individuelle

L'optimum du producteur permet la construction de la fonction d'offre. La fonction d'offre de l'entreprise représente la quantité produite en fonction du prix quand l'entreprise maximise le profit. L'entreprise a intérêt à mettre en vente un produit sur le marché si elle parvient à couvrir au moins ses coûts de production et à dégager une marge de profit. En concurrence pure et parfaite, le prix d'offre se fixe au niveau du coût marginal, et il est utile de rappeler que l'entreprise ne peut agir sur le prix du marché qui est pour elle une donnée.

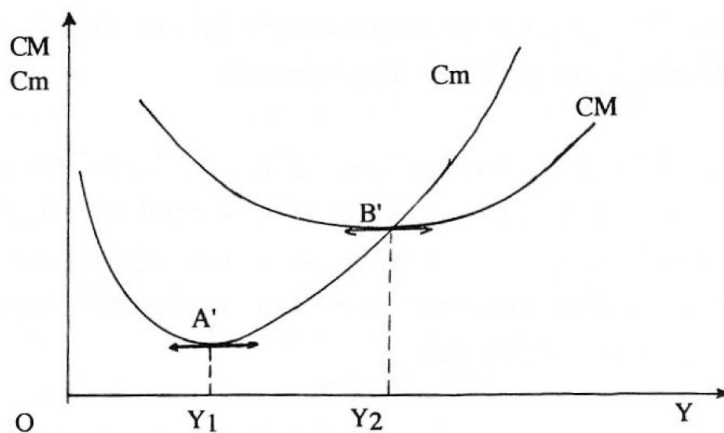
Dès lors, on peut distinguer deux situations :

- Si le prix du marché est supérieur au coût marginal, l'entreprise a intérêt à produire davantage pour pouvoir vendre davantage, et ceci fait augmenter le coût marginal.
- Si le prix du marché est inférieur au coût marginal, l'entreprise fait alors des pertes puisque la production d'une unité supplémentaire du bien lui coûte plus cher que cela ne lui rapporte. Dans ce cas, elle a intérêt à diminuer

la quantité produite pour que le coût marginal diminue et s'ajuste au prix du marché.

La courbe d'offre du producteur représente la quantité offerte pour chaque niveau de prix possible d'un bien. En fonction de l'hypothèse de la loi des rendements décroissants, la courbe d'offre est croissante en fonction du prix.

Ainsi, la courbe d'offre du producteur est donnée par la partie de la courbe de coût marginal située au-dessus de la courbe de coût moyen. La courbe de coût marginal coupe la courbe de coût moyen en son minimum.



La théorie microéconomique retient également la notion d'*élasticité de l'offre* : elle se définit comme la sensibilité de l'offre consécutive à une variation des prix d'offre.

L'élasticité-prix de l'offre aura tendance à être positive si, à mesure que les prix augmentent, le producteur aura tendance à accroître sa production.

- Ainsi une offre élastique signifie que l'offre réagit très fortement aux variations de prix : ainsi, tout changement de prix en pourcentage entraîne un pourcentage plus élevé de la variation de production offerte.
- À l'inverse, une offre inélastique signifie que tout changement de prix en pourcentage entraîne une variation en pourcentage plus faible de la production.

Le niveau d'élasticité de l'offre est déterminé par différents facteurs : si l'entreprise dispose de stocks excédentaires pour augmenter facilement les quantités offertes (ou non) ; le

temps qu'elle met à réagir aux variations de prix ; si les travailleurs sont disponibles rapidement sur le marché du travail.

■ La courbe d'offre du marché

La courbe d'offre du marché est construite par agrégation des courbes individuelles d'offre (parfois appelée courbe d'offre agrégée). À chaque niveau de prix, elle fait correspondre la somme des offres de tous les producteurs sur le marché.

Sur la courbe d'offre, il faut distinguer d'une part le *déplacement le long de la courbe d'offre* (lorsque la quantité offerte se modifie en fonction du prix) et d'autre part le *déplacement de la courbe d'offre* provoqué par un choc d'offre (technologie et/ou prix des facteurs).

LES INTERDÉPENDANCES ENTRE AGENTS ET OPÉRATIONS ÉCONOMIQUES

« La masse totale du produit annuel de la terre et du travail d'un pays, ou, ce qui revient au même, la somme totale du prix de ce produit annuel, se divise naturellement, comme on l'a déjà observé, en trois parties : la rente de la terre, les salaires du travail et les profits des capitaux, et elle constitue un revenu à trois différentes classes du peuple : à ceux qui vivent de rentes, à ceux qui vivent de salaires et à ceux qui vivent de profits. Ces trois grandes classes sont les classes primitives et constituantes de toute société civilisée, du revenu desquelles toute autre classe tire en dernier résultat le sien. » - Adam Smith, 1776

« Le fait seul de la formation d'un produit ouvre, dès l'instant même, un débouché à d'autres. » - Jean-Baptiste Say, 1803

« La rente est la portion du produit de la terre que l'on paye aux propriétaires pour avoir le droit d'exploiter des facultés productives et impérissables du sol. » - David Ricardo, 1817

« La macroéconomie traite du tableau d'ensemble – à savoir des “macro-agrégats” constitués par le revenu national, le volume de l'emploi et les niveaux généraux des prix. » - Paul Anthony Samuelson, 1948

SOMMAIRE

[4.1 Agents économiques et opérations économiques](#)

[4.2 La production](#)

[4.3 Mesurer la production](#)

[4.4 Le PIB : une mesure de la production globale](#)

[4.5 Les limites du PIB en tant qu'indicateur de production globale](#)

[4.6 Les indicateurs complémentaires du PIB](#)

[4.7 La répartition des revenus](#)

[4.8 Des inégalités de revenus primaires à la redistribution](#)

[4.9 L'évolution du partage de la valeur ajoutée entre les facteurs de production](#)

4.1 AGENTS ÉCONOMIQUES ET OPÉRATIONS ÉCONOMIQUES

Commençons par définir des catégories d'agents économiques afin de savoir « qui fait quoi » : quelles sont les fonctions économiques spécifiques de chacun ? Quels sont les échanges qui en découlent ? En quoi ces différents échanges créent-ils des interdépendances ?

4.1.1 *Les agents économiques*

Un **agent économique** est soit un individu fictif, soit un groupe d'individus faisant des choix économiques, sur la base d'un certain degré d'autonomie dans la prise de décision. Le terme d'agent économique étant un terme générique qui peut désigner des réalités très hétérogènes, les économistes et statisticiens vont procéder à des **regroupements** sur la base de divers critères permettant de constituer des groupes à forte homogénéité interne, intragroupe et à forte hétérogénéité externe, intergroupes. Le principal critère permettra de regrouper les agents économiques en fonction de leur **activité économique principale**.

De cette manière, les millions d'opérations économiques que réalisent quotidiennement les millions d'agents économiques seront plus lisibles, plus compréhensibles...

Cette simplification du discours théorique et de la représentation statistique est indispensable pour pouvoir observer et analyser la réalité complexe. Cette simplification n'est pas une abstraction pure, dans la mesure où chacun d'entre nous ne fait pas tout lui-même pour satisfaire ses besoins. Nous l'avons vu, les spécialisations sont inévitables et rendent nécessaires les échanges, consubstantiels des interdépendances.

4.1.1.1 Des unités institutionnelles aux secteurs institutionnels

En France, l'**Insee**³, dans le cadre du SCN (Système de Comptabilité Nationale) et conformément au SEC 2010 (Système Européen de Comptabilité d'Eurostat, qui s'impose à tous les pays de l'Union européenne), **va définir ces différentes catégories d'agents et d'opérations économiques**. Pour commencer, au sens comptable de l'Insee, un agent économique est une **unité institutionnelle (UI)**. Celle-ci correspond à l'élément de base du système. *« Les unités, dites institutionnelles, constituent les unités de base de la comptabilité nationale : ce sont des unités susceptibles de posséder elles-mêmes des actifs, de souscrire des engagements, de s'engager dans des activités économiques et de réaliser des opérations avec d'autres unités. »* (SEC 2010) L'UI a pour caractéristique principale d'être un **centre de décision autonome dans l'exercice de sa fonction principale**. L'unité institutionnelle est **résidente** si elle exerce des activités économiques sur le territoire économique national depuis au moins un an, quelle que soit sa nationalité. Elle est non résidente si ses activités économiques sur le territoire national durent depuis moins d'un an.

Ainsi, l'économie nationale désigne l'ensemble des UI résidentes, y compris, donc, les filiales d'entreprises étrangères ou les travailleurs immigrés exerçant en France depuis plus d'un an.

« Les secteurs institutionnels regroupent les unités institutionnelles ayant des comportements économiques similaires caractérisés par leur fonction principale et la nature de leur activité ». L'Insee définit **six secteurs institutionnels (SI)** : les ménages, les sociétés non financières (SNF), les sociétés financières (SFi), les administrations publiques (APU), les institutions sans but lucratif au service des ménages (ISBLSM) et le reste du monde, soit cinq SI résidents et un SI non résident.

4.1.1.2 Les six secteurs institutionnels

Le premier critère de classement des unités institutionnelles en SI est **la fonction économique principale**. Le second critère correspond à la ressource principale de l'unité économique.

– Pour l'Insee, un **ménage**, au sens du recensement de la population, désigne l'ensemble des personnes qui partagent la même résidence principale, sans que ces personnes soient nécessairement unies par des liens de parenté. Un ménage peut être constitué d'une seule personne. Il y a égalité entre le nombre de ménages et le nombre de résidences principales. Le ménage est un centre de décision économique dans l'exercice de sa fonction principale : la consommation finale de biens et services. Autrement dit, l'autre partie du revenu est consacrée à l'épargne. Il y a aussi la fourniture de facteurs de production (force de travail et capitaux) à d'autres SI. Retenons que ce SI est spécifique dans la mesure où tous les individus d'une société sont membres d'un ménage !

Ainsi, un célibataire, un couple marié ou non, pacsé ou non, avec ou sans enfants, des colocataires... sont des ménages. Dans le cas d'entrepreneurs, les membres d'un ménage peuvent produire des biens (marchands) ou des services financiers et non financiers marchands. Les personnes vivant dans des habitations mobiles comme les bateliers, les sans-abri, et les personnes vivant en communauté (foyers de travailleurs, maisons de retraite, résidences universitaires, maisons de détention...) sont considérées comme vivant hors ménage.

Pourtant, les prisonniers, les pensionnaires d'une maison de retraite... constituent un seul ménage (SEC 2010 pages 50 et 51).

– Les **sociétés non financières (SNF)** sont des sociétés et quasi-sociétés⁴ dont la fonction principale consiste à produire des biens ou des services non financiers marchands pour les autres agents économiques, mais aussi pour elles-mêmes. Les opérations de répartition et

les opérations financières sont séparées de celles de leurs propriétaires.

– Les **sociétés financières (SFi)** sont des sociétés et quasi-sociétés dont l'activité principale consiste à produire des services financiers marchands. On y retrouve les assurances, les banques et les autres établissements de crédit, les caisses d'épargne, les organismes de placement collectif en valeurs mobilières (Sicav, fonds communs de placement), la banque centrale et le Trésor public. La fonction principale des sociétés financières est d'assurer le financement de l'économie. Pour cela, elles jouent plusieurs rôles complémentaires : un rôle d'intermédiaire entre les agents à capacité et à besoin de financement ; un rôle de transformation de l'épargne des ménages, souvent disponible à court terme, en ressources disponibles à long terme pour les agents à besoin de financement (sociétés et APU) ; enfin, un rôle de création de la monnaie via le crédit nécessaire au fonctionnement de l'économie⁵.

– Les **administrations publiques (APU)** regroupent toutes les organisations dont l'activité principale consiste à produire des services non marchands pour les UI de tous les SI ou à redistribuer les revenus (indirectement les productions) entre tous les SI. Les APU sont principalement financées par des prélèvements obligatoires (taxes, impôts et cotisations sociales) et versent en contrepartie des prestations sociales.

On distingue **trois niveaux d'APU** : les administrations publiques centrales (**APUC** : État central avec présidence de la République, ministères, parlement... ; l'État déconcentré (comme les préfectures) et les administrations publiques locales (**APUL**) ou territoriales (commune avec le conseil municipal, département avec le conseil départemental, région avec le conseil régional). Ce sont les différentes lois de décentralisation qui fixent les compétences des unes et des autres, sous contrainte du principe de subsidiarité. Au sens large, quand on désigne l'État, on fait référence à l'ensemble des APU. Au sens

étroit, l'État correspond aux APUC. Il y a enfin les administrations de sécurité sociale (**ASSO**), qui assurent le fonctionnement du système de protection sociale des individus contre les différents risques sociaux.

Les administrations publiques centrales (APUC) regroupent l'État au sens strict : Gouvernement, Parlement, Magistrature et les organismes divers d'administration centrale (ODAC) qui sont une catégorie d'établissements publics, au sens du SCN, qui regroupe des organismes auxquels l'État a donné une compétence fonctionnelle spécialisée au niveau national. Ces ODAC sont contrôlés et financés majoritairement par l'État, et ont une activité principalement non marchande. Les ODAC peuvent être regroupés en dix grandes fonctions : services généraux des administrations publiques, défense, ordre et sécurité publics, affaires économiques, protection de l'environnement, logement et équipements collectifs, santé, loisirs, culture et culte, enseignement et protection sociale.

Les administrations publiques locales (APUL) regroupent les **collectivités locales**, les collectivités territoriales à compétence générale (communes, départements et régions), les groupements de communes à fiscalité propre (communautés urbaines, communautés d'agglomération et communautés de communes) et certaines activités des syndicats de communes.

Les **organismes divers d'administration locale** (ODAL) sont des établissements publics locaux : centres communaux d'action sociale (CCAS), caisses des écoles, services départementaux d'incendie et de secours (SDIS),... les établissements publics locaux d'enseignement (collèges, lycées d'enseignement général et professionnel même si les salaires des enseignants sont versés par l'État et ne constituent donc pas une dépense des administrations publiques locales), les associations récréatives et culturelles financées majoritairement par les collectivités territoriales, les chambres consulaires (commerce et industrie, agriculture et métiers).

Source : Insee

– Les **institutions sans but lucratif au service des ménages (ISBLSM)** correspondent à l'ensemble des unités privées dotées de la personnalité juridique dont la fonction principale est de produire des biens et services non marchands au profit des ménages. Ce sont pour l'essentiel des associations, les églises, mosquées, synagogues, temples... les partis politiques et les syndicats. Ces ISBLSM sont principalement financées par contributions volontaires en espèces ou en nature effectuées par les ménages en leur qualité de consommateurs, de versements provenant des administrations publiques, ainsi que de revenus de la propriété.

– Le **reste du monde (RDM)** est un agent encore plus abstrait, dans la mesure où il regroupe tous les agents économiques non-résidents, qui ne sont pas membres des cinq premiers SI. Ce SI permet de définir et mesurer l'ensemble des échanges entre les agents résidents et les agents non résidents (ceux du RDM). Il y a les échanges

de biens, de services, mais aussi les échanges monétaires et financiers.

Cette nomenclature des secteurs institutionnels est nécessaire, mais ce regroupement des unités institutionnelles en secteurs institutionnels ne suffit pas. Encore faut-il regrouper les millions d'opérations économiques, afin de pouvoir mieux visualiser les interdépendances liées aux échanges économiques divers entre ces catégories d'agents.

4.1.2 *Les opérations économiques*

L'Insee distingue trois catégories d'opérations économiques : les opérations sur les produits, les opérations de répartition et les opérations financières.

4.1.2.1 Les opérations sur les produits

Les opérations sur les produits concernent la production et l'utilisation des biens et services marchands ou non marchands pour satisfaire des besoins divers.

- La **production** a été très largement définie précédemment. Sous contrainte de rareté, elle consiste à produire les biens et services nécessaires à la satisfaction des besoins de tous les agents économiques, issus de tous les SI. Les différentes productions donnent lieu à de nombreux échanges entre les SI.
- **L'investissement** productif désigne l'opération d'acquisition de capital fixe productif par les entreprises (et sociétés) et les APU. L'Insee le mesure à partir de la formation brute de capital fixe (FBCF). La FBCF privilégie surtout les biens matériels (acquisition de machines, la construction de logements, les travaux publics. Cependant, les dépenses de R&D sont comptabilisées dans la FBCF depuis le SEC 2010). L'essentiel de la FBCF est réalisé par les entreprises sous forme de biens d'équipement. Cela dit, l'investissement immatériel (logiciels) occupe une place croissante. Il ne faut pas oublier les investissements directs à l'étranger (IDE) des entreprises, qui accroissent

la taille des firmes multinationales. Enfin, l'investissement immobilier des ménages se retrouve dans cette catégorie⁶.

- **Les consommations intermédiaires** correspondent à la valeur des biens et services transformés ou entièrement consommés au cours du processus de production.

- La **consommation finale** des ménages correspond à l'ensemble des achats de biens et services pour satisfaire leurs divers besoins. Depuis, le système européen de comptabilité de 1995 (SEC 95), la comptabilité nationale (SCN) calcule deux agrégats pour la consommation des ménages : la dépense de consommation finale (DCF) et la consommation finale effective (CFE). La dépense de consommation concerne les dépenses directes des ménages. On y retrouve la part des dépenses de santé, d'éducation, de logement restant à leur charge, après remboursement éventuel. On y inclut les loyers imputés, c'est-à-dire les loyers que les ménages propriétaires de leur résidence principale se versent à eux-mêmes. Les loyers imputés sont la contrepartie de la production de service de logement des propriétaires qui occupent leur logement : on parle de production pour emploi final propre (PEFP) des ménages⁷.

La consommation effective des ménages concerne l'ensemble de leur consommation de biens et services. La part socialisée de la consommation, telle que les consommations dites individualisables (santé et éducation) au sens où le bénéficiaire peut en être précisément défini, y est comprise. Faire un détour par les Dépenses de consommation individuelle et collective des APU.

En revanche, les dépenses des APU pour la défense nationale et la sécurité intérieure ne sont pas incluses dans la consommation effective des ménages, car elles sont non individualisables.

- Les **exportations** et les **importations** de biens et services constituent le commerce extérieur. Les exportations désignent l'ensemble des biens et des

services vendus (pour l'essentiel) par des résidents à des non-résidents. Les importations correspondent à l'ensemble des biens et des services achetés (pour l'essentiel) par des résidents à des non-résidents. Les importations sont des achats. Les exportations sont des ventes. Logiquement, pour pouvoir importer-acheter, il faut avoir vendu-exporté. Dans le cas contraire, il est nécessaire que les résidents s'endettent vis-à-vis des non-résidents.

4.1.2.2 Les opérations de répartition

Les opérations de répartition décrivent l'origine, la distribution et l'utilisation finale des revenus primaires issus de la production, ainsi que la redistribution des revenus de transfert. Comme nous l'avons vu, la répartition se déroule en deux temps : la répartition primaire des revenus entre les facteurs de production et la répartition secondaire des revenus via le processus de redistribution. Le processus de redistribution permet de passer du revenu primaire au revenu disponible.

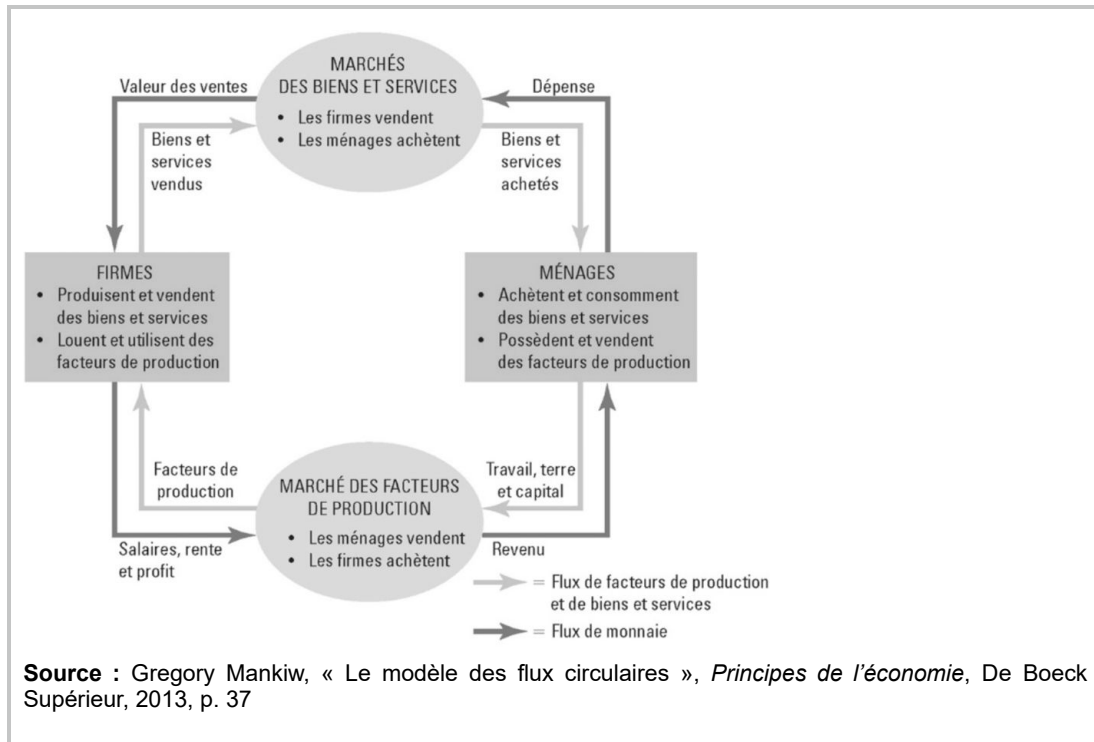
4.1.2.3 Les opérations financières

Les opérations financières, décrites de manière détaillée par ailleurs, correspondent à l'ensemble des opérations qui mettent en relation directement (via les marchés de capitaux) ou indirectement (via des intermédiaires financiers) les agents à capacité de financement et les agents à besoin de financement (*cf.* chapitre 6).

4.1.3 Synthèse des interdépendances entre agents économiques et opérations économiques

Dans la représentation qui suit, proposée par Gregory Mankiw, nous observons aisément les interdépendances entre agents économiques via les opérations économiques qu'ils réalisent entre eux sur différents marchés. Pourquoi en est-il ainsi ? Tout simplement parce que la logique des

échanges sur les différents marchés est une illustration réelle de ces interdépendances. Ainsi, dans le cas d'une situation économique de plein emploi, les revenus des ménages et des entreprises augmentent, ce qui est bon pour les uns et les autres. Logiquement, au niveau du **plein emploi, sans accroissement de la productivité, la production et les revenus stagnent.**



GRAPHIQUE 4.1.

Le circuit économique simplifié, une modélisation de ces interdépendances

Les recettes fiscales et sociales des administrations publiques sont élevées. Il est possible de baisser les taux, sans pour autant baisser les recettes, et les dépenses peuvent ne pas baisser, voire augmenter. Simultanément, il est plus facile de réduire les déficits publics et la dette. Le bonheur des uns fait le bonheur des autres, et réciproquement. *A contrario*, dans une situation de croissance molle et de chômage élevé, les revenus des ménages et des entreprises ralentissent ou baissent. Les recettes fiscales et

sociales des administrations publiques diminuent et il est alors plus difficile de réduire les déficits publics et la dette.

Au niveau macroéconomique, les conditions de l'offre et de la demande sont inextricablement liées. La santé économique et financière des entreprises est dépendante de celle des ménages et de celle des administrations publiques, et vice versa. Les situations de l'économie à court terme et à long terme sont réciproquement dépendantes. Les économies ouvertes le sont également, car les exportations/ventes des uns sont les importations-achats des autres...

4.2 LA PRODUCTION

4.2.1 *Définition générale*

La production est un concept générique qui permet de designer des réalités multiples et hétérogènes. Commençons donc par préciser les sens de ce terme courant, afin de mieux les comprendre et d'en reconnaître les différentes formes.

4.2.2 *Les différents sens de la notion de production*

Au sens courant, la production est une **activité humaine et socialement organisée** (par un ensemble de règles diverses) consistant à fabriquer, en utilisant et en transformant des ressources naturelles ou d'autres produits intermédiaires, des biens ou des services. En ce sens, toute activité humaine de ce type est une activité productive.

4.2.2.1 **De la production domestique...**

Lorsqu'une personne cultive des légumes dans son jardin pour sa propre consommation, elle produit des biens. Elle produira un service lorsqu'elle préparera son repas. De surcroît, un cours d'économie donné par David au fils de Laurent n'est pas une production au sens économique, car David ne sera pas rémunéré par Laurent.

4.2.2.2 **... à la production au sens économique**

Dans ce cas, pour qu'il y ait production, il faut qu'il y ait une référence directe ou indirecte au marché. Autrement dit, il conviendra que le produit s'échange sur un marché, ou qu'il soit produit à partir de facteurs de production (du travail...) rémunérés sur un marché. Un cours d'économie donné par David ou Laurent dans un lycée public ou une école privée est une production au sens économique. Ils sont logiquement rémunérés pour ce service de haute qualité. Pour autant, cette rémunération ne transite pas par un marché puisqu'elle est la contrepartie de prélèvements obligatoires. Cette rémunération n'est donc pas liée directement à la confrontation entre une offre et une demande, mais à l'obtention d'un concours.

Dans ces conditions, la production économique est une activité humaine socialement organisée consistant à fabriquer ou produire des biens ou des services s'échangeant sur un marché ou produits avec des facteurs de production rémunérés sur un marché.

Retenons encore que si la production est une activité, nous verrons un peu plus loin, via différents instruments de mesure, qu'elle est aussi le résultat de cette activité.

4.2.3 *Les différents types de production*

Concrètement, la production résulte d'abord de l'activité d'une grande multitude d'unités de production, aux activités très diverses, qui peuvent être des entreprises privées ou publiques, des administrations publiques...

- **La production d'une unité de production**

- **La production globale** désigne l'ensemble des productions de biens et services de toutes les unités de production qui est destiné à satisfaire les besoins de tous les agents économiques.

4.2.4 *Les biens et les services*

On peut utiliser trois types de critères pour les distinguer. Les produits sont-ils matériels et stockables ? Sont-ils produits et

consommés simultanément ou non ?

Les **biens** sont des produits matériels et stockables. Ils sont produits et consommés à différents moments. Il peut donc exister un décalage temporel entre leur production et leur utilisation. Par exemple, une bouteille de cidre doux et un smartphone sont des biens pour les membres d'un ménage donné. Un verre de cidre dégusté dans une crêperie, un abonnement Internet ou téléphonique sont des services.

Les **services** sont immatériels, non stockables. Leur production et leur utilisation sont simultanées. La crêpe servie à la crêperie relève du service, car on considère que ce qui est consommé ce n'est pas la simple crêpe, mais un service global de restauration. La crêpe n'est qu'une des composantes de ce service. Le lait, la farine, les œufs, l'électricité, etc. sont des consommations intermédiaires, c'est-à-dire des biens ou services qui sont utilisés au cours du processus de production du service de restauration que propose la crêperie.

4.2.5 *Productions marchandes et productions non marchandes*

Par convention, tous les biens sont marchands. Ils sont échangés sur des marchés à des prix de marché. Le terme « biens collectifs » désignera en fait des services collectifs.

Les services marchands sont également échangés sur un marché à des prix de marché. Conventionnellement, la production marchande est vendue (donc achetée) à un prix économiquement significatif, c'est-à-dire un prix couvrant plus de 50 % des coûts de production. L'échange marchand permet de déterminer la valeur monétaire des produits échangés.

Attention : En toute rigueur, il faudrait parler de services principalement marchands, car pour certaines activités coexistent des parties marchandes et non marchandes ; certains services sont considérés comme toujours marchands (par exemple, les transports), d'autres comme toujours non marchands (par exemple, l'administration générale).

Source : Insee

Les services non marchands n'ont pas de prix de marché. Ils ne sont donc ni vendus ni achetés. On considère qu'une unité produit des services non marchands lorsqu'elle les fournit gratuitement ou à des prix qui ne sont pas économiquement significatifs. Ces activités de services se rencontrent dans les domaines de l'éducation, de la santé, de l'action sociale et de l'administration.

Mais ils ne sont pas gratuits. Rien n'est jamais gratuit en économie. Comme le dit l'adage des économistes, « there is no free lunch » (il n'y a pas de repas gratuit). Les services non marchands sont financés par les impôts versés par la collectivité (par l'ensemble des UI ou des SI, excepté les APU, aux APU), car tout le monde a intérêt à ce que tout le monde y ait accès. Chacun a intérêt, ici, à ce que le besoin de tous soit satisfait. Comme la « main invisible » du marché ne permet pas de faire converger les intérêts individuels et l'intérêt collectif, c'est à la main bien visible de l'État de réparer ces défaillances de marché.

Sur la distinction entre production non marchande, biens collectifs et biens publics, on pourra se reporter au chapitre 7.

4.2.6 *Biens de consommation finale, biens de production, biens de consommation intermédiaire, biens d'équipement*

Les biens de consommation finale (CF) concernent les ménages et la satisfaction finale de leurs besoins. La consommation est dite finale, car le bien est détruit, plus ou moins rapidement, avec la consommation. Ainsi, la crêpe sera détruite plus rapidement que le contenu de la bouteille de Jack Daniel's, fort heureusement. De surcroît, les ménages consomment des biens durables qui relèvent de la DCF comme un congélateur, une télévision, un ordinateur, une voiture...

Les biens de production prennent la forme de biens d'équipement (four du boulanger, camion du transporteur routier...) ou encore de biens de consommation intermédiaire.

Les biens de consommation intermédiaire (CI) sont transformés ou utilisés au cours du processus de production dans une entreprise ou une administration publique. Ils ne sont utilisés qu'une seule fois. Ainsi la farine, le lait, les œufs... sont des CI pour la crêperie, mais des biens de CF pour le mangeur de crêpes.

Un même bien peut donc faire l'objet d'une consommation finale pour un ménage ou d'une consommation intermédiaire pour une entreprise. Au niveau des entreprises, les biens de production désignent des biens ou services utilisés au cours du processus de production.

Parmi eux, les biens d'équipement correspondent à du capital fixe qui sera utilisé plusieurs fois (comme la plaque pour faire cuire les crêpes de la crêperie), alors que les biens ou services de consommation intermédiaire correspondent à du capital circulant et ne seront utilisés qu'une seule fois.

4.2.7 Les différentes catégories d'organisations productives

La production est assurée dans le cadre d'**organisations productives**.

Il s'agit d'organisations dans la mesure où un ensemble de personnes ont des objectifs communs et vont mettre en œuvre divers moyens (humains, techniques, financiers...) pour les atteindre. Il y a organisation, car l'activité est encadrée par des règles, une hiérarchie... Une organisation productive est simplement une organisation particulière caractérisée dont l'objectif est de produire des biens ou des services pour satisfaire des besoins divers.

Parmi les organisations productives, on peut rencontrer des entreprises publiques ou privées, des administrations publiques, des associations.

On utilise trois **critères** principaux pour distinguer les organisations productives : le **statut juridique**, qui est privé ou public, le **caractère marchand ou non marchand** de la production, l'**objectif lucratif ou non** de l'activité.

Il convient de ne pas confondre marchand et lucratif. En effet, les organisations lucratives (dont l'objectif principal est la recherche du profit) sont nécessairement marchandes. En revanche, il existe des organisations privées marchandes non lucratives, comme les mutuelles ou les coopératives de production. Les administrations publiques sont à but non lucratif et non marchandes. Les entreprises privées, à but lucratif, sont marchandes. Les entreprises peuvent être privées ou publiques, à but lucratif ou à but non lucratif.

Toutes les organisations productives marchandes sont des entreprises. Une entreprise est dite publique, non pas en fonction du caractère marchand de sa production, mais en fonction de l'identité du propriétaire principal : l'État. À titre d'exemples, les grandes entreprises publiques sont classées dans le SI des SNF.

Une association est une organisation productive privée à but non lucratif. On perçoit intuitivement, au travers de ces quelques définitions, que de nombreuses interactions, interdépendances et échanges, du fait des diverses spécialisations, existent entre ces différentes organisations productives et ces différentes opérations productives.

4.3 MESURER LA PRODUCTION

Comme toute opération économique, la production doit être mesurée. La mesure est au cœur de la construction des « faits économiques » qui n'existent pas spontanément, qui ne se révèlent pas directement et par nature. Les faits économiques correspondent tout simplement à des statistiques fabriquées sur la base de nombreuses conventions et critères.

4.3.1 *La production d'une entreprise est une production marchande*

Pour commencer, on pourra mesurer la production d'une entreprise par son **chiffre d'affaires** (CA), avant de calculer sa valeur ajoutée. En fait, la mesure la plus pertinente de la

production d'une unité de production est la valeur ajoutée. Il faut distinguer la production de la valeur ajoutée (VA). La VA n'est pas la mesure de la production, mais la mesure de la création de richesse liée à l'activité de production.

Le CA n'est qu'une approximation de la mesure de la création de richesse. Mais il mesure la valeur de la production, car il implique une double comptabilisation de certaines productions. On obtiendra :

Chiffre d'affaires = valeur des ventes réalisées par une entreprise

Chiffre d'affaires = prix de vente unitaire x nombre de produits vendus

Soit $CA = p \times Q$

Valeur ajoutée = valeur de la production (CA) – valeur des consommations intermédiaires

4.3.2 *La production d'une administration publique est une production non marchande*

La production d'une administration publique est, comme nous l'avons indiqué, évaluée de manière marchande, mais indirectement, car elle n'est pas échangée sur un marché. Il n'y a donc ni prix, ni acheteurs, ni vendeurs, mais elle n'est pas gratuite. Comme la production non marchande n'a pas de prix, même si elle n'est pas gratuite, il a fallu trouver un moyen pour la mesurer. La convention veut qu'elle soit mesurée par le **coût des facteurs** de production qui, eux, s'échangent à un prix sur divers marchés. Ainsi, la production d'un fonctionnaire est « mesurée » par sa rémunération. Si son traitement augmente peu, on pourrait en déduire que sa production augmente peu. Quand le ministère de la Fonction publique gèle les traitements de fonctionnaires, peut-on en déduire que leur production stagne ?

4.3.3 De la production en valeur à la production en volume

Cette première mesure, en valeur, de la production n'est pas convenable, dans la mesure où elle intègre le niveau des prix et donc, en évolution, l'inflation. Nous pouvons donc être confrontés à un phénomène d'illusion monétaire ou nominale, dans la mesure où la production et la croissance peuvent être surestimées du fait du niveau plus ou moins élevé et de la hausse plus ou moins forte des prix. Il est donc indispensable de déduire (retrancher, neutraliser...) l'effet prix pour ne conserver que l'effet quantité. En effet, souvenons-nous qu'une économie fonctionne afin de produire les quantités de biens et services destinés à satisfaire les divers besoins de tous les agents économiques en matière de consommation, d'investissement...

Pour le dire simplement, ce qui compte en économie n'est pas le (pQ), ou la valeur de la production, mais le (Q), son volume. Ainsi, un ménage n'est pas plus satisfait au regard de ses besoins quand la valeur de ce qu'il a dans son caddy est deux fois plus élevée. Il le serait plutôt si, pour un prix stable, le volume de marchandise dans son caddy avait doublé.

$$\text{Production en volume} = \frac{\text{Production en valeur}}{\text{Indice des prix}}$$

Nous pouvons retenir les équivalences (en colonnes) suivantes :

Mesure « en volume » = Mesure « en valeur réelle » =
Mesure « en monnaie constante »

Mesure « en valeur » = Mesure « en valeur nominale » =
Mesure « en monnaie courante »

4.3.4 De la production particulière d'une unité de production... à la production globale

Concrètement, la production résulte d'abord de l'activité d'une grande multitude d'unités institutionnelles, aux activités très diverses, qui peuvent être des entreprises privées ou publiques, des administrations publiques...

- **La production d'une unité institutionnelle** va correspondre à la quantité de produits (biens ou services) obtenus au cours d'une période donnée. Sur le plan comptable, il s'agit souvent de l'année.

- **La production globale (ou production macroéconomique)** désigne l'ensemble des productions de biens et services de toutes les unités de production qui est destinée à satisfaire les besoins de tous les agents économiques.

4.3.5 De la production effective à la production potentielle

- **La production effective** correspond à la production réalisée à un moment donné. Mais son niveau ne donne aucune indication de l'efficacité de l'utilisation des facteurs de production et de leur combinaison.

- **La production potentielle** désignera la production maximale possible, compte tenu de l'utilisation totale des facteurs de production, quantitativement et qualitativement. À ce niveau de production, il est impossible de produire plus, d'une manière ou d'une autre. Il n'y a aucun gaspillage de capacité de production.

4.4 LE PIB : UNE MESURE DE LA PRODUCTION GLOBALE

Quels sont les principaux intérêts et les principales limites du PIB en tant qu'indicateur de la richesse créée, autrement dit de la capacité de l'économie à produire les biens et services permettant de satisfaire les besoins de tous les agents économiques ?

4.4.1 Le PIB est un agrégat

Le PIB est un agrégat, donc, par définition, il va faire la somme (agréger) de productions très hétérogènes. Le produit intérieur brut (PIB) permet de mesurer la production globale de l'économie, en faisant plus précisément la somme des valeurs ajoutées de biens et services d'une économie par tous les agents économiques qui ont une activité de production. À travers la mesure de la production globale de biens et services, le PIB permet d'évaluer la capacité de l'économie à satisfaire les besoins de tous les agents économiques (ménages, entreprises).

EN FRANCE, EN 2017, LA VALEUR NOMINALE

- du PIB est de 2291,7 milliards d'euros courants ;
- de la consommation effective des ménages est de 1592,8 milliards d'euros, 69,5 % du PIB :
- dont 1191,0 milliards d'euros pour la dépense de consommation finale, 51,9 % du PIB
- et 353,6 milliards d'euros pour la consommation individualisable des APU, 15,4 % du PIB
- de la consommation collective des APU est de 186,5 milliards d'euros, 8,1 % du PIB ;
- de la FBCF est de 515,9 milliards d'euros, 22,5 % du PIB ;
- des exportations est de 707,7 milliards d'euros, 30,9 % du PIB ;
- des importations est de 733,0 milliards d'euros, 31,9 % du PIB.

Source : Insee Première, Les comptes de la Nation en 2017

Le PIB est un indicateur statique, *stricto sensu*, mais il nous permettra de calculer la croissance économique, indicateur dynamique. On pourra mesurer la croissance économique, et donc l'amélioration, l'augmentation de cette capacité à satisfaire ces besoins. En calculant le taux de variation du PIB, on mesure la croissance économique. Le taux de croissance correspond donc au taux de variation du PIB (cf. chapitre sur la croissance économique).

Le PIB est, par construction, un indicateur quantitatif ! Néanmoins, ces productions peuvent avoir des impacts qualitatifs positifs ou négatifs.

La plupart des pays calculent le PIB. Pour autant, on peut rencontrer d'autres agrégats proches : PNB et RNB.

4.4.2 *Le PIB n'est pas le PNB, ni le RNB, ni le RND*

4.4.2.1 Du PIB au PNB

Depuis le changement de système de comptes intervenu en 1999 (passage au SEC 95 et abandon du SECN), le produit national brut ou PNB n'est plus calculé. La différence entre le PIB et le PNB reposait sur un critère géographique.

Pour le PIB, c'est le critère de la résidence des unités institutionnelles qui était déterminant. Le PIB était la somme des valeurs ajoutées des agents résidents.

Pour le PNB, c'est le critère de la nationalité (peu importe la résidence) qui était déterminant. Le PNB était la somme des valeurs ajoutées des agents « nationaux ».

Pour passer du PIB au PNB, il suffisait d'ajouter au premier les soldes des revenus reçus ou versés au reste du monde. Le **PNB** était la somme des valeurs ajoutées brutes réalisées par les unités institutionnelles françaises (critère de nationalité) sur le territoire métropolitain et à l'étranger : solde des revenus versés et reçus du reste du monde.

PNB = Produit national brut = PIB + revenus de facteurs venant de l'extérieur (du travail, de la propriété, de l'entreprise), c'est-à-dire réalisés à l'étranger par des Français (reçus du reste du monde) – revenus de facteurs versés à l'étranger (du travail, de la propriété, de l'entreprise gagnés en France par des étrangers et transférés dans leur pays (versés au reste du monde) = PIB + ou – solde des revenus de facteurs avec l'étranger.

4.4.3 *Du PIB au PIN : mesures en valeur brute ou en valeur nette*

Pourquoi la production intérieure (comme la valeur ajoutée...) est-elle calculée brute ? Quelle en est la signification ?

Toute production se traduit par l'utilisation d'un stock de capital alimenté par un flux d'investissement qui en augmente

le stock, pendant que son utilisation en réduit le stock (au sens de la capacité productive). Il est tout simplement **très compliqué d'estimer** le degré de dépréciation du stock de capital lié à la consommation de capital fixe, et donc de calculer **la production intérieure nette (PIN)**. Le PIN, ou produit intérieur net, serait un indicateur très très imprécis et discutable. Les difficultés méthodologiques et comptables sont difficiles à surmonter. En effet, la CCF est l'indicateur le moins fiable de la comptabilité nationale. Il existe de très nombreuses conventions permettant de calculer l'usure, la dépréciation des actifs... Pour autant, depuis le SEC 2010, l'Insee publie les comptes en brut et en net.

4.4.4 *Du PIB en valeur au PIB en volume*

Comme indiqué précédemment, au niveau de l'économie globale, il est indispensable de distinguer la production en valeur de la production en volume. Cette obligation s'applique logiquement au PIB.

$$\text{PIB en volume} = \frac{\text{PIB en valeur}}{\text{Indice des prix}}$$

TABLEAU 4.1.
PIB en euros courants et en euros constant en France

	2015	2016	2017
PIB en euros courants	2198,4	2228,6	2291,7
PIB en euros constants	2173,7	2199,2	2246,7

Source : Insee, Informations Rapides, 15 mai 2018

4.4.5 *PIB effectif et PIB potentiel*

Enfin, arrêtons-nous un instant sur des concepts incontournables qui conduisent à une distinction cruciale entre PIB effectif et PIB potentiel. Cette distinction est fondamentale, dans la mesure où elle a une portée majeure

quand on envisage la situation de l'économie à un moment donné et les décisions qui doivent être prises en vue de mieux préparer l'avenir pour affronter la rareté. Le **PIB effectif** correspond à la quantité de biens et services produits par les facteurs de production utilisés, quantitativement et qualitativement, à un moment donné. Toute sous-utilisation de ceux-ci constitue un gaspillage de ressources productives et une réduction de la capacité à satisfaire les besoins des agents économiques. En effet, si le taux de chômage est trop élevé, si le taux d'emploi de la population en âge de travailler est trop faible, si l'utilisation du capital productif est trop faible..., l'économie s'éloigne alors de sa frontière des possibilités de production et de sa capacité maximum (du moment) à pouvoir satisfaire les agents de l'économie. Ce **gaspillage** de ressources productives va provoquer un écart de production (*output gap*) avec le PIB potentiel estimé au même moment. Le **PIB potentiel** correspond à la production maximum possible, quantitativement et qualitativement, compte tenu de la population en âge de travailler (capacité de production de la quantité de travail disponible) et de son niveau de formation, de la quantité de capital productif et du niveau technologique.

Une économie se trouve alors, dans un premier temps, dans la même situation qu'un sportif qui souhaite, non pas progresser, mais se maintenir à son plus haut niveau du moment. L'objectif est bel et bien de ramener, ou au moins de rapprocher, le PIB à son niveau potentiel en réduisant l'écart de production ou *output gap*. Comment ?

Tout simplement en faisant croître le PIB effectif, voire en accélérant la croissance effective. Pour cela, il va falloir trouver les moyens de faire baisser le nombre de chômeurs et le taux de chômage, d'augmenter le taux d'emploi et le taux d'utilisation des capacités de production.

À court terme, et cela sera développé plus tard, c'est en favorisant la hausse de la demande globale (politiques conjoncturelles...) qu'on pourra initier un tel cercle vertueux, pour autant que l'environnement économique général soit

propice à la consommation, à l'investissement et à l'embauche. Nous verrons un peu plus loin l'importance d'une distinction analogue, qui en découle, entre **croissance effective** et **croissance potentielle**, mais qui se révèle utile et nécessaire.

4.4.6 *Les trois approches du PIB et le circuit économique de base*

Il existe **trois manières de calculer le PIB**. Le résultat sera exactement le même, mais l'angle sera différent, et donc les enseignements aussi.

Pour l'Insee, le produit intérieur brut (PIB) aux prix du marché est un agrégat représentant le résultat final de l'activité de production des unités productrices résidentes. Il peut se mesurer de trois manières :

- Selon une approche **par la production**, le PIB est égal à la **somme des valeurs ajoutées brutes** des différents secteurs institutionnels ou des différentes branches d'activité, augmentée des impôts sur les produits (le principal impôt sur les produits est la TVA) moins les subventions sur les produits (lesquels ne sont pas affectés aux secteurs et aux branches d'activité).
- Selon une approche **par les revenus (primaires)**, le PIB est égal à la **somme des emplois des comptes d'exploitation** des secteurs institutionnels : rémunération des salariés, impôts sur la production et les importations moins les subventions, excédent brut d'exploitation et revenu mixte.
- Selon une **approche par la demande**, le PIB est égal à la **somme des demandes (emplois ou utilisations final(e)s)** intérieures de biens et de services (consommation finale effective, formation brute de capital fixe, variations de stocks), plus les exportations, moins les importations.

Les trois manières de calculer le PIB relèvent de ces trois logiques et correspondent à un modèle économique de base. Le circuit économique de base repose sur ces trois types d'opérations fondamentales. Ce circuit permet de mettre en évidence les interdépendances qui peuvent exister entre agents économiques à travers les opérations économiques qu'ils réalisent entre eux. On obtient l'enchaînement suivant:

La production => (crée) les revenus => (permettent) les dépenses => (donnent vie à) la production...

L'approche par la production (offre) est donc reliée à l'approche par les revenus (rémunération des facteurs) et à l'approche par les dépenses (demande). En effet, la production est obtenue à partir de facteurs de production qui sont rémunérés (revenus) pour leur participation à cette activité de production. Ces revenus seront dépensés (immédiatement ou de manière différée, en cas d'épargne) pour acheter les biens et services produits.

Le PIB, indicateur de production globale, est donc un indicateur macroéconomique dont les composantes ou déterminants macroéconomiques (ou moteurs de la croissance économique) apparaissent dans la dernière formule.

La valeur du PIB est la même, quel que soit le mode de calcul : production, revenu, demande.

4.4.7 Le PIB et équilibre ressources = emplois (ERE) en biens et services pour une économie

Retenons déjà que la troisième manière de calculer le PIB nous permettra de mettre en évidence « l'équilibre (ressources = emplois) en biens et services pour une économie ».

Dans le cas d'une économie ouverte, dans une version simplifiée (sans variation de stock...) de la formule, nous

obtenons ceci :

$$\text{PIB} + \text{M} = \text{C} + \text{I} + \text{G} + \text{X}$$

Le point 17.1.1 permet de revoir l'équation des déséquilibres sans distinction entre agents publics et agents privés

Cette relation est une identité, égalité toujours vérifiée, qui nous permet de répondre à la question simple suivante : dans une économie, quelles sont les origines des produits et quelles en sont les utilisations ?

Cet équilibre, qui est une identité comptable, nous permet aussi de comprendre quels sont les moteurs (déterminants, composantes, macroéconomiques) du PIB et la différence entre égalité ex ante et ex post quand on aborde l'analyse macroéconomique.

Tout ceci étant précisé, nous ne devons pas oublier que le PIB est cependant un indicateur de production globale (et donc de croissance économique limitée), imprécis, incomplet. Nous devons donc présenter un peu plus loin des indicateurs complémentaires, bien plus qu'alternatifs ou substituables au PIB.

4.5 LES LIMITES DU PIB EN TANT QU'INDICATEUR DE PRODUCTION GLOBALE

Maintenant que les diverses manières de mesurer le PIB sont présentées, venons-en aux limites du PIB en tant qu'indicateur de la production globale, afin de voir dans quelle mesure le PIB remplit sa mission, qui consiste à mesurer la capacité de l'économie à produire les biens et services permettant d'améliorer la satisfaction des besoins des agents économiques.

4.5.1 *La production domestique, légale, n'est pas mesurée*

Elle est, en effet, très difficilement mesurable. Comment mesurer, comment donner un prix, une valeur marchande à la préparation d'un repas et à la satisfaction qu'on peut en retirer ? Il en est de même pour le ménage, le repassage, le bricolage, les soins domestiques... Il s'agit bien d'une production au sens courant, mais elle n'est pas comptabilisée dans le PIB, car ce n'est pas une production au sens économique, elle n'utilise pas de travail rémunéré et n'est pas vendue sur un marché... Selon Alfred Sauvy, démographe et économiste français, cela peut poser quelques problèmes, car, *ceteris paribus* (toutes choses égales par ailleurs), le changement de cadre pour une même production (un service de ménage marchand ou le ménage fait par une personne chez elle pour son propre confort) se traduit par une hausse ou une baisse de la production globale. Ainsi, il suffit que « Monsieur épouse sa femme de ménage » pour faire baisser le PIB, ou que « Madame épouse son chauffeur » (selon la formule d'Alfred Sauvy). Le travail marchand de la femme de ménage qui était compté dans le PIB, car déclaré comme travail salarié, devient alors du travail domestique, non compté dans le PIB.

4.5.2 *Les productions illégales ne sont pas mesurées*

De surcroît, le PIB ne permet pas de mesurer les activités de production illégales, celles de l'économie informelle, souterraine, ou économie de l'ombre. Il s'agit de l'ensemble des activités productives qui échappent au contrôle de l'État et qui n'apparaissent pas dans les statistiques officielles. Les activités illégales regroupent le « **travail clandestin** » ou « travail au noir » et les **trafics divers** (drogue, prostitution...). **Cette limite est encore plus sensible pour les pays en voie de développement**, où la part de la production domestique et du travail clandestin est plus élevée que dans les pays développés à économie de marché (PDEM).

Des corrections ont récemment été apportées, à la demande d'Eurostat, pour les pays européens. Cela s'est alors traduit pas une hausse sensible de certains PIB, comme en Espagne, avec la prise en compte de la prostitution... La prise en compte, officiellement, des revenus d'activités illégales (trafics de drogue et prostitution) dans le calcul du PIB serait inscrite dans le nouveau système européen des comptes nationaux (SEC 2010) publié par Eurostat, l'institut statistique communautaire. Ce nouveau standard européen est en vigueur depuis septembre 2014.

Ainsi, des pays comme l'Autriche, l'Estonie, la Finlande, la Norvège, la Slovaquie ou la Suède, l'Italie, la Grande-Bretagne et l'Espagne intègrent les revenus de la prostitution ou de la drogue dans l'activité économique.

En France, l'Insee précise que les activités économiques illégales ne peuvent être comptabilisées qu'à partir du moment où toutes les unités concernées y participent d'un commun accord. Par principe, l'Insee ne comptabilise pas les activités dans lesquelles le consentement des parties n'est pas clairement établi (prostitution).

Le problème est que les comptes nationaux doivent respecter le SEC (pour être en conformité avec les normes comptables européennes), afin d'harmoniser leurs calculs et de pouvoir faire des comparaisons fiables et calculer la contribution au budget de l'UE.

4.5.3 Les externalités sont mal prises en compte dans le calcul du PIB

Le PIB augmente sous l'effet de productions entraînant des externalités positives ou négatives

Les externalités ou effets externes, qui peuvent être positifs ou négatifs, correspondent aux effets des activités-échanges entre deux ou plusieurs agents économiques sur d'autres qui n'y participent pas directement (sans contreparties financières). Dans le cas d'externalités négatives, il en résulte que le PIB augmente grâce à des productions qui permettent

de satisfaire les besoins des uns, mais en dégradant la satisfaction des autres, ou des mêmes d'ailleurs. Ces effets externes négatifs sont mal pris en compte dans le PIB. Ce type de situation n'est donc pas un optimum de Pareto.

Par exemple, les nuisances dues à la production ne sont pas prises en compte, dans la mesure où le PIB augmente alors même que, toutes choses égales par ailleurs, la satisfaction globale baisse.

Il en est ainsi de la production d'une usine chimique polluante – donc un enrichissement pour le pays, ce qui renvoie à un aspect positif (un gain, une création de richesses, de VA), dont les effets en termes de pollution de l'eau ou de l'air sont passés sous silence. Il y a une dégradation de l'environnement, donc, d'une certaine façon, un appauvrissement (aspect négatif), une perte (d'air pur, d'eau potable...) et donc une réduction de la capacité des individus à satisfaire leurs besoins.

De manière analogue, l'augmentation de la production d'alcool et de tabac va faire croître le PIB, mais la dégradation de la santé privée et publique qu'elle entraîne (et qui représente une perte en matière de satisfaction des besoins et de coûts sociaux) ne sera pas prise en compte. Il y a donc augmentation du coût d'opportunité de l'allocation des ressources rares.

Selon la même logique, les externalités positives ne sont pas ou sont mal évaluées. Il en est ainsi des cours payés par un individu qui va ensuite transmettre les connaissances qu'il a accumulées à d'autres personnes, parfois par le simple biais de discussions. Le problème est que cela peut désinciter au financement de formations, car, tel un « passager clandestin », on peut alors bénéficier du voyage sans avoir payé le billet.

Le PIB va permettre de comptabiliser positivement les activités de réparation des externalités négatives

Ici, le PIB augmente quand on répare ce qui a été dégradé du fait d'autres productions, qui devaient améliorer la satisfaction

des besoins. Après avoir compté positivement certaines activités destructrices, les comptes nationaux comptent comme une richesse produite le produit des activités destinées à réparer les effets pervers des externalités négatives.

Par exemple, l'activité de production d'une usine chimique se traduit par la pollution de l'eau et nécessite la construction d'une station d'épuration. Autrement dit, on la comptabilise comme activité productive positive, malgré les nuisances qui en résultent, et on lui rajoute une activité productive de réparation de cette nuisance, comptabilisée également positivement. La production mesurée par le PIB augmente lorsque les destructions augmentent (excepté en période de guerre : le stock de capital et le PIB baissent), car il faut alors accroître la production d'activités réparatrices... Peut-il y avoir, dans ces conditions, un gain net lors des échanges ?

Ainsi, l'installation d'une station d'épuration des eaux sera perçue comme un enrichissement, alors qu'elle ne fait que compenser la pollution de l'eau. Ou encore, une campagne de lutte contre la consommation de tabac ou d'alcool sera considérée comme un enrichissement, alors qu'il ne s'agit que de conséquences de nuisances. Dit autrement, selon le PIB, un pays qui a une eau naturellement pure est moins « riche » qu'un pays où l'eau est purifiée ou dépolluée par l'homme qui l'avait polluée préalablement.

En résumé, plus on produit en dégradant, en nuisant d'une part, plus on doit compenser, réparer d'autre part. Mais, dans les deux cas, le PIB augmente. Or, compte tenu du coût d'opportunité, en situation de rareté, ces deux types de production font augmenter le PIB sans pour autant améliorer *in fine* notre niveau de satisfaction. Il y a donc **surestimation** de la production globale et de notre capacité à satisfaire des besoins en produisant des biens et des services. Et il y a **sous-estimation** de la production globale et de notre capacité à satisfaire des besoins en produisant des biens et des services quand on ne mesure pas l'apport des productions domestiques... Donc, d'une part, on sous-estime

le PIB et donc la croissance du PIB et, d'autre part, on surestime le PIB et la croissance. Pour autant, il ne faut pas jeter le bébé avec l'eau du bain. Autrement dit, il ne faut pas rejeter le PIB en bloc. Le PIB ne veut pas tout dire, mais il ne veut pas rien dire. Les erreurs de mesure présentées pour le PIB vont logiquement impacter le PIB par habitant et la croissance économique.

4.6 LES INDICATEURS COMPLÉMENTAIRES DU PIB

L'existence de limites du PIB nous invite à nous référer à d'autres indicateurs.

Les limites du PIB, et donc du PIB par habitant, ont conduit des économistes et des organismes nationaux et internationaux à proposer d'autres indicateurs. La liste des indicateurs complémentaires, voire substituables, serait longue. On peut donc se limiter à retenir l'IDH, l'indicateur de richesse développé par la Banque mondiale et les propositions de la CMPEPS.

4.6.1 *L'IDH, l'indicateur de développement humain*

L'IDH, l'indicateur de développement humain du PNUD (Programme des Nations Unies pour le Développement)

Comme le PIB ne permet pas de mesurer le niveau de développement, le bien-être, etc., il existe pour cela des indicateurs *ad hoc*. On retiendra essentiellement l'IDH et ses dérivés comme l'ISDH, les IPH... **L'IDH, l'indicateur de développement humain du PNUD, a été créé en s'inspirant des travaux de l'économiste Amartya Sen.** L'objectif de l'IDH est d'estimer le niveau de développement humain des pays, sans se limiter à la dimension économique mesurée par le PIB ou le PIB par habitant.

L'Indice de développement humain est un indice composite regroupant trois dimensions fondamentales du développement humain.

1. L'espérance de vie à la naissance exprime la capacité à vivre longtemps et en bonne santé. Elle permet d'estimer la satisfaction des besoins matériels essentiels tels que l'accès à une alimentation saine, à l'eau potable, à un logement décent, à une bonne hygiène et aux soins médicaux.
2. La durée moyenne de scolarisation et la durée attendue de scolarisation expriment la capacité à acquérir des connaissances. Le niveau de savoir ou niveau d'éducation est estimé par la durée moyenne de scolarisation pour les adultes de plus de 25 ans et la durée attendue de scolarisation pour les enfants d'âge scolaire. Il estime le degré de satisfaction des besoins immatériels tels que la capacité à participer aux prises de décision sur le lieu de travail ou dans la société.
3. Le revenu national brut par habitant exprime la capacité à avoir un niveau de vie décent.

=> La valeur de l'IDH est comprise entre 0 et 1.

Pour mesurer le développement humain de manière plus complète, le Rapport sur le développement humain présente quatre autres indices composites. L'IDH ajusté aux inégalités revoit l'IDH en fonction de l'étendue des inégalités. L'Indice de développement de genre compare les valeurs de l'IDH pour les femmes et pour les hommes. L'Indice d'inégalité de genre met en évidence l'autonomisation des femmes. L'Indice de pauvreté multidimensionnelle mesure les aspects de la pauvreté autres que le revenu.

Source : Bureau du Rapport sur le développement humain.

NIVEAUX DE L'IDH EN 2017

Monde	0,728
Pays en développement	0,681
Afrique subsaharienne	0,537
États arabes	0,699
Asie de l'Est et Pacifique	0,733
Asie du Sud	0,638
Amérique latine et Caraïbes	0,758
Europe et Asie centrale	0,771
Pays les moins avancés	0,524
Petits États insulaires en développement	0,676
OCDE	0,895
Développement humain très élevé	0,894
Développement humain élevé	0,757
Développement humain moyen	0,645
Développement humain faible	0,504

Source : PopulationData.net

<https://www.populationdata.net/palmares/idh/>

4.6.2 Mieux mesurer les performances économiques et du progrès social

La commission Stiglitz-Sen-Fitoussi (SSF), officiellement nommée « **Commission sur la mesure des performances économiques et du progrès social** » (CMPEPS), a remis un rapport au président de la République en 2008. Cette commission a été créée officiellement en 2007 par Nicolas

Sarkozy, sous l'impulsion de Jean-Paul Fitoussi (encore président de l'OFCE et secrétaire général de l'Association internationale de sciences économiques). Elle était présidée par le prix Nobel d'économie Joseph Stiglitz et le vice-président Amartya Sen, également prix Nobel d'économie. Le rapport de la CMPEPS a été suivi de nombreux travaux au niveau international.

Le rapport de la commission SSF de 2008 établit une distinction entre évaluation du bien-être présent et évaluation de sa soutenabilité, c'est-à-dire de sa capacité à se maintenir dans le temps. Le bien-être présent dépend à la fois des ressources économiques, comme les revenus, et des caractéristiques non économiques de la vie des gens : ce qu'ils font et ce qu'ils peuvent faire, leur appréciation de leur vie, leur environnement naturel. La soutenabilité de ces niveaux de bien-être dépend de la question de savoir si les stocks de capital qui importent pour notre vie (capital naturel, physique, humain, social) seront ou non transmis aux générations à venir.

Le bien-être est pluridimensionnel

Pour cerner la notion de bien-être, il est nécessaire de recourir à une définition pluridimensionnelle. À partir des travaux de recherche existants et de l'étude de nombreuses initiatives concrètes prises dans le monde, la commission SSF a répertorié les principales dimensions qu'il convient de prendre en considération. En principe au moins, ces dimensions devraient être appréhendées simultanément :

- Les conditions de vie matérielles (revenu, consommation et richesse) ;
- La santé ;
- L'éducation ;
- Les activités personnelles, dont le travail ;
- La participation à la vie politique et la gouvernance ;
- Les liens et rapports sociaux ;
- L'environnement (état présent et à venir) ;

– L'insécurité, tant économique que physique.

Toutes ces dimensions modèlent le bien-être de chacun ; pourtant, bon nombre d'entre elles sont ignorées par les outils traditionnels de mesure des revenus.

La **commission SSF** a développé une réflexion sur les moyens « d'échapper à une approche trop quantitative, trop comptable de la mesure de nos performances collectives » et d'élaborer de nouveaux indicateurs de richesse.

À propos du PIB, les recommandations suivantes organisent un programme de recherches qui conduisent notamment, au sein de l'Insee, à :

1. Se référer aux revenus et à la consommation plutôt qu'à la production.
2. Prendre en compte le patrimoine en même temps que les revenus et la consommation.
3. Adopter le point de vue des ménages.
4. Accorder davantage d'importance à la répartition des revenus, de la consommation et des richesses.
5. Élargir les indicateurs de revenus aux activités non marchandes.

Les indicateurs et les difficultés de mesure étant désormais connus, nous pouvons utiliser ces indicateurs pour mesurer la croissance (dynamique du PIB) et pour en observer les évolutions à long terme, puis à court terme, avant d'en développer les différentes explications (cf. chapitres 8, 9, 10 et 11 sur la croissance économique).

4.7 LA RÉPARTITION DES REVENUS

Comment les revenus sont-ils distribués dans l'économie ?

Le partage de la valeur ajoutée, les inégalités de revenus (salaires, profits, rentes), l'impôt sur le revenu des ménages (IRPP ou impôt sur le revenu des personnes physiques), l'impôt sur les bénéfices des entreprises, l'impôt sur le patrimoine, l'impôt sur la fortune... sont au cœur de l'actualité économique quotidienne. Il est impossible de prétendre

comprendre les rouages fondamentaux du fonctionnement de l'économie sans comprendre les mécanismes de la répartition des revenus et les problèmes qui se posent pour les individus, les ménages, les entreprises, les administrations publiques, pour l'économie et la société considérées dans leur ensemble.

Souvenons-nous que Jean-Baptiste Say dès 1803, puis David Ricardo en 1817, considéraient déjà que cette question de la répartition des revenus était centrale, dans la mesure où elle est inextricablement liée à celle de la production et de ses conditions. P. A. Samuelson (1948) considérait que la répartition était l'un des trois problèmes économiques fondamentaux (pour qui produire⁸ ?).

4.7.1 *Les revenus ont plusieurs origines*

Les revenus ont plusieurs origines dans une économie, car, *lato sensu*, la répartition des revenus se réalise en plusieurs étapes.

On distingue généralement la **répartition primaire** des revenus, qui résulte de la rémunération des facteurs de production ayant participé à la production par leur travail (revenus perçus par les travailleurs) ou leur apport en capital (revenus perçus par les détenteurs-propriétaires du capital productif), et la **répartition secondaire**, qui est le résultat du processus de **redistribution**. La répartition primaire des revenus sera donc suivie d'une redistribution, car la première distribution des parts du gâteau doit être corrigée, en raison de l'existence d'inégalités (multidimensionnelles et inhérentes au fonctionnement des économies capitalistes de marché), jugées trop grandes, d'écarts de revenus considérés comme inefficaces d'un point de vue économique ou injustes du point de vue de la justice sociale.

Au niveau macroéconomique, la répartition des revenus (avant et après redistribution) va automatiquement influencer les composantes de la demande globale de biens et services (pour qui produire ? Et que produire ?) et la capacité d'offre

globale de l'économie (comment produire ?). À court terme, cette répartition aura des effets sur le niveau du PIB effectif et la croissance effective et, à long terme, des conséquences sur le PIB potentiel et la croissance potentielle. Cette répartition primaire aura également des effets sur les déséquilibres macroéconomiques et financiers dans des économies de plus en plus ouvertes, où donc les interdépendances sont nombreuses : les valeurs absolues et relatives liées à ce partage ne seront pas sans effets sur les équilibres ou déséquilibres macroéconomiques et financiers qui en découlent.

Pour ces raisons, au moins, **la répartition primaire donnera lieu à un conflit récurrent portant sur le partage de la valeur ajoutée**. Quant à la redistribution, elle sera jugée à l'aune du degré de réduction des inégalités qui en découle, de son efficacité en termes d'incitations économiques et de l'amélioration éventuelle de la justice sociale qui en résulte.

4.7.2 *Une répartition indirecte des biens et services*

Pour commencer, retenons que **la répartition des revenus correspond simplement à une répartition indirecte des biens et services** qui ont été produits, grâce à la combinaison de facteurs de productions, pour satisfaire tous les besoins de tous les agents économiques. Les revenus ne sont qu'un moyen d'accéder aux biens et services pour la consommation, l'investissement, la constitution d'un patrimoine...

De surcroît, dans une économie de libre-échange et de libre choix, recevoir un revenu nous permet d'exercer cette capacité de choisir et d'échanger. En effet, pour se procurer la plupart des biens et services nécessaires à la satisfaction de leurs besoins, les agents économiques doivent obtenir des revenus. Les revenus reçus sont des sommes de monnaie versées en contrepartie de leur participation à l'activité

productive, de la possession d'un patrimoine ou encore via le processus de redistribution des revenus primaires.

4.7.3 *La répartition primaire des revenus*

En France, en 2016, pour un PIB de 2228,9 milliards d'euros courants, la rémunération globale des salariés s'élevait à 1159,7 milliards d'euros courants, soit 52 % du PIB. La rémunération globale des entreprises, l'excédent brut d'exploitation, était de 777,4 milliards, soit 35 % du PIB. (Source : Insee, TEF 2018 p. 109)

À la lecture de ces quelques données, on prend rapidement la mesure de l'importance de la répartition des revenus, qui sont la contrepartie des activités de production et qui permettent les dépenses immédiates et l'épargne (dépenses futures). Rappelez-vous l'adage cité précédemment selon lequel « *la production crée les revenus, les revenus permettent les dépenses et les dépenses conduisent à acquérir les produits* ».

Il indique que la dynamique économique va, en partie, dépendre de cette distribution et des équilibres macroéconomiques qui en découlent, dans une économie où les interdépendances entre opérations économiques, et donc entre agents économiques, sont irréductibles, incontournables.

4.7.4 *De la valeur ajoutée à la répartition primaire des revenus*

La répartition primaire des revenus est le résultat du partage de la valeur ajoutée entre les facteurs de production qui ont contribué à sa réalisation. Pour rappel, la valeur ajoutée est un supplément de valeur créée (une production nouvelle), au cours du processus de production. On la mesure en retranchant la valeur des consommations intermédiaires utilisées pour produire de la valeur : la valeur des biens et services vendus.

Valeur ajoutée = valeur de la production – valeur des consommations intermédiaires

La valeur ajoutée génère logiquement des revenus primaires pour tous les agents économiques ayant participé à la production.

Des revenus du travail sont versés au travailleur et des revenus du capital rémunèrent les détenteurs du capital productif et du patrimoine. Les deux principaux facteurs de production étant le **travail** et le **capital** productif, la répartition primaire des revenus concerne essentiellement les ménages et les unités de production au sein desquelles les entreprises occupent une place majeure.

Les **entreprises** qui réalisent un profit (généralement évalué par l'excédent brut d'exploitation), correspondant à la différence entre le chiffre d'affaires réalisé et les charges supportées, pourront conserver, *in fine*, une partie du bénéfice réalisé. Il est donc indispensable de **ne pas confondre profits et bénéfices**.

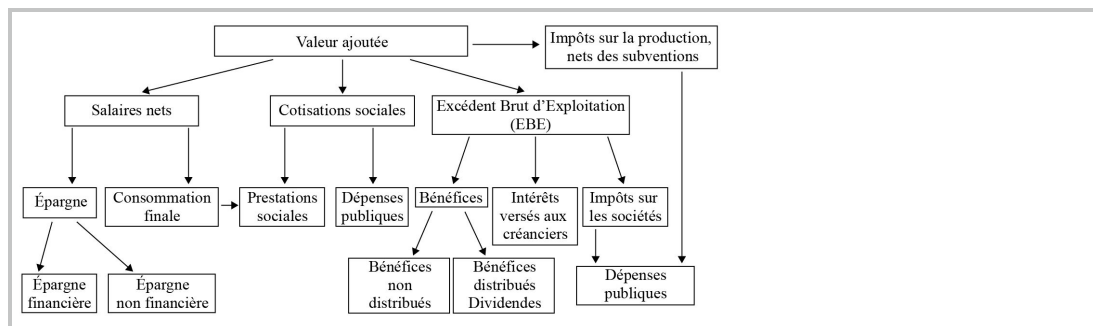
L'EBE (excédent brut d'exploitation) est la mesure du profit en comptabilité nationale. L'EBE est égal à la valeur ajoutée, diminuée de la rémunération des salariés, des autres impôts sur la production et augmentée des autres subventions sur la production. Le **profit** est un revenu primaire (partage de la valeur ajoutée) qui rémunère le facteur capital.

Quand on rapporte l'EBE à la valeur ajoutée, on obtient le taux de marge, qui mesure la part des profits dans la valeur ajoutée et qui sera déterminant pour le niveau d'investissement.

Le terme **bénéfice** est un terme de comptabilité privée qui désigne la différence entre les recettes de l'entreprise et l'ensemble de ses coûts. Pour l'analyse des revenus primaires issus de la valeur ajoutée, on préférera le terme profit.

En ce qui concerne les **ménages**, on peut distinguer **trois types de revenus primaires**. Les ménages perçoivent des

revenus d'activité comme les salaires et traitements (pour les fonctionnaires). Les **revenus du travail** rémunèrent le travail en tant que facteur de production. Des **revenus de la propriété** tels que les intérêts, dividendes, loyers, etc. peuvent venir compléter leurs revenus. Les **revenus mixtes** désignent ce que perçoivent les entrepreneurs individuels. Ils sont qualifiés de mixtes, car ils relèvent simultanément de la rémunération du travail et de la rémunération du capital productif (profit en tant qu'entrepreneur). Ces revenus sont ceux des entrepreneurs individuels et professions libérales. Les revenus du capital permettent de rémunérer les propriétaires du capital productif des entreprises et l'entreprise elle-même. Ainsi en est-il des profits non distribués par l'entreprise qu'elle pourra utiliser, selon ses arbitrages, pour des investissements productifs ou des placements.



GRAPHIQUE 4.2.

Le partage de la valeur ajoutée et l'utilisation des revenus

Il est utile de distinguer cette **répartition fonctionnelle** des revenus entre facteurs de production d'une **répartition personnelle** (entre individus ou ménages) et d'une **répartition sociale** (entre catégories sociales).

4.7.5 Salaire brut, cotisations sociales et salaire net

Les revenus primaires versés aux facteurs de production peuvent être décomposés de la manière suivante. Le **salaire** est la rémunération du travail, convenue entre un salarié et son employeur. On appelle **salaire mensuel de base** (SMB)

le salaire brut, primes et heures supplémentaires non comprises. Son montant correspond généralement à celui de la première ligne du bulletin de paye d'un salarié. Le **salaire brut** correspond à l'intégralité des sommes perçues par le salarié au titre de son contrat de travail avant déduction des cotisations sociales et avant versement des prestations sociales. Le **salaire net** (de prélèvements sociaux) est le salaire que perçoit effectivement le salarié. Les cotisations sociales ne sont pas comptées, y compris la CSG (contribution sociale généralisée) et la CRDS (contribution au remboursement de la dette sociale).

Les principales sources d'information sur les salaires sont les suivantes : les déclarations annuelles de données sociales (DADS), l'enquête sur le coût de la main-d'œuvre et la structure des salaires (ECMOSS) ainsi que l'enquête sur l'activité et les conditions d'emploi de la main-d'œuvre (Acemo). Cette dernière est réalisée par la direction de l'animation de la recherche, des études et des statistiques (Dares) du ministère en charge de l'emploi.

Le **coût salarial** (ou **salaire super brut**) est constitué par l'ensemble des dépenses qui incombent à l'entreprise pour l'emploi d'un salarié. Il se répartit en :

- un coût direct, composé principalement des salaires bruts auxquels s'ajoutent différents avantages salariaux (rémunérations, primes, congés payés, commissions et honoraires, avantages en nature, intéressement et participation, y compris cotisations sociales) ;
- un coût indirect formé essentiellement des cotisations patronales légales et conventionnelles et de diverses charges (formation professionnelle, frais de transport, œuvres sociales).

Pour le **coût salarial ou salaire super brut**, on a :

$\begin{aligned} \text{Rémunération} &= \text{Coût salarial} = \text{SB} + \text{CSS} = \text{SN} + \text{CSP} + \text{CSS} \\ &\text{avec CS} = \text{CSS} + \text{CSP} \end{aligned}$

En France, selon les données de l'Insee, pour l'ensemble des ménages en 2017, le revenu primaire brut (RPB) est de 1 603,6 milliards d'euros. La rémunération des salariés est de 1 216,2 milliards (75,8 % de ce RPB), le revenu du patrimoine

est de 266,1 milliards (16,6 %) et le revenu mixte brut, de 121,4 milliards (7,5 %). Les cotisations sociales (484,9 milliards) représentent 30,2 % du total du revenu primaire brut.

Logiquement, ce partage issu de la répartition primaire des revenus va déterminer les montants et les parts qui seront consacrés par les ménages à leur consommation, à leur épargne, à leurs investissements immobiliers et à leurs placements financiers, mais également l'investissement productif des entreprises (utilisation de l'épargne brute)... D'un point de vue macroéconomique, cela va affecter le montant du PIB et son taux de croissance, le niveau de l'emploi et du chômage... Il existe donc bien de nombreuses interactions entre les opérations de production, de répartition des revenus et de dépenses diverses (cf. logique du circuit économique de base).

4.7.6 *Le salaire minimum et la répartition primaire*

Il ne faut, pour autant, pas oublier que certains mécanismes institutionnels peuvent influencer fortement et durablement la répartition primaire des revenus.

De même, toutes choses égales par ailleurs, l'existence d'un salaire minimum, pour contrecarrer une certaine dualisation du marché du travail, peut influencer le niveau des salaires (via un effet grille hiérarchie des salaires), donc le partage de la valeur ajoutée et la répartition primaire des revenus. En effet, *ceteris paribus*, si la part de la valeur ajoutée qui sert à rémunérer le facteur travail augmente, alors celle qui permet de rémunérer le capital baisse, et vice versa. Cela n'est pas sans conséquence sur la demande globale, l'offre globale et donc sur la croissance économique, l'emploi...

En France, le salaire minimum brut mensuel (pour 151,67 heures de travail) était de 1 308,88 euros au 1^{er} mai 2008 (soit 8,63 euros de l'heure) ; il était de 1 445,38 euros au 1^{er} mai 2014 (soit 9,53 euros de l'heure).

Au 1^{er} janvier 2019, le SMIC brut est de 1 521,22 euros (soit 10,03 euros de l'heure).

Au 1^{er} janvier 2019, le SMIC net est de 1 171,34 euros (soit 7,72 euros de l'heure).

4.8 DES INÉGALITÉS DE REVENUS PRIMAIRES À LA REDISTRIBUTION

4.8.1 *Des inégalités dans la répartition primaire des revenus*

La répartition primaire est rarement économiquement efficace, ni même socialement acceptable, au sens où elle est en général plus ou moins inégalitaire (du point de vue de la répartition individuelle), voire considérée comme injuste. Elle ne concerne, en outre, que les agents économiques qui contribuent directement à la production ou disposent d'un patrimoine qui leur procure des revenus. Ceux qui sont inactifs, et plus largement les personnes qui ne sont pas (ou plus) en âge de travailler et produire et qui ne disposent d'aucun patrimoine de rapport, ne sont pas concernés. Ils ne peuvent pourtant pas être exclus, de fait, du partage du gâteau.

Dans ces conditions, les pouvoirs publics décident de procéder à une correction de cette répartition primaire par le biais du processus de redistribution. Ce **processus de redistribution** multidimensionnelle est simultanément un **processus de correction** du partage initial (primaire) du gâteau de biens et services produits et un **processus de distribution** d'une partie de ce gâteau vers ceux qui n'y ont pas accès directement, en raison de l'absence ou de l'insuffisance de revenus primaires. La redistribution conduit à modifier le partage des revenus primaires et à aboutir à une répartition secondaire des revenus.

En résumé, la redistribution se traduit par le passage de la répartition primaire à une répartition secondaire, le passage des revenus primaires au revenu disponible. En pratique, la **redistribution** correspond à un ensemble d'opérations et de mécanismes qui consistent à modifier la répartition primaire des revenus. Ces opérations recouvrent d'une part les

prélèvements obligatoires effectués par l'État, les collectivités locales et les organismes de protection sociale (ASSO en comptabilité nationale) sur les revenus primaires (impôts, cotisations sociales) et, d'autre part, le versement aux ménages de **revenus de transferts (prestations sociales)**, en espèces ou en nature.

4.8.2 De la redistribution au revenu disponible : les deux étapes de la redistribution

Le processus de redistribution se décompose schématiquement en **deux étapes**. Les **prélèvements obligatoires (PO)** permettent le versement de revenus de **transferts** (prestations sociales), le financement de services non marchands (redistribution *indirecte*). Les revenus des ménages sont modifiés de telle façon que leur revenu disponible peut devenir supérieur ou inférieur à leur revenu primaire, en fonction du niveau de départ. Si les ressources prélevées via les PO sont insuffisantes et que les dépenses à financer ne baissent pas, les ASSO doivent recourir à l'emprunt et leur dette croît.

Selon l'Insee, « *les revenus de transfert correspondent rigoureusement à des transferts, en espèces ou en nature, aux **ménages** qui sont destinés à alléger la charge financière que représente pour ceux-ci la protection contre un certain nombre de risques ou de besoins ; ils sont effectués par l'intermédiaire de régimes organisés de façon collective ou, en dehors de ces régimes, par des unités des administrations publiques ou des ISBLSM (institution sans but lucratif au service des ménages). Les prestations délivrées par les institutions de protection sociale prennent plusieurs formes :*

- *prestations sociales (au sens strict) : versement en espèces (pension de retraite, allocations, indemnités...) ou en nature (remboursement de soins ou de médicaments...) ;*
- *prestations de services sociaux : accès gratuit ou à tarif réduit à des services comme l'hospitalisation publique, l'hébergement en maison de retraite, réduction des tarifs des*

transports en commun pour les familles nombreuses ou les personnes âgées ».

Le **revenu disponible** d'un ménage, qui est le **résultat de cette redistribution**, est égal à la somme des revenus primaires (les revenus d'activité, les revenus du patrimoine) et des revenus de transfert dont on soustrait quatre impôts directs (l'impôt sur le revenu, la taxe d'habitation, les contributions sociales généralisées (CSG) et la contribution à la réduction de la dette sociale (CRDS)).

Pour faire simple, on obtient ainsi :
$$\boxed{RD = RP - PO + PS}$$

L'Insee calcule le RDB, le revenu disponible brut, « *revenu à la disposition des ménages pour consommer et épargner. Il comprend l'ensemble des revenus d'activité, des revenus du patrimoine et les prestations sociales, auxquels sont soustraits les impôts directs et les cotisations sociales. C'est une grandeur macroéconomique* ».

4.8.3 Du revenu disponible ajusté au revenu arbitrage

L'Insee calcule aussi le **revenu disponible ajusté (RDA)** des ménages, soit le revenu disponible augmenté des transferts sociaux en nature (qui sont la contrepartie des consommations individualisables incluses dans les dépenses des APU et des ISBLSM).

Enfin, les statisticiens calculent le **revenu arbitrage**. Il suffit de retrancher du RDA les **dépenses pré-engagées** du revenu disponible, c'est-à-dire l'ensemble des dépenses des ménages réalisées dans le cadre d'un contrat difficilement renégociable à court terme.

Ces dépenses pré-engagées pèsent lourd, surtout dans le budget des ménages les plus modestes. Il s'agit des dépenses liées au logement (y compris les loyers imputés), relatives à l'eau, au gaz, à l'électricité et aux autres combustibles utilisés dans les habitations, des abonnements aux services de télécommunications, des frais de cantine, de l'accès aux services de télévision (redevance télévisuelle,

abonnements à des chaînes payantes), des assurances (hors assurance-vie) et des services financiers (y compris, dans le cas de la comptabilité nationale, les services d'intermédiation financière indirectement mesurés). Ces dépenses réduisent la capacité réelle d'arbitrage des ménages dans l'utilisation de leurs revenus pour satisfaire des besoins divers.

4.8.4 Le résultat de la redistribution : le revenu disponible des ménages : du RP au RDB

Globalement, on observe que le RDB est plus faible que le RP. Au sens strict, la redistribution réduit, au niveau macroéconomique, le revenu disponible des ménages pour leurs choix (sous contraintes). Les prélèvements se mesurent au niveau agrégé, sans distinction des revenus par catégories de revenus. Les cotisations sociales pèsent plus lourd (33,7 % du RDB) que les impôts (15,6 % du RDB) dans le financement de la redistribution.

Comment expliquer ce choix ? On peut donc s'interroger sur le système des prélèvements-transferts et ses effets sur les inégalités et la justice sociale. Les prélèvements sont-ils trop élevés, les prestations sont-elles trop élevées ou trop faibles ? Cette redistribution est-elle économiquement efficace en matière d'allocation des ressources rares ? Quels sont ses effets sur la croissance effective et sur la croissance potentielle ? Les réponses à ces questions seront abordées dans les chapitres qui suivent. En particulier, les questions liées aux inégalités et à la justice sociale feront l'objet de développements spécifiques. Mais, à ce stade, il est important de ne pas oublier que la résolution des trois problèmes économiques fondamentaux implique des réponses interdépendantes. Les choix en matière de redistribution ne sont pas neutres sur l'allocation des ressources rares et donc sur le niveau de production et sur la croissance maximum possible.

4.9 L'ÉVOLUTION DU PARTAGE DE LA VALEUR AJOUTÉE ENTRE LES FACTEURS DE PRODUCTION⁹

4.9.1 *Le partage de la valeur ajoutée sur longue période*

En France, « plusieurs phases peuvent être distinguées depuis 1949. La part des salaires est relativement stable jusqu'au premier choc pétrolier, puis elle augmente jusqu'au début des années 1980. Ceci est suivi d'une baisse qui la ramène, à la fin des années 1980, légèrement en dessous de sa valeur d'avant le choc pétrolier. La hausse consécutive au premier choc pétrolier est en général attribuée au maintien de règles antérieures de progression des salaires dans un contexte de croissance ralentie. Les politiques de désindexation et le contrechoc pétrolier ont ensuite conduit au mouvement inverse. Depuis cette date, le partage apparaît plutôt stable... ».

(Source : « Partage de la valeur ajoutée, partage des profits et écarts de rémunérations en France, rapport au Président de la République », mission présidée par Jean-Philippe Cotis, directeur général de l'Insee)

Logiquement, comme les parts des salaires et des profits représentent l'essentiel de la VA, lorsque l'une baisse, l'autre augmente. Encore une fois, cela se traduira au niveau macroéconomique par des déséquilibres entre offre globale et demande globale et entre composantes de la demande globale. Ce partage n'est donc pas neutre pour le niveau du PIB et pour la croissance économique

4.9.2 *L'évolution de l'utilisation de l'EBE*

L'utilisation du profit brut, mesurée par l'EBE (cf. *Graphique 4.2. Le partage de la valeur ajoutée et l'utilisation des revenus*), et son évolution est tout aussi cruciale, car elles seront déterminantes de l'évolution de l'investissement et

donc de la capacité d'offre de l'économie, de son potentiel productif. On peut distinguer cinq grands types d'utilisation du profit brut : le paiement de l'impôt sur les sociétés (IS) ; le versement d'intérêts (net des intérêts reçus) ; la distribution de revenus aux propriétaires du capital (nette des mêmes revenus reçus) ; des opérations diverses regroupées en une seule catégorie (prestations sociales versées au titre des régimes d'employeurs, nettes des cotisations reçues, primes d'assurance-dommage, nettes des indemnités reçues, autres transferts courants divers) et l'épargne, qui correspond au revenu qui reste disponible pour l'autofinancement des investissements ou l'alimentation de la trésorerie.

3 Insee : Institut national de la statistique et des études économiques.

4 Les quasi-sociétés sont des unités qui n'ont pas de personnalité juridique propre, mais disposent d'une comptabilité séparée comme les succursales ou des unités comme la Régie des alcools (dont la personnalité juridique est celle de l'État).

5 La distinction entre société et entrepreneur est essentielle en comptabilité nationale. En effet, l'entrepreneur appartient aux SI des ménages alors que les sociétés appartiennent aux secteurs des SNF ou SF. L'entrepreneur est une personne physique alors que la société est une personne morale.

6 La formation de capital fixe étant mesurée « brute », l'usure ou obsolescence (usure morale) des actifs fixes utilisés n'est pas prise en compte, car elle est enregistrée dans la consommation de capital fixe (CCF). On obtient $FNCF = FBCF - CCF$ avec N pour Nette.

7 La DCF est entièrement payée par les ménages alors que la CFE incorpore des consommations qui ne sont pas payées par les ménages.

8 Il convient de ne pas confondre répartition (primaire) et redistribution. Le processus de formation du revenu disponible implique que la redistribution permet de passer du revenu primaire au revenu disponible.

9 Pour rappel, le produit intérieur brut (PIB) peut se calculer à partir de la somme des valeurs ajoutées brutes au prix de base des différents secteurs institutionnels (ou des différentes branches d'activité), en ajoutant les impôts sur les produits (12,9 % de la VA en 2016) et en retirant les subventions correspondantes (1,0 % de la VA en 2016).

Les impôts sur les produits sont la TVA, les impôts sur les importations et autres impôts sur les produits (TIPP, taxes sur les tabacs, sur les alcools, etc.).

L'ENTREPRISE DANS L'ÉCONOMIE DE MARCHÉ

« La tâche de chef est très spéciale : celui qui peut la résoudre n'a pas besoin d'être sous d'autres rapports ni intelligent, ni intéressant, ni cultivé, ni d'occuper en aucun sens une "situation élevée" ; il peut même sembler ridicule dans les positions sociales où son succès l'amène par la suite. Par son essence, mais aussi par son histoire (ce qui ne coïncide pas nécessairement), il est hors de son bureau typiquement un parvenu, il est sans tradition, aussi est-il souvent incertain, il s'adapte, anxieux, bref il est tout sauf un chef. Il est le révolutionnaire de l'économie – et le pionnier involontaire de la révolution sociale et politique. » - Joseph Schumpeter, 1911

« Le rôle de l'entrepreneur consiste à réformer ou à révolutionner la routine de production en exploitant une invention ou, plus généralement, une possibilité technique inédite. » - Joseph Schumpeter, 1942

« La volonté des firmes de s'agrandir a été un trait caractéristique du capitalisme depuis son début ; à la vérité, s'il n'en avait pas été ainsi, le capitalisme n'aurait jamais existé. » - Joan Robinson, 1976

SOMMAIRE

[5.1 Qu'est-ce qu'une entreprise ?](#)

[5.2 La diversité des entreprises](#)

[5.3 Le rôle déterminant de l'entrepreneur](#)

5.4 Quels sont les indicateurs de la performance d'une entreprise ?

5.5 Les transformations de la gouvernance d'entreprise depuis le XIX^e siècle

5.1 QU'EST-CE QU'UNE ENTREPRISE ?

5.1.1 *L'entreprise, un acteur central de l'économie*

Pour des millions de personnes dans le monde, la vie quotidienne est influencée par les choix économiques et sociaux effectués par les **entreprises** : celles-ci définissent l'organisation du travail, sont à l'origine des innovations, exploitent les matières premières, mobilisent des ressources humaines, financières et technologiques pour produire des biens et des services. L'entreprise privée constitue « l'institution cardinale du capitalisme », selon l'économiste François Perroux : en tant que principal centre de création de richesses, elle est un agent économique essentiel. Les médias relaient régulièrement les informations sur « la vie des entreprises » lorsqu'il s'agit d'évoquer leurs performances ou leurs déboires dans l'économie mondialisée d'aujourd'hui, évoquent le « moral des patrons » et des investisseurs, et de nombreuses revues y sont consacrées. L'entreprise est un objet d'études complexe, au sens où elle est à la fois un lieu de production, une organisation et une institution : elle forme un ensemble tourné vers l'objectif de la production, mais en tant que communauté humaine, elle est un système d'acteurs, créateurs de règles et de valeurs, un lieu où se nouent des contrats, des négociations, qui abrite des relations sociales de coopération, mais où peuvent aussi éclater des conflits.

Pour les économistes, l'entreprise est une **unité économique autonome** qui, grâce à la combinaison de divers facteurs de production, produit des biens et des services et distribue des revenus en contrepartie de l'utilisation de ces facteurs. La **production** est à la fois un processus et un résultat de ce processus, que l'on peut exprimer en quantités physiques (tonnes d'acier pour la sidérurgie, nombre de voitures d'un

même modèle pour un constructeur automobile, quantité d'un type de puces informatiques pour un constructeur informatique, nombre de clients servis dans un restaurant, etc.). La production consiste donc en une combinaison de facteurs (nature, ressources humaines, capital) dans un même lieu, (l'entreprise), en vue d'obtenir des biens ou des services destinés à la satisfaction des besoins des agents économiques. Pour produire, l'entreprise a besoin de temps, de savoir, de savoir-faire, de matières premières, d'équipements, d'un environnement qui facilite cette activité productive, etc. L'utilisation de ces facteurs donne lieu à une **rémunération** qui permet aux fournisseurs de disposer de revenus pour faire face à leurs dépenses présentes (ou futures), en épargnant une partie du revenu. Pour l'entreprise, ces différentes rémunérations sont des coûts de production.

Quelle que soit l'entreprise, la question du calcul économique se pose : sur les choix de la **combinaison productive**, d'une part, et sur la détermination du volume de production à réaliser, d'autre part.

Les économistes définissent l'entreprise avec les caractéristiques suivantes :

- L'entreprise produit des biens ou des services ;
- L'entreprise produit en vue de vendre et de réaliser un bénéfice ;
- Les recettes de l'entreprise proviennent de la vente du bien ou du service produit ;
- Pour produire, l'entreprise utilise dans des proportions variables des facteurs de production (travail et capital) qu'elle doit rémunérer ;
- La production et la vente de l'entreprise engendrent des coûts, dont notamment la rémunération des facteurs de production ;
- Le profit de l'entreprise représente la différence entre les recettes et les coûts.

En produisant, l'entreprise crée de la richesse. Cette richesse est ce qu'elle ajoute à ce qu'elle achète : c'est ce que l'on appelle la **valeur ajoutée**, qui est égale à la différence entre la valeur de la production (le chiffre d'affaires) et la valeur des biens et services utilisés pour cette production (les consommations intermédiaires).

Qui sont les principaux bénéficiaires de cette richesse créée ? Il s'agit des salariés (salaires et cotisations sociales) dans le cadre de la participation aux résultats ; de l'État, qui met à la disposition des entreprises des infrastructures et des services publics et à qui l'entreprise verse des impôts et des taxes ; des apporteurs de capitaux (banquiers et actionnaires) ; de l'entreprise elle-même pour investir (autofinancement).

5.1.2 Pourquoi existe-t-il des entreprises ?

L'entreprise se distingue des autres organisations productives par son caractère **marchand**. Pour les économistes, elle est, depuis les premières analyses de Ronald Coase en 1937 sur « la nature de la firme », un **mode de coordination des activités économiques** alternatif au marché. Coase montre notamment que l'entreprise et le marché sont deux formes alternatives de coordination économique : la coordination des comportements est assurée par le système des prix sur le marché, alors qu'au sein de l'entreprise, elle est assurée par un système hiérarchique (avec des rapports de pouvoir entre propriétaires, cadres dirigeants et autres salariés). L'entreprise peut par exemple choisir de recourir ponctuellement au marché pour bénéficier de la compétence d'un travailleur, ou décider d'embaucher durablement cette personne en tant que salarié. Dans le premier cas, sa rémunération dépendra du prix sur le marché du travail, tandis que dans le second, celle-ci sera fonction de règles formelles ou informelles fixées dans l'entreprise.

Coase explique que le recours au marché ou la production au sein de l'entreprise engendrent des coûts :

– Le recours au marché entraîne des **coûts de transaction** : essentiellement des coûts liés à la découverte des prix adéquats (recherche des prix les plus bas pour l'achat des consommations intermédiaires, par exemple), des coûts de négociation et de conclusion du contrat (un accord peut prendre du temps et nécessiter le recours à des compétences juridiques coûteuses). Ces coûts sont principalement le fruit de l'incertitude et des problèmes d'accès à l'information. De leur niveau dépend le choix de l'entrepreneur de produire au sein de l'entreprise (ce que l'on appelle « internaliser »), ou de déléguer à une autre entreprise (c'est-à-dire « externaliser »).

– La production en interne engendre des **coûts d'organisation** : au-delà du coût de fabrication (coût des facteurs de production et des consommations intermédiaires), un certain nombre de coûts sont induits par la définition précise des tâches, des règles et des rapports de subordination hiérarchique rationnellement établis, inhérents à la bureaucratie.

La **théorie des coûts de transaction** a été développée par O. Williamson (prix Nobel 2009) dans les années 1970, à partir des travaux précurseurs de R. Coase. Elle place l'étude des institutions au centre de l'analyse économique, en conservant l'hypothèse de rationalité des agents. L'existence des firmes en tant qu'institutions alternatives au marché se justifie lorsque ce dernier présente une efficacité moindre, du fait de la présence de coûts de transactions et de la possibilité de comportements opportunistes. Ces derniers consistent à ne pas respecter les engagements pris par l'agent dans le cadre d'une relation marchande. Par exemple, un salarié fournit peu d'efforts au travail, ou un assuré ne donne pas les informations correctes sur sa situation. Cette théorie explique que la comparaison des coûts d'organisation avec les coûts de transaction permet de déterminer le mode de coordination le plus adapté : si une transaction marchande supplémentaire présente un coût supérieur à celui engendré

par la production par l'entreprise, alors l'agent économique choisira d'internaliser cette activité (c'est-à-dire la réaliser dans le cadre hiérarchique de l'entreprise). Ainsi, quand l'entreprise se développe et accroît son volume de production, il arrive un moment où une taille trop importante devient un handicap : le processus productif devient très complexe à organiser, l'information circule mal et le rendement des facteurs de production baisse. Il existe ainsi une « **taille optimale** » pour une entreprise, qui correspond à la taille pour laquelle les rendements d'échelle sont maximaux. Pour R. Coase, la firme est donc un mode de coordination des relations économiques entre les agents qui repose sur la hiérarchie. Au sein des entreprises, ce sont les consignes données aux différents agents par les détenteurs du pouvoir hiérarchique qui assurent la coordination des activités. Ce mode de coordination est alternatif au marché, où la coordination est assurée par les prix relatifs, qui sont un vecteur d'information orientant les décisions des agents. Selon les cas, il est donc plus avantageux de passer par le marché ou par la firme.

Le développement des sociétés de services en ingénierie informatique (SSI) illustre parfaitement ce phénomène et la volonté de beaucoup d'entreprises de se recentrer sur leur cœur de métier (c'est-à-dire ce qu'elles font le mieux, comme fabriquer des véhicules pour Peugeot et Renault, par exemple) en sous-traitant l'activité informatique, comme la mise au point des logiciels. Dans un mouvement inverse, des entreprises choisissent d'en racheter d'autres pour internaliser leurs transactions (fusions-acquisitions), ou rattraper leur retard technologique et compléter leur gamme de produits (rachat par Apple des casques et écouteurs audio Beats Electronics en 2014 pour trois milliards de dollars).

5.1.3 *La combinaison productive et son efficacité*

La production de l'entreprise résulte de la combinaison des facteurs de production que sont le travail, le capital et les ressources naturelles.

Le capital renvoie à une triple dimension : physique, financière et humaine.

- Le **capital physique** est constitué de tous les biens courants qui concourent à la production d'autres biens. On distingue le capital circulant, c'est-à-dire les biens qui sont transformés ou consommés au cours du processus de production et le capital fixe, c'est-à-dire les biens qui restent une fois la production réalisée (les biens de production ou biens d'équipement). La notion de capital physique ne renvoie donc pas seulement aux machines utilisées durant plusieurs cycles de production, mais aussi à l'énergie utilisée et aux pièces détachées consommées.

- Le **capital financier** est constitué des capitaux dont peut disposer l'entreprise pour acquérir son capital physique : ce peut être celui des particuliers qui mettent directement leur argent à la disposition de l'entreprise en achetant des titres (comme les actions et les obligations) ou en souscrivant des parts du capital ; ce peut être le capital prêté par les banques (à court ou à long terme), ou les fonds de l'entreprise gardés en réserve, soit pour son fonctionnement (son fonds de roulement) soit pour son développement (par autofinancement, surtout constitué par les bénéfices non distribués).

- Le **capital humain** désigne l'expérience accumulée sous forme de savoirs, de savoir-faire, et plus généralement l'ensemble des compétences des travailleurs dans l'entreprise, qui dépendent de la formation initiale, mais aussi de la formation continue des individus.

Généralement, l'entrepreneur cherche à **maximiser le profit**, dans un état donné de la technique. Or, celui-ci change en permanence : pour pouvoir bénéficier de ces évolutions, l'entreprise investit : l'investissement consiste à augmenter ou renouveler le capital de l'entreprise. Il existe des

investissements matériels (comme l'achat de nouvelles machines ou le renouvellement du parc informatique) et immatériels, ces derniers comprenant les frais de recherche et développement (R&D), la formation du personnel, l'acquisition de logiciels, les dépenses de publicité et de marketing. Comme la formation du capital de l'entreprise se réalise par l'investissement, on évoque la formation brute de capital fixe (FBCF) dans la comptabilité nationale. L'investissement permet de produire plus et mieux, mais il est toujours risqué et coûteux : il dépend des anticipations de l'entreprise et de sa confiance dans l'avenir. Les coûts de production comprennent la rémunération des facteurs de production (capital fixe et travail), à laquelle s'ajoute le coût des matières premières et autres consommations intermédiaires.

L'entreprise effectue alors un choix de combinaison productive. Les **facteurs de production** peuvent être complémentaires ou substituables ou encore, cas le plus général, à la fois complémentaires et substituables.

- Lorsqu'ils sont parfaitement complémentaires, la question du choix de la combinaison productive ne se pose pas, puisqu'ils ne peuvent être associés que dans une certaine proportion ;
- Lorsqu'ils sont substituables, ils peuvent être associés dans des proportions variables et plusieurs combinaisons productives sont alors possibles pour réaliser un volume de production donné.

Pour choisir la meilleure combinaison productive, l'entreprise compare les **coûts** associés à chacune d'elles. On montre ainsi que le choix est déterminé par les coûts relatifs des facteurs de production, la hausse du prix d'un facteur incitant les entreprises à lui en substituer un autre (par exemple, si le coût du travail augmente, l'entreprise est incitée à substituer du capital au travail). La réflexion sur la combinaison productive conduit à s'interroger sur son efficacité, donc sur la

productivité de chacun des facteurs, dont l'expression peut être physique ou en valeur.

5.1.4 *La mesure des coûts et la recherche du profit de l'entreprise*

L'entreprise cherche à déterminer le **volume optimal de production**. Le choix du volume de production fait intervenir la comparaison entre les coûts et les recettes, au fur et à mesure qu'augmente le volume de la production. En effet, pour produire, il faut des matières premières, des équipements et de la main-d'œuvre. L'entreprise choisit le volume de production qui lui permet de maximiser son profit tout en étant en mesure d'écouler la production réalisée.

La motivation première de *l'entreprise capitaliste* demeure la recherche du profit, et cette capacité dépend de sa performance intrinsèque, de sa stratégie et de ses choix, mais aussi du marché sur lequel elle intervient. La quantité optimale à vendre pour le producteur rationnel est celle qui permet de réaliser le profit maximum. Le profit total de l'entreprise est égal à la différence entre la recette totale et le coût total ; il peut se calculer également en multipliant le profit moyen unitaire par le nombre d'unités produites. Le profit marginal est le profit supplémentaire engendré par la production d'une unité supplémentaire, il est égal à la différence entre la recette marginale (le prix de vente) et le coût marginal.

Pour maximiser son profit, l'entreprise doit produire une quantité telle que sa recette marginale soit égale à son coût marginal, autrement dit elle a intérêt à augmenter sa production tant que le coût marginal est inférieur à la recette marginale (profit marginal positif). Au-delà, le profit marginal devient négatif et l'entreprise, motivée par la maximisation du profit, n'a pas intérêt à accroître davantage sa production.

En termes comptables, le profit correspond aux recettes de l'entreprise qui ne sont pas utilisées pour acheter des

consommations intermédiaires ou pour rémunérer les salariés. Il sert en particulier à rémunérer le propriétaire de l'entreprise (c'est notamment ce que l'on appelle dans les grandes entreprises la création de valeur pour l'actionnaire).

Pour les économistes, le profit est le revenu résiduel de l'entreprise, celui qui reste quand elle a payé tous les facteurs de production.

5.2 LA DIVERSITÉ DES ENTREPRISES

5.2.1 *Comment classer les entreprises ?*

Au sens de l'Insee, l'entreprise est « *la plus petite combinaison d'unités légales et constitue une unité organisationnelle de production de biens et de services, jouissant d'une certaine autonomie de décision, notamment pour l'affectation de ses ressources courantes* ».

En effet, une entreprise peut être constituée d'unités (usine, siège social...) situées dans des endroits différents (on parle d'établissements) et, même si une administration est une unité de production, elle fournit des services non marchands et, de ce fait, n'est pas une entreprise. Le point commun entre une entreprise et une administration est, toutefois, qu'il s'agit **d'organisations productives**, c'est-à-dire un mode d'allocation des ressources et de coordination des activités qui a pour objet la production et qui est, à ce titre, fondé sur des règles. Ainsi, les organisations à activité marchande comme une société anonyme, une banque, une exploitation agricole, une coopérative de production constituent bien des entreprises. Toutefois, les unités à caractère non marchand, telles qu'un ministère par exemple, ne peuvent pas être des entreprises : dans ce cas, il s'agit d'une administration publique.

La nature de la production d'une entreprise constitue un critère de classement indispensable au sein du système productif : on distingue la situation des entreprises selon la filière, le secteur et la branche. Selon l'Insee, un *secteur*

d'activité regroupe des entreprises de fabrication, de commerce ou de service qui ont la même activité principale (au regard de la nomenclature d'activité économique considérée), alors qu'une *branche* (ou branche d'activité) regroupe des unités de production homogènes, c'est-à-dire qui fabriquent les mêmes types de produits (ou rendent des services) qui appartiennent à la même catégorie. La *filière* désigne, quant à elle, l'ensemble des activités complémentaires qui concourent, d'amont en aval, à la réalisation d'un produit fini. Selon l'Insee, on parle ainsi de « *filière électronique* » (du silicium à l'ordinateur en passant par les composants) ou de « *filière automobile* » (de l'acier au véhicule en passant par les équipements). La filière intègre d'ailleurs en général plusieurs branches. Mais on peut également classer les entreprises selon la taille et le statut juridique.

5.2.2 Les grandes catégories d'entreprises

Le décret du 18 décembre 2008 définit également **quatre nouvelles catégories de tailles d'entreprises**. Pour cela, il s'appuie sur des critères portant à la fois sur les effectifs, le chiffre d'affaires et le total de bilan. Recourir à ces **trois critères** permet de rendre compte de la réalité d'entreprises commerciales, financières ou de *holdings* qui jouent un rôle économique majeur par leur chiffre d'affaires ou par les actifs qu'elles détiennent, bien qu'elles n'emploient que peu de salariés.

Les quatre catégories d'entreprises selon des critères économiques

Quatre catégories de tailles sont donc distinguées : les **micro-entreprises**, les petites et moyennes entreprises (**PME**), les entreprises de taille intermédiaire (**ETI**) et les grandes entreprises (**GE**). Aux termes du décret, les PME incluent les micro-entreprises. Pour décrire plus facilement la population des entreprises sans recouvrement entre

catégories, on parlera donc des « *PME non micro-entreprises* ».

Les quatre catégories forment également une partition assez équilibrée de la valeur ajoutée. En revanche, les activités exercées, les moyens engagés, le rôle des groupes et les dimensions des marchés auxquels elles s'adressent sont très différents.

- La catégorie des **micro-entreprises** est constituée des entreprises qui occupent moins de dix personnes et qui ont un chiffre d'affaires annuel ou un total de bilan n'excédant pas deux millions d'euros.
- La catégorie des **petites et moyennes entreprises (PME)** est constituée des entreprises qui occupent moins de 250 personnes et qui ont un chiffre d'affaires annuel n'excédant pas 50 millions d'euros ou un total de bilan n'excédant pas 43 millions d'euros.
- La catégorie des **entreprises de taille intermédiaire (ETI)** est constituée des entreprises qui n'appartiennent pas à la catégorie des petites et moyennes entreprises, qui occupent moins de 5 000 personnes et qui ont un chiffre d'affaires annuel n'excédant pas 1 500 millions d'euros ou un total de bilan n'excédant pas 2 000 millions d'euros.
- La catégorie des **grandes entreprises (GE)** est constituée des entreprises qui ne sont pas classées dans les catégories précédentes. Plus encore que par le nombre de leurs salariés, les grandes entreprises (GE) supplantent celles des autres catégories par l'importance des moyens qu'elles engagent.

TABLEAU 5.1.
Les quatre catégories d'entreprises selon l'Insee

Secteur marchand non agricole	Effectif nombre de salariés	Taille du bilan	Chiffre d'affaires
Micro-entreprise	0 à 9	Inférieur à 2 millions d'euros	Inférieur à 2 millions d'euros

Petite et Moyenne Entreprise (PME)	Inférieur à 250 salariés (de 10 à 249)	Inférieur à 43 millions d'euros	Inférieur à 50 millions d'euros
Entreprise de Taille Intermédiaire (ETI)	Inférieur à 5 000 salariés (de 250 à 4 999)	Inférieur à 2 000 millions d'euros	Inférieur à 1 500 millions d'euros
Grande Entreprise (GE)	Supérieur à 5 000 salariés	Supérieur à 2 000 millions d'euros	Supérieur à 1 500 millions d'euros

L'intensité capitalistique, qui rapporte un indicateur du capital de l'entreprise (habituellement les immobilisations corporelles) à l'effectif salarié, diffère fortement selon les types d'entreprises retenues : elle est nettement plus élevée dans les entreprises de taille intermédiaire et plus encore dans les grandes entreprises, qui font appel à du personnel en moyenne plus qualifié. Les entreprises de taille intermédiaire (ETI) et les grandes entreprises (GE) ont ainsi, en moyenne, une productivité du travail plus forte (mesurée par la valeur ajoutée par salarié) et elles versent des salaires bruts plus élevés que la moyenne des entreprises.

Toutefois, les transformations du système productif ont fait du **groupe** une unité pertinente : il est constitué d'un ensemble d'entreprises contrôlées par une société appelée « *tête de groupe* » selon l'Insee. À l'heure actuelle, les investissements les plus importants sont souvent décidés au niveau du groupe plutôt que par l'entreprise indépendante.

5.2.3 *La classification des entreprises selon la taille*

La taille économique d'une entreprise peut être déterminée par le montant du chiffre d'affaires, la taille du bilan, et surtout le nombre de salariés.

« En 2015, l'ensemble des secteurs marchands non agricoles comptabilise 4,0 millions d'entreprises au sens unités légales et entreprises profilées pour un chiffre d'affaires hors taxes

global de 3 800 milliards d'euros et une valeur ajoutée de 1 100 milliards d'euros, soit 56 % de la valeur ajoutée de l'ensemble de l'économie française [...]. 1,1 million de micro-entrepreneurs et micro-entreprises au sens fiscal bénéficiant de formalités comptables et fiscales simplifiées — Entreprises ayant choisi les régimes fiscaux micro BIC (bénéfices industriels et commerciaux) ou micro BNC (bénéfices non commerciaux) — (pas de bilan, ni de compte de résultat, ni de liasse fiscale) ne sont pas retenus ici. Leur poids économique est très faible (1,1 % du chiffre d'affaires des entreprises de moins de 10 salariés) et les données statistiques disponibles sont moins détaillées. Les secteurs principalement marchands non agricoles et non financiers, hors micro-entrepreneurs et micro-entreprises au sens fiscal, rassemblent 2,35 millions d'entreprises, au sens unités légales et entreprises profilées (figure 1). Ces entreprises réalisent 3 655 milliards d'euros de chiffre d'affaires, 1 000 milliards d'euros de valeur ajoutée, soit 50 % de la valeur ajoutée de l'ensemble de l'économie, et emploient 11,86 millions de salariés en équivalent temps plein.

Alors que les 2,20 millions d'entreprises de moins de 10 salariés concentrent environ 20 % des salariés, du chiffre d'affaires et de la valeur ajoutée des secteurs principalement marchands, 4 200 entreprises de 250 salariés ou plus en concentrent entre 40 % et 45 %. »

(Source : Insee, Les entreprises en France en 2015, novembre 2017)

L'une des principales questions auxquelles l'entreprise est confrontée est celle de sa *taille optimale* : la grande taille permet à l'entreprise de réaliser des économies d'échelle, mais elle a l'inconvénient d'entraîner la hausse de certains coûts. Il est possible de classer les entreprises en les regroupant selon leur taille, même s'il est parfois difficile d'agréger des unités de production exerçant des activités différentes ou n'ayant pas les mêmes structures de production. Différents critères peuvent être utilisés pour répertorier les entreprises en fonction de leur dimension : le

volume des facteurs de production mis en œuvre (nombre de salariés, montant du capital investi), le chiffre d'affaires ou encore la rentabilité (valeur ajoutée, bénéfice). On peut classer les entreprises en fonction de nombreux facteurs : en fonction des *effectifs salariés* (mesurés généralement par les salariés permanents, même si le développement de formes particulières d'emploi et l'informatisation qui réduit l'emploi peuvent biaiser les résultats) ; en fonction des *indicateurs de création de richesse* (valeur ajoutée ou excédent brut d'exploitation) ; ou en fonction de divers *indicateurs financiers* (capitaux propres, ou capitalisation boursière pour les grandes entreprises cotées en Bourse). L'Insee retient l'effectif employé, tandis que le droit social français raisonne également en fonction de seuils fixés à partir d'un certain nombre de salariés en matière d'obligations des entreprises (par exemple, la loi oblige à créer un comité d'entreprise quand les effectifs dépassent les 50 salariés).

En 2015, selon l'Insee, l'ensemble des secteurs marchands non agricoles comptabilisait 3,82 millions d'entreprises dont, 3,67 millions de MIC, 139 941 PME, 5 753 ETI et 287 GE.

En France en 2015, 96 % des entreprises sont des MIC, 3,66 % des entreprises sont des PME, 0,15 % des entreprises sont des ETI et 0,0075 % des entreprises sont des GE.

TABEAU 5.2.

Nombre et pourcentages de salariés par catégories d'entreprises, en France, en 2015

Secteur marchand non agricole	Nombre d'entreprises 3 820 122 millions	En % du total	Effectifs salariés au 31/12/2015 14 497 000	% de salariés au 31/12/2015
MIC	3 674 141	96 %	2 745 000	18,9 %
PME	139 941	3,66 %	4 259 000	29,4 %
ETI	5753	0,15 %	3 657 000	25,2 %
GE	287	0,0075 %	4 325 000	29,8 %

Champ : unités légales et entreprises profilées (hors micro-entrepreneurs et

micro-entreprises au sens fiscal) des secteurs principalement marchands non agricoles et non financiers.

Source : Insee, *Les entreprises en France en 2015*, novembre 2017

5.3 LE RÔLE DÉTERMINANT DE L'ENTREPRENEUR

5.3.1 *L'entrepreneur comme force motrice du capitalisme*

L'entrepreneur est sans nul doute un personnage clé de l'économie. L'histoire du capitalisme depuis la révolution industrielle est d'abord émaillée d'**entrepreneurs de légende** : du barbier perruquier Richard Arkwright, qui se lance dans la filature dans les années 1760 à Steve Jobs d'Apple, Mark Zuckerberg de Facebook, ou Elon Musk de Tesla au début du ^{xxi}^e siècle, les *success stories* se succèdent depuis trois siècles et les grandes entreprises ont longtemps porté le nom de leur illustre créateur à l'instar de Ford, Peugeot, Tata, etc. L'analyse économique retient pour cerner ce personnage légendaire les diverses fonctions de l'entreprise moderne : tantôt le financier, tantôt l'innovateur, tantôt le gestionnaire avisé se succèdent pour incarner le personnage de l'entrepreneur. Preneur de risques ? Faiseur de projets ? Révolutionnaire de l'économie ? Découvreur d'opportunités ? Ses **qualités** dominantes font encore débat de nos jours. Mais si le rôle de l'entrepreneur peut apparaître déterminant dans la phase de démarrage de l'entreprise puis au cours de son développement, celle-ci reste une aventure collective, dont la réussite repose également sur un environnement et des institutions plus ou moins favorables à l'entrepreneuriat : moyens financiers (personnels, familiaux, bancaires), poids des institutions (famille, État et grandes entreprises), ressources scientifiques et techniques, état de la société (revenu, goûts des consommateurs, infrastructures).

Chez les auteurs précurseurs de l'économie, comme Adam Smith ou David Ricardo, les entrepreneurs sont issus des

deux **classes** qui détiennent le capital : les capitalistes ou les propriétaires fonciers. L'entrepreneur est donc d'abord une entité concrète pour expliquer un phénomène abstrait, comme la croissance ou la crise économique. Certains économistes du XIX^e siècle, comme Karl Marx, ont analysé le comportement de l'entrepreneur (ou du capitaliste) en le replaçant dans un cadre social plus général, qui est celui de la société capitaliste. Pour autant, Marx préfère le terme de « capitaliste » pour le désigner. Son analyse met en évidence la propriété privée des moyens de production et la nécessité de faire du profit. Pour Marx, l'entrepreneur n'existe que parce que la société capitaliste lui permet d'exister. Cette société est basée sur le profit maximum. L'entrepreneur ne peut exister sans s'enrichir, au risque sinon de la faillite. Pour cela, il doit sans cesse « révolutionner » les moyens de production et l'innovation devient un levier de la concurrence. Selon Marx, en favorisant la croissance des firmes, la concurrence favorise la socialisation de la production, qui nécessitera au final une transformation révolutionnaire des sociétés pour faire coïncider le régime de propriété avec la production collective. Ainsi se dissout l'entrepreneur, en même temps que la société capitaliste.

Paradoxalement, la démarche individualiste (de la « main invisible » à l'*homo oeconomicus*) de la théorie économique néoclassique ne laisse que peu de place à l'entrepreneur (on évoque simplement « l'équilibre du producteur » dans la théorie microéconomique). Tout au plus, l'entrepreneur est celui qui permet d'articuler le marché des facteurs de production à celui des produits dans l'équilibre général, mais ce rôle n'est pas vraiment théorisé. L'entrepreneur dans la théorie néoclassique n'est pas un individu exceptionnel, qui se distinguerait par des facultés spécifiques : son rôle d'intermédiaire entre des marchés, du travail, des biens et services est reconnu, mais il est difficile d'établir sa contribution à la création de richesse. Pour John Maynard Keynes, en revanche, les entrepreneurs favorisent par leurs comportements peureux et moutonniers (Keynes évoque des

« *esprits animaux* » ou les « *esprits déréglés des milieux d'affaires* ») la diffusion des crises économiques. L'entrepreneur renvoie donc plutôt, chez lui, à un concept collectif en liaison avec le principe de la demande anticipée. Marginal dans l'analyse néoclassique dominante, l'entrepreneur va pourtant apparaître chez quelques économistes comme un être exceptionnel, mû par des mobiles originaux et des capacités hors du commun, « faiseur de projets » (Jeremy Bentham) et « preneur de risques » (Richard Cantillon et Jean-Baptiste Say).

C'est à l'économiste autrichien Joseph Schumpeter qu'il revient d'avoir donné à l'entrepreneur cette place devenue mythique dans le développement du capitalisme, en l'identifiant à la **fonction d'innovation**. L'entrepreneur « schumpetérien » est en effet un esprit, une fonction plus qu'une personne que l'on pourrait identifier. Cette fonction est de rompre le « flux circulaire » (la routine) en détournant les facteurs de production des usages anciens vers de nouvelles combinaisons productives. Il rompt cette routine grâce au crédit des banques, ce qui lui permet de diriger le capital hors de son usage établi (reproduction du système à l'identique) pour l'affecter à de *nouvelles* combinaisons productives (technique, organisation, gestion). L'entrepreneur est donc celui dont l'existence fait que le changement vient de l'intérieur du processus économique, dans une dynamique endogène. Il ne s'adapte pas à un environnement (qui lui donnerait ses facteurs et absorberait ses produits), comme l'expose la théorie néoclassique, mais c'est lui qui, au contraire, par son intervention, façonne les structures économiques en modifiant les techniques de production et les préférences des consommateurs. La fonction de l'entrepreneur, selon Schumpeter, est donc uniquement d'innover : l'entrepreneur est ce personnage hors du commun, au cœur des déséquilibres économiques et du **processus de « destruction créatrice »** qui révolutionne perpétuellement les structures du capitalisme. Pour Schumpeter, l'entrepreneur est pourtant destiné à disparaître

et le capitalisme entrepreneurial avec lui. C'est la thèse qu'il développe dans *Capitalisme, Socialisme et démocratie* (1942). La routinisation de l'activité d'innovation (travail d'équipe et de laboratoire), la diminution des résistances à l'innovation dans des sociétés habituées à innover (et qui n'ont plus besoin d'être domptées par l'entrepreneur) font que le progrès technique tend à se bureaucratiser dans les grandes entreprises, et Schumpeter annonce alors le « *crépuscule de la fonction d'entrepreneur* ».

5.3.2 L'entrepreneur a besoin d'un environnement institutionnel favorable

L'entrepreneuriat, depuis ses débuts, est également une affaire de **réseau social**. Par exemple, à l'époque de la révolution industrielle, l'inventeur de la machine à vapeur, James Watt, n'a pu arriver à une application rentable qu'en s'associant d'abord avec un autre inventeur, John Roebuck, puis avec le producteur de boucles de ceinture Matthew Boulton et grâce à l'aide du sidérurgiste John Wilkinson dans les années 1770 (toutes ces personnes faisant partie d'une société savante, la *Lunar Society*). C'est donc bien sur un capital social, celui de ses réseaux, que s'appuie d'abord l'entrepreneur pour lancer son projet, et à l'origine de toute entreprise existe une coopération entre plusieurs personnes. L'entrepreneur dispose d'un **capital social** (acquis par ses ressources financières, informationnelles, relationnelles, tiré de son éducation), qu'il va valoriser en créant son entreprise à condition que la société, le marché, les institutions lui en donnent l'opportunité. Si le rôle de l'entrepreneur est de risquer son capital en créant une entreprise pour l'accroître, tous les individus ne disposent pas des mêmes ressources pour jouer ce rôle. La création d'une entreprise, d'une part, et la réussite, d'autre part, sont conditionnées par des ressources financières, des connaissances théoriques et empiriques et des relations personnelles et institutionnelles que le candidat entrepreneur sera en mesure de mobiliser pour son projet. Contrairement à ce que postule la théorie

économique, l'entrepreneuriat ne va pas de soi, sa réussite est conditionnée par un ensemble de facteurs extérieurs présents ou pas dans la société et la probabilité d'être entrepreneur n'est pas également ouverte à tous les individus.

Donc :

- À l'**échelle microéconomique**, l'entrepreneuriat dépend aussi de la classe sociale de l'entrepreneur, de son parcours éducatif et professionnel, de son vécu et de ses réseaux d'amitié et de connaissances. L'entrepreneur, de surcroît, est souvent appelé à s'appuyer sur son entourage pour compenser ses carences et compléter ses ressources.

- Au **niveau macroéconomique**, la qualité et le nombre d'entrepreneurs dans un milieu donné dépendent largement de la situation économique générale (boom, crise...) et sociale (stabilité, troubles...), des politiques économiques envisagées ou effectivement mises en place (encouragement à la création d'entreprise, forte protection des travailleurs, niveau et structure de la fiscalité...), de la configuration de l'offre (taille des marchés, localisation des entreprises...) aussi bien que de la demande (évolution du revenu des ménages, existence de besoins latents ou insatisfaits...), de l'existence et du niveau de diffusion des connaissances scientifiques et techniques, de la division de la production entre grandes et petites entreprises, etc. ; autant de facteurs qui vont permettre à l'entrepreneur de valoriser son potentiel de ressources. Les entrepreneurs insistent souvent eux-mêmes sur l'environnement réglementaire et fiscal.

L'entrepreneur n'est pas isolé dans l'économie et la société, et l'initiative individuelle doit pouvoir s'appuyer sur des institutions puissantes comme l'État, la finance et la famille. Tout un environnement institutionnel est ainsi bâti pour favoriser la création d'entreprise : aides de l'État et des associations, stratégies délibérées des grandes entreprises et des institutions bancaires et financières.

5.4 QUELS SONT LES INDICATEURS DE LA PERFORMANCE D'UNE ENTREPRISE ?

5.4.1 *Le bilan*

Le bilan représente la situation patrimoniale de l'entreprise à une date précise : le plus souvent, il est établi à la fin d'une année complète (en général le 31 décembre), soit ce que les comptables appellent un « exercice ». Il est à la fois un document comptable obligatoire et un outil d'analyse de la structure financière et de son équilibre financier. C'est un document qui décrit, à un moment donné, tout ce que l'entreprise possède (son « actif ») et tout ce qu'elle doit (son « passif »). De la même façon, on peut dire que le bilan décrit les ressources financières (au passif) et les emplois que l'entreprise fait de ses ressources (à l'actif). Les emplois sont nécessairement égaux aux ressources ; l'actif est toujours égal au passif.

Les petites entreprises au sens comptable (c'est-à-dire celles qui ne dépassent pas, pendant deux exercices successifs, deux ou trois seuils suivants : quatre millions de total de bilan, huit millions de chiffre d'affaires net, 50 salariés), peuvent désormais présenter un bilan simplifié.

On peut ainsi présenter la structure du bilan d'une entreprise :

L'**actif** représente l'ensemble des emplois. On distingue :

- L'actif immobilisé : les biens et les créances destinés à être utilisés ou à rester de façon durable dans l'entreprise : immobilisations incorporelles (licences, brevets, fonds de commerce), immobilisations corporelles (terrains, immeubles, installations techniques, machines, etc.), immobilisations financières (actions, parts sociales détenues, prêts consentis par l'entreprise à des tiers, etc.).
- L'actif circulant : biens et créances liés au cycle d'exploitation et qui n'ont pas vocation à être maintenus durablement dans l'entreprise : stocks (matières premières, produits en cours de production, produits finis

que l'entreprise a en stock), créances (factures clients en attente de règlement), instruments de trésorerie, disponibilités financières (avoirs détenus en caisse ou sur des comptes bancaires).

- Les charges à répartir sur plusieurs exercices : comptes d'attentes destinés à étaler les charges lors de l'inventaire. Conformément au plan comptable général (PCG), les charges à répartir ne concernent que les frais d'émission d'emprunts.

- Les écarts de conversion d'actif : les pertes de charges latentes.

Le **passif** représente les ressources de financement. On distingue :

- Les capitaux propres : les moyens de financement mis à la disposition de l'entreprise de façon permanente : capital, résultat de l'exercice, subventions d'investissement.

- Les provisions : destinées à couvrir des dettes probables liées à des risques et à des charges (litiges, grosses réparations, etc.).

- Les dettes : moyens de financement externes : emprunts, dettes fournisseurs, dettes sociales.

- Les écarts de conversion de passif : les gains de charges latentes.

Le résultat

Le bilan de fin d'exercice permet de calculer le résultat de l'exercice par différence entre :

Total actif (colonne net) – Total passif = résultat

Le résultat se place au passif dans la rubrique « capitaux propres » :

- En (+), il s'agit d'un bénéfice (actif > passif) ;

- En (–), il s'agit d'une perte (actif < passif).

Le résultat doit être identique à celui déterminé au compte de résultat.

Actif	Passif
– Actif immobilisé : immobilisations incorporelles (exemple : brevets) + immobilisations corporelles (exemple : biens d'équipement) + immobilisations financières (exemple : participation dans d'autres entreprises) – Actif circulant : stocks + créances détenues par des tiers (exemple : crédits) + disponibilités (argent disponible en banque ou en caisse) – Comptes de régularisation TOTAL	– Capitaux propres : capital (apport du ou des propriétaires) + réserves (bénéfices antérieurs non distribués aux propriétaires) – Provisions pour risques et charges – Dettes : à plus ou moins longue échéance (exemples : emprunts bancaires, obligations émises sur les marchés financiers) – Comptes de régularisation TOTAL

TABLEAU 5.4.

Signification simplifiée du bilan d'une entreprise

Actif	Passif
Ce que l'entreprise possède	Ce que vaut l'entreprise (capitaux propres...)
Ce qui est dû à l'entreprise (créances)	Ce que doit l'entreprise (dettes)

5.4.2 Le compte de résultat

Le compte de résultat décrit l'activité de l'entreprise au cours d'un exercice : c'est à la fois un document comptable obligatoire et un outil d'analyse de l'activité, de la performance et de la rentabilité de l'entreprise. Le compte de résultat fait partie des comptes annuels. Il regroupe les produits et les charges de l'exercice, indépendamment de leur date d'encaissement ou de paiement. Il est établi à partir des soldes des comptes de gestion : les comptes de charges

et les comptes de produits. Le compte de résultat détermine le résultat de l'exercice (bénéfice ou perte) par différence entre les produits et les charges.

Les **charges** correspondent à des coûts engagés par l'entreprise pour exercer son activité. On distingue parmi celles-ci :

- Les charges d'exploitation : les coûts occasionnés par l'activité normale de l'entreprise : achats et consommations, charges de personnel, impôts.
- Les charges financières : les coûts liés à la politique financière de l'entreprise : intérêts des emprunts, escomptes accordés, pertes dues aux variations des taux de change.
- Les charges exceptionnelles : les coûts qui ne sont pas liés à l'activité normale de l'entreprise : amendes, pénalités.
- La participation des salariés aux résultats de l'exercice : charge obligatoire pour toute entreprise ayant au moins 50 salariés.
- L'impôt sur les bénéfices : charge d'impôt supportée par les entités soumises à l'impôt sur les sociétés (société à responsabilité limitée, société anonyme).

Les **produits** sont les revenus générés par l'activité de l'entreprise.

On distingue parmi ceux-ci :

- Les produits d'exploitation : les ressources produites par l'activité normale de l'entreprise : ventes, subventions d'exploitation.
- Les produits financiers : revenus financiers procurés par des placements, escomptes obtenus et reprises sur provisions.
- Les produits exceptionnels : produits occasionnés, essentiellement, par la cession d'éléments de l'actif.

TABLEAU 5.5.

La structure du compte de résultat de l'entreprise

Charges (coûts)	Produits (revenus)

– Charges d'exploitation – Charges financières – Charges exceptionnelles – Impôt sur les bénéfices	– Produits d'exploitation – Produits financiers – Produits exceptionnels
Résultat de l'exercice (bénéfice)	Résultats de l'exercice (perte)
TOTAL GÉNÉRAL	TOTAL GÉNÉRAL

Le **résultat** (bénéfice ou perte) permet de savoir si l'entreprise s'est enrichie ou appauvrie :

- produits > charges = bénéfice ;
- produits < charges = perte.

Le résultat est placé du côté opposé à sa nature pour équilibrer le compte de résultat. Le résultat déterminé au compte de résultat doit être identique à celui du bilan.

5.4.3 Quelques ratios de performance de l'entreprise

La rentabilité désigne la capacité qu'a une activité à dégager un revenu supérieur à celui qui est engagé pour mener à bien cette activité.

- La **rentabilité commerciale** : ce ratio permet de calculer la capacité de l'entreprise à dégager un revenu satisfaisant sur les ventes (excédent brut d'exploitation/chiffre d'affaire).
- La **rentabilité financière** : les actionnaires ou associés comparent les fonds qu'ils ont investis dans l'entreprise et les bénéfices qu'ils en retirent ; ils calculent ainsi la rentabilité des capitaux propres (résultats de l'exercice/capitaux propres) et le bénéfice annuel associé à la détention d'une action, soit le bénéfice par action (résultat/nombre d'actions émises).
- L'actionnaire, propriétaire du capital, peut mesurer l'évolution du bénéfice par action et de la valeur de l'action en bourse grâce au « **Price earning ratio** » (**PER**), soit le rapport entre le cours de l'action et le bénéfice par action.

5.5 LES TRANSFORMATIONS DE LA GOUVERNANCE D'ENTREPRISE DEPUIS LE XIX^E SIÈCLE

5.5.1 *Qu'est-ce que la gouvernance de l'entreprise ?*

L'entreprise n'est pas une machine ni une simple chaîne de décision qui partirait d'un centre de décision parfaitement unifié, descendant par capillarité jusqu'aux moindres échelons de réalisation des projets. La fiction d'une entreprise où le pouvoir serait contrôlé de manière rationnelle par un entrepreneur, souvent supposé être le propriétaire, à l'image des entreprises familiales, ne résiste pas à l'observation des faits. On appelle **gouvernance de l'entreprise** l'ensemble des mécanismes qui ont pour effet de délimiter les pouvoirs des dirigeants et d'influencer leurs décisions. C'est le système de répartition des droits et de responsabilités des principaux acteurs de l'entreprise (les « parties prenantes ») : dirigeants, administrateurs, actionnaires et salariés. Le mode de gouvernance de l'entreprise permet de déterminer ses objectifs et les moyens de les atteindre, et d'évaluer ses performances.

Si l'usage de ce terme est relativement récent, il désigne un problème crucial pour les entreprises sur le long terme : comment associer des capitaux, et organiser la répartition des pouvoirs au sein de l'entreprise, de façon à en assurer le développement ? Le terme de « **gouvernance** » se différencie de celui de « gouvernement », qui suppose une organisation centralisée et hiérarchisée du pouvoir : abstraction faite des rapports hiérarchiques au sein de l'entreprise, celle-ci suppose une coopération de plusieurs parties prenantes (propriétaires du capital, managers, salariés, fournisseurs), déterminées à partager un objectif commun (*affectio societatis*), en dépit de divergences d'intérêts qui peuvent entraîner des conflits (comme sur le

partage des richesses créées). L'efficacité d'un système de gouvernance d'entreprise se mesure alors à sa capacité de réduire les conflits et les coûts de transaction internes préjudiciables à la réalisation de ses objectifs.

5.5.2 La naissance de l'entreprise capitaliste moderne (XIX^e siècle)

Avec la révolution industrielle au XIX^e siècle, on assiste à l'émergence de l'entreprise sous sa forme capitaliste moderne, caractérisée par la réunion, dans un même lieu, de salariés à qui le propriétaire confie un travail et apporte le capital nécessaire. Des transformations juridiques importantes ont été nécessaires à l'apparition de l'entreprise moderne, notamment pour faire émerger un marché libre du travail. En France, la Révolution française est une étape décisive, avec la loi Le Chapelier (1791), qui abolit les anciennes corporations, et en même temps interdit toute coalition ouvrière, et donc la pratique du syndicalisme. Cette loi est abolie par la loi Waldeck-Rousseau de 1884, qui autorise l'activité syndicale. Le développement des sociétés s'accélère cependant au cours du XIX^e siècle. Les sociétés en nom collectif (sociétés de personnes associées et responsables sur leur patrimoine), les sociétés en commandite et les sociétés anonymes apparaissent en France dans le code du commerce de 1807, mais les sociétés anonymes sont soumises à une autorisation jusqu'en 1867, et de ce fait le développement de sociétés par actions n'intervient réellement qu'à la fin du XIX^e siècle. Le phénomène est plus ou moins rapide selon les formes nationales d'industrialisation : par exemple, le développement de grandes sociétés est plus rapide aux États-Unis et en Allemagne dans la deuxième moitié du XIX^e siècle.

5.5.3 « L'ère des managers », du début du XX^e siècle aux années 1970

La direction d'entreprise a connu de profondes transformations au cours du xx^e siècle, en raison de la concentration progressive qui pousse à la constitution de vastes structures, dans lesquelles s'opère une distinction entre la propriété de l'entreprise et sa direction. Dans une société anonyme de grande taille, les actionnaires n'exercent pas un contrôle quotidien sur l'entreprise, ce contrôle est dévolu à un manager, c'est-à-dire à un cadre salarié, spécialisé dans les tâches de direction. J. Burnham (1905-1987) parle de « *révolution managériale* » selon le titre de son livre publié dès 1941, pour caractériser ce phénomène. De même, J. K. Galbraith (1908-2006) décrit dans *Le nouvel État industriel* (1967) la formation d'une technostructure, c'est-à-dire le groupe des cadres dirigeants qui détiennent le pouvoir économique réel, par opposition aux actionnaires, dont le pouvoir est plus formel et ne s'exerce pas quotidiennement. Selon W. Baumol (né en 1922), les managers ont ainsi une fonction d'utilité propre, celle de maximiser le chiffre d'affaire sous contrainte de profit minimum.

Dans son ouvrage *La main visible des managers* (1977), l'historien A. Chandler montre que le pouvoir croissant des managers conduit les firmes à orienter et structurer les marchés par leurs stratégies à long terme et leurs choix organisationnels. Le fonctionnement des marchés n'a donc rien de spontané, il est le produit des stratégies managériales rationnelles. Historiquement, le poids du contrôle managérial est bien visible, tout particulièrement aux États-Unis. Il en découle une tendance de plus en plus marquée à la séparation entre la logique managériale et la logique actionnariale, qui peuvent avoir des objectifs et des intérêts contradictoires.

5.5.4 Les transformations contemporaines de la gouvernance d'entreprise

La première moitié du xx^e siècle jusqu'aux années 1970 a été dominée par un capitalisme « managérial », caractérisé par

un certain équilibre entre les parties prenantes. Selon R. E. Freeman, les parties prenantes sont les « *individus ou groupes d'individus qui peuvent influencer ou être influencés par la réalisation des objectifs d'une organisation* » (Freeman, 1984). Par conséquent, les intérêts de ces agents peuvent être affectés par le comportement de l'entreprise, mais, en retour, leur réaction peut durablement affecter la firme elle-même. D'après la théorie des parties prenantes, l'entreprise doit donc se développer en partenariat avec ses différentes parties prenantes, plutôt que de privilégier la seule relation actionnariale.

L'époque contemporaine a été marquée par une remise en cause du modèle managérial de gestion et par le développement d'un nouveau modèle de gouvernance, davantage soucieux de la rentabilité pour les propriétaires du capital, de telle manière que l'on a pu évoquer parfois un « **retour de l'actionnaire** » :

- Dès la fin des années 1960, la **perte de compétitivité des grands oligopoles industriels** et financiers, la crise de la production de masse et de l'organisation du travail fordiste, le ralentissement des gains de productivité et la perte de rentabilité des titres financiers détenus par les actionnaires ont favorisé une évolution de la gouvernance de l'entreprise. Le pouvoir excessif des « managers » dans les grandes entreprises est fortement critiqué, d'autant que leur rentabilité subit un déclin continu. La « direction par objectifs » prônée par P. Drucker dès la fin des années 1960, est symptomatique de cette évolution : elle consiste à fixer des objectifs aux salariés avec des mécanismes incitatifs en fonction de leur réalisation, plutôt que leur donner des tâches à exécuter, de façon à les impliquer davantage dans leur travail.

- Un certain nombre de travaux universitaires et de rapports d'institutions internationales ont préconisé un rééquilibrage de la gouvernance des entreprises. L'économiste M. Friedman avait déjà hiérarchisé les priorités de l'entreprise capitaliste en affirmant de manière

polémique dans un article du *New York Times* que celle-ci, selon lui, n'a « *qu'une responsabilité sociale, et une seule, vis-à-vis de son actionnaire : utiliser ses ressources et s'engager dans des activités destinées à accroître ses profits* » (Friedman, 1970).

Dans un article de 1972, A. A. Alchian et H. Demsetz ont montré que pour être gérée efficacement, l'entreprise doit être contrôlée, en dernière instance, par ses actionnaires-propriétaires. Selon leur raisonnement, dans une production commune, on ne peut que difficilement isoler la contribution de chacun au résultat collectif : dans ces conditions, les propriétaires du capital, les actionnaires, doivent avoir le pouvoir ultime de nommer les dirigeants, de contrôler leur action et de les sanctionner le cas échéant, car ils ne peuvent percevoir des dividendes que si l'entreprise dégage un surplus une fois toutes les charges acquittées. Pour inciter les dirigeants à bien gérer la société anonyme et maximiser la rentabilité des capitaux, il faut donc qu'ils représentent les intérêts des actionnaires. Dès la fin des années 1970, l'*American Law Institute* énonce les grands principes de la *corporate governance*, censée encadrer le pouvoir des dirigeants (mécanisme de surveillance et d'information, rôle des conseils d'administration, mesures incitatives touchant à la rémunération des dirigeants, etc.), tandis que divers rapports ultérieurs diffuseront ces principes dans le monde de l'entreprise (rapport Cadbury en 1992, rapport Viénot en 1995, rapport de l'OCDE en 1999). Ces « codes de bonne conduite » se généraliseront dans la conduite des grands groupes sous la forme de normes et de « bonnes pratiques » à respecter sous la pression des actionnaires, et dans un contexte de surveillance des agences de notation qui évaluent la gouvernance des firmes.

– Le développement d'une économie de marchés financiers, dans les années 1990, a également transformé le rapport entre managers et actionnaires des grandes firmes. Les entreprises dépendent davantage pour leur

financement des marchés financiers, et le rapport de force se renverse au profit des actionnaires. De plus, la place croissante prise par les investisseurs institutionnels (fonds de pension, fonds mutuels, compagnies d'assurance) dans le capital des entreprises a conduit à une concentration de l'actionnariat, qui retrouve davantage de pouvoir de contrôle. Cela se traduit, dans les grandes firmes, par la remise en cause des managers aux résultats jugés insuffisants, et l'imposition de normes de profit élevées en faveur des actionnaires. C'est la base de la gouvernance « orientée actionnaire » (ou modèle « *shareholder* »), qui subordonne la direction d'entreprise à l'intérêt des actionnaires, en mettant en place des mécanismes de contrôle et d'incitations pour que les dirigeants fassent des choix en faveur des actionnaires. L'objectif de la firme est alors de favoriser la création de valeur, soit faire en sorte que le placement réalisé par l'actionnaire ait une rémunération supérieure à son coût d'opportunité (par exemple le rendement moyen des capitaux sur les marchés financiers). Différentes mesures sont utilisées, comme le développement des *stock-options* qui donnent aux managers une option d'achat à terme d'actions de leur société à un prix fixé à l'avance, ce qui les incite à tout mettre en œuvre pour que le cours de l'action augmente. La gouvernance « orientée actionnaire » a accompagné un changement profond de stratégie des firmes au cours des années 1990, avec le mouvement de « recentrage sur le métier » : les firmes créent davantage de valeur si elles se focalisent sur ce qui constitue leur avantage spécifique (le « cœur de métier »), et délèguent ou vendent le reste de leurs activités à d'autres firmes.

5.5.5 Le modèle de gouvernance orienté « actionnaires » en débat

Ce **modèle de la corporate governance** a fait l'objet de diverses critiques :

- D'une part, il a conduit au développement d'un capitalisme globalisé faiblement régulé soumis à de violentes crises financières et il a favorisé une gestion « court-termiste » à l'anglo-saxonne (un retour rapide sur investissement), au détriment des projets durables de développement, et au détriment des salariés.
- D'autre part, une autre critique porte sur le fait de privilégier uniquement l'actionnaire, au détriment des autres « parties prenantes » (ou « *stakeholders* »), c'est-à-dire des différents agents affectés directement ou indirectement par le comportement de la firme : les salariés (qui portent le risque alors que la part des dividendes versée aux actionnaires ne cesse de croître), les banques, mais aussi les clients et fournisseurs, les riverains, les représentants de la société civile (ONG, syndicats et associations de consommateurs) ou les administrations publiques. Avec notamment pour les États les enjeux autour de l'évasion fiscale et la question des paradis fiscaux posée aujourd'hui par le G20 et l'OCDE après la survenue de différents scandales financiers d'ampleur mondiale (Luxembourg Leaks, Panama Papers).
- Selon certains auteurs comme B. Segrestin et A. Hatchuel (*Refonder l'entreprise*, Seuil, 2012), on assiste aujourd'hui à une véritable « crise de l'entreprise », due aux limites du modèle de gouvernance orienté actionnaires, et qui met en cause la pérennité même des entreprises à long terme, en raison des risques trop élevés pris par leurs dirigeants. Il faudrait donc selon eux « refonder l'entreprise » en s'appuyant sur l'ensemble des parties prenantes, et notamment les salariés dans le cadre d'une nouvelle démocratie d'entreprise, où les dirigeants seraient habilités par les salariés. Un certain nombre d'analystes plaident pour un renforcement de la représentation des salariés au conseil d'administration des grandes entreprises et un accroissement des pratiques de « profit sharing » avec les salariés (partage des profits de l'entreprise).

5.5.6 *La responsabilité sociale de l'entreprise (RSE)*

L'entreprise n'est pas un agent économique déconnecté de la société. L'entreprise et la société sont en interaction : l'action de l'entreprise a des implications sur la société, et celle-ci a des implications sur la vie de l'entreprise. Si au XIX^e siècle, les patrons, au nom d'un certain paternalisme, pouvaient avoir des préoccupations sociales en faveur de leurs travailleurs, au nom d'une certaine philosophie humaniste, la responsabilité sociale des entreprises (traduit de l'anglais *social responsibility*) ne concerne pas aujourd'hui le seul comportement de l'entreprise avec son personnel (le « *social* »). Selon la Commission européenne, la responsabilité sociale de l'entreprise (RSE) est « *la responsabilité des entreprises vis-à-vis des effets qu'elles exercent sur la société* ». La responsabilité sociale de l'entreprise (*corporate social responsibility* aux États-Unis) suppose le principe selon lequel l'entreprise a des obligations non seulement envers ses actionnaires (les propriétaires du capital), mais également envers de ses salariés et de la société civile. On évoque parfois une entreprise « *citoyenne* », qui affirme respecter un certain nombre de valeurs qu'elle entend partager avec la société au sein de laquelle son action se développe. Depuis une vingtaine d'années, les préoccupations de la société à l'égard des enjeux sociaux, sociétaux et environnementaux n'ont cessé de s'amplifier et les entreprises ont été soumises à des pressions grandissantes de l'opinion publique pour mener des actions dans ces domaines au nom du développement durable. Les crises écologiques (affaires Erika, Exxon, Valdès, BP) et les divers scandales financiers (Enron, Parmalat), ainsi que les diverses menaces environnementales (déforestation, pluies acides, effet de serre), conjugués aux risques et aux conditions de travail déplorables constatés dans certains ateliers de confection travaillant pour des marques internationales de vêtements

(effondrement de l'immeuble Rana Plaza au Bangladesh en avril 2013, causant la mort de 1138 ouvrières et ouvriers et plus de 2000 blessés) ont amené les grandes entreprises à reconsidérer leurs manières de conduire leurs affaires.

Un certain nombre d'outils de mesure ont été développés, comme la « *notation sociale* » qui consiste en une évaluation par des cabinets d'audit social indépendants des comportements de l'entreprise dans divers domaines (environnement, égalité hommes/femmes, gestion du personnel, etc.), ou la promotion d'une normalisation des critères d'évaluation utilisés dans le domaine sociétal (norme « SA 8000 » dans le domaine du droit du travail sur la base des conventions de l'Organisation Internationale du Travail (OIT), norme « ISO 26 000 » relative au comportement « socialement » responsable des entreprises dans le cadre du développement durable, et qui permet aux entreprises certifiées de le faire savoir sur leurs produits. Dans un rapport de France Stratégie publié en janvier 2016, intitulé *Responsabilité sociale des entreprises et compétitivité*, S. Benhamou et M.-A. Diaye évoquent la crise financière de 2008 comme un accélérateur des initiatives en faveur de la RSE, afin de repenser le modèle de l'entreprise du XXI^e siècle et susciter de nouvelles dynamiques de croissance durable et inclusive, même s'ils déplorent que ce thème soit encore méconnu d'une grande partie des entreprises. Les auteurs de l'étude montrent notamment que la RSE n'est pas, bien au contraire, un obstacle à la performance économique.

Pour les économistes, la RSE n'induit d'ailleurs pas un mode de gouvernance particulier de l'entreprise, en particulier parce qu'elle peut fort bien s'exercer dans le cadre de la gouvernance « actionnariale », ou bien être ignorée dans le cadre d'autres types de gouvernance pourtant respectueux d'un équilibre avec les autres parties prenantes, elle implique un débat sociétal et politique sur l'organisation générale du système de gouvernance d'entreprise, et plus largement, sur la place de l'entreprise dans la société. L'entreprise est alors confrontée au défi de maximiser son profit sous contraintes

sociale et environnementale, car ne pas le faire serait prendre le risque de détériorer son image de marque et donc ses profits futurs. Pour expliquer l'essor des stratégies de RSE, la littérature économique, appuyée notamment sur les travaux de J. Tirole (né en 1953, prix Nobel en 2014) et R. Benabou, s'est développée principalement autour de la question des imperfections de marché et des incitations. Les travaux sur ces enjeux identifient trois principaux arguments incitant les entreprises à offrir des biens publics ou à corriger leurs impacts négatifs de façon volontaire : 1) éviter une future réglementation contraignante, 2) répondre aux pressions de la société civile et 3) s'acquitter d'un devoir moral ou altruiste. Deux autres motifs économiques sont généralement évoqués : la RSE peut être un outil de stratégie concurrentielle pour augmenter ses parts de marché en attirant de nouveaux consommateurs et pour se différencier de ses concurrents ; la RSE peut en fait répondre à une demande émanant de trois catégories de parties prenantes internes : les actionnaires (en particulier les investisseurs socialement responsables), les employés et les administrateurs. Par exemple, la dégradation de l'environnement par une firme peut conduire les acteurs de la société civile à mener des campagnes contre elle (dénigrement sur les réseaux sociaux, boycott), ce qui risque de menacer son activité même.

Cependant, cette politique de la RSE est vue par ses défenseurs comme un moyen de réconcilier l'entreprise et la société, alors qu'elle est perçue par ses détracteurs comme un simple outil de communication au service des objectifs de l'entreprise, et notamment la motivation du personnel pour consolider la cohésion interne. En mars 2018 en France, Jean-Dominique Senard, président du groupe Michelin, et Nicole Notat, ancienne secrétaire générale de la CFDT, ont remis un rapport d'étude après une mission intitulée « Entreprise et intérêt général » à Bruno Le Maire (ministre de l'Économie et des Finances), Nicolas Hulot (ministre d'État, ministre de la Transition écologique et solidaire), Muriel

Pénicaud (ministre du Travail) et Nicole Belloubet (Garde des Sceaux, ministre de la Justice). Les propositions ont alimenté la loi du Plan d'action pour la croissance et la transformation des entreprises (PACTE) votée en 2018. Le texte porte sur la gouvernance de l'entreprise et la nécessité de changer son « **objet social** », afin qu'il soit moins tourné vers la recherche du profit, mais il a aussi pour ambition d'améliorer le dialogue social et la prise en compte des parties prenantes, en augmentant le nombre d'administrateurs salariés notamment. Il a également modifié certains articles importants du Code civil pour préciser la « **raison d'être** » **des entreprises** : l'article 1833 du Code civil a été amendé pour consacrer la notion jurisprudentielle d'« intérêt social » et pour affirmer la nécessité pour les sociétés de prendre en considération les enjeux sociaux et environnementaux inhérents à leur activité. Et l'article 1835 du Code civil a été aussi été réécrit pour reconnaître la possibilité aux sociétés qui le souhaitent de se doter d'une « raison d'être » dans leurs statuts, au titre des engagements pris en matière de responsabilité sociale et environnementale. L'objectif est ici d'inscrire la « raison d'être » de l'entreprise dans un projet davantage soucieux du long terme et de l'intérêt collectif.

LA MONNAIE ET LE FINANCEMENT DE L'ÉCONOMIE

« Il n'est pas dans l'économie quelque chose de plus insignifiant que la monnaie, si on la considère autrement que comme un mécanisme pour faire vite et commodément ce qu'on ferait moins vite et moins commodément si elle n'existait pas. » - John Stuart Mill, 1848

« Si les changements dans le taux des prêts [bancaires] interviennent simultanément et uniformément avec les changements correspondants dans le taux réel de l'intérêt, alors – en dehors de l'influence directe de la production d'or – aucun changement dans le niveau des prix des biens ne peut arriver, et encore moins un changement cumulatif à la hausse. » - Knut Wicksell, 1906

« Le problème n'est jamais d'expliquer une "valeur absolue" de la monnaie mais seulement de savoir comment et quand la monnaie influe sur les valeurs relatives des biens et à quelles conditions elle n'affecte pas ces valeurs relatives ou, pour reprendre une expression heureuse de Wicksell, quand la monnaie est neutre par rapport aux biens. Ce n'est pas une monnaie stable en valeur mais une monnaie neutre qui doit être le point de départ de l'analyse théorique des influences de la monnaie sur la production. » - Friedrich Hayek, 1931

« L'importance essentielle de la monnaie vient de ce qu'elle est un lien entre le présent et le futur. » - John Maynard Keynes, 1936

[6.1 Pourquoi utiliser la monnaie ?](#)

[6.2 Monnaie et activité économique](#)

[6.3 À quoi servent les banques et les marchés financiers ?](#)

[6.4 Les banques : des acteurs majeurs du financement de l'économie](#)

[6.5 Le pouvoir de création monétaire des banques](#)

[6.6 La régulation de la création monétaire](#)

6.1 POURQUOI UTILISER LA MONNAIE ?

6.1.1 *La monnaie : une institution au cœur de l'économie de marché*

La monnaie est l'un des instruments les plus utilisés dans notre vie quotidienne dans le cadre de l'économie de marché. Sur le plan étymologique, la monnaie vient du latin *moneta* que l'on peut traduire par « celle qui avertit » ou encore « celle qui donne son avis ». Ce surnom était aussi donné à la déesse Junon à qui l'on attribuait le pouvoir d'annoncer les événements à venir (*Judo moneta*), et, dans le temple qui lui était dédié, situé sur le Capitole, à Rome, étaient frappées des pièces de monnaie que l'on a fini par désigner par le terme de *moneta*. Le terme de monnaie nous renvoie donc historiquement à son utilisation usuelle : des pièces d'un faible montant. Cependant, pour l'économiste, cette notion s'applique aux instruments de paiement au-delà des seuls pièces et billets qui ne constituent qu'une faible part de la monnaie en circulation. **La monnaie est donc l'ensemble des moyens de paiement dont disposent les agents économiques pour régler leurs transactions.** Dans certaines sociétés dites « primitives », les échanges, très limités, se réalisaient sous forme de troc. Les produits s'échangeaient contre d'autres produits. Lorsque les échanges se sont développés, le troc est devenu impossible et un bien quelconque (le coquillage, par exemple) a servi d'intermédiaire. Les produits s'échangeaient alors contre ce bien particulier, appelé monnaie d'échange, et celui-ci permettait ensuite l'achat d'autres biens. La monnaie est donc un bien particulier, reconnu et accepté par tous, destiné à faciliter les échanges.

6.1.2 *Les trois grandes fonctions de la monnaie :*

– Une fonction d'**unité de compte**. Elle permet d'exprimer la valeur de tous les biens et services dans une unité commune (par exemple l'euro) ; le préalable à l'existence de la monnaie est l'échange et réside dans la volonté de se procurer ce dont on est démuné en cédant une partie de ce dont on est pourvu. Le troc pourrait très bien aboutir à ce résultat, mais au prix d'une grande quantité d'évaluations : le recours à la monnaie est dès lors nécessaire lorsque la quantité d'objets à évaluer progresse. La monnaie **simplifie alors les évaluations**, car elle constitue un numéraire qui permet d'exprimer unanimement les prix. Déjà, le philosophe Aristote insistait sur l'introduction de la monnaie pour la commensurabilité des objets et pour le rôle d'étalon de mesure : la monnaie est bien cette référence à laquelle il est possible de comparer tous les autres biens. La monnaie sous la forme d'unité de compte est apparue au moment où naissait la capacité d'enregistrer par écrit les comptes (avec le développement de l'écriture). Elle aide à l'évaluation et permet d'afficher une moindre quantité de prix que le troc : une inflation basse (et une variabilité faible des prix) est souvent défendue dans le débat monétaire pour sa capacité à assurer une mesure stable de la valeur des choses au fil du temps. Les modifications de la quantité de monnaie en circulation peuvent alors perturber ce rôle d'étalon de mesure, puisque, d'une année sur l'autre, une même quantité de monnaie peut ne pas permettre l'achat d'une même quantité de biens, sans pour autant que cela conduise à abandonner la référence monétaire.

– Une fonction d'**intermédiaire des échanges**. Dans une économie de troc, si un individu A veut vendre des pièces d'étoffe et acheter une table, il doit trouver un vendeur de tables qui désire des pièces d'étoffe pour réaliser son échange. On dit alors qu'il faut qu'il y ait *double coïncidence des besoins*. Par l'intermédiaire de la monnaie, en revanche, l'individu A peut vendre ses pièces d'étoffe à n'importe qui et, en échange de la monnaie

reçue, peut acheter une table à n'importe quel vendeur ; pour jouer ce rôle d'instrument de transaction, la monnaie doit prendre une forme qui permette de la multiplier en fonction des besoins d'échanges, de la diviser en unités élémentaires et de la transporter facilement, à condition d'être acceptée par tous les coéchangistes. La monnaie est donc indispensable pour assurer l'existence d'une société au sein de laquelle les individus exercent des activités très diversifiées : on dit que la monnaie est indispensable à la *division du travail*.

– Une fonction de **réserve de valeur**. L'individu A, qui a vendu ses pièces d'étoffe, peut conserver un temps la monnaie qu'il détient pour acheter dans un temps futur différents biens. Dans une économie de troc, les actes de vente et d'achat sont simultanés. Dans une économie monétaire, il est possible de conserver le revenu issu d'une vente et d'exprimer une demande de bien ou service ultérieurement. L'idée de réserve de valeur intervient dès que les opérations d'échange sont disjointes dans le temps : la monnaie que j'obtiens en échange d'un bien ou d'un service vendu doit me permettre d'effectuer, plus tard, l'achat d'un bien ou d'un service d'une valeur équivalente. Sa qualité d'actif liquide lui permet d'être convertible instantanément en n'importe quel bien ou service : la monnaie achète tous les biens. L'inflation peut alors compromettre cette fonction de la monnaie : j'obtiendrai à l'avenir, avec la même quantité de monnaie, une quantité moins grande de biens et services. La déflation est également dangereuse : si les prix baissent, les consommateurs reportent leurs achats et attendent de nouvelles baisses de prix.

6.1.3 Les formes contemporaines de la monnaie

Dans les économies contemporaines, on distingue habituellement la monnaie divisionnaire, la monnaie fiduciaire

et la monnaie scripturale.

- La **monnaie divisionnaire** correspond aux pièces, lesquelles ne représentent désormais qu'à peine plus qu'une très faible part des paiements en France (et dans de nombreux pays) aujourd'hui. La monnaie divisionnaire est, dans une certaine mesure, une monnaie métallique dont la valeur intrinsèque est différente de sa valeur faciale. Il s'agit donc d'une forme de monnaie fiduciaire.

- La **monnaie fiduciaire** est une monnaie fondamentalement basée sur la confiance qu'on a de son usage et de sa valeur d'échange. La valeur de la monnaie fiduciaire est fondée sur la confiance dans la mesure où sa valeur intrinsèque peut être très faible par rapport à sa valeur faciale, qui détermine sa valeur d'échange (« fiduciaire » vient du latin *fiducia* qui signifie confiance). Il existe un décalage plus ou moins important (selon la valeur des billets) entre la valeur intrinsèque et la valeur faciale. La confiance va expliquer que nous soyons capables d'utiliser des monnaies à faible valeur intrinsèque. Ce décalage existe donc aussi pour les pièces, mais dans une bien moindre mesure, étant donné la faible valeur faciale de ces dernières. C'est cette confiance qui nous permet d'accepter d'utiliser une monnaie dont la valeur intrinsèque est relativement si faible. Il y a, d'une part, la confiance fondée sur sa pérennité et le fait que tout le monde va l'utiliser durablement, et, d'autre part, la confiance dans la stabilité de son pouvoir d'achat et donc dans le maintien de la stabilité des prix. L'État est donc au cœur des processus monétaires. Une monnaie est forcément attachée à un État, et l'éclatement des États se traduit souvent par la disparition de leur monnaie. C'est la confiance dans la pérennité de la communauté à laquelle ils appartiennent qui conduit les individus à accepter ces moyens de paiement en échange de leurs biens ou en rémunération de leur travail. La monnaie fiduciaire (les billets) représente moins de 10 % des paiements en France aujourd'hui.

– La **monnaie scripturale** est la principale forme de monnaie dans les économies développées et émergentes de nos jours. La monnaie scripturale (« scriptural » vient du latin *scriptus* qui signifie « écriture ») est le résultat d'un jeu d'écritures sur les comptes des clients et de la banque. La monnaie scripturale est donc purement immatérielle. Elle ne peut être détenue manuellement et transportée avec nos pièces et nos billets dans notre porte-monnaie. La monnaie scripturale représente désormais environ 90 % de nos paiements en France. Les paiements sont effectués par les débits et crédits sur les comptes bancaires des agents économiques concernés. Ainsi s'opère l'essentiel de la circulation de la monnaie dans les économies contemporaines.

La distinction usuelle entre monnaie fiduciaire et monnaie scripturale ne doit pas nous tromper, car les deux formes de monnaie reposent sur la confiance.

La monnaie scripturale ne peut être mobilisée ou utilisée qu'en utilisant divers instruments tels que le chèque, la carte de paiement bancaire, le virement, l'autorisation de prélèvement... Ces supports traduisent l'ordre donné à la banque de réaliser le paiement engagé par son client lors d'un échange marchand. C'est ainsi que va circuler la monnaie scripturale, par virement d'une somme du compte courant d'un agent économique vers celui d'un autre, dans la même banque ou dans une autre banque.

6.1.4 *Une monnaie de plus en plus dématérialisée ?*

Quand on observe les évolutions des formes modernes et contemporaines de la monnaie, on remarque un processus général de dématérialisation de la monnaie. Autrement dit, la monnaie prend des formes de moins en moins matérielles et de plus en plus immatérielles. Dans le même temps, on observe un second processus de dématérialisation des instruments d'utilisation de la monnaie scripturale. Le

processus de dématérialisation des formes de la monnaie se traduit par une réduction progressive, mais tendancielle, de l'utilisation des pièces et des billets et une croissance simultanée de l'utilisation des instruments de mobilisation de la monnaie scripturale.

Le processus de dématérialisation des instruments de mobilisation de la monnaie scripturale s'observe aussi dans la mesure où les instruments de paiement papier, tels que le chèque et le virement, ont tendance à être de moins en moins utilisés tandis que la carte bancaire l'est de plus en plus. Or, plus la carte bancaire est utilisée (matériellement ou via Internet), plus la dématérialisation des instruments d'utilisation de la monnaie scripturale est grande. En effet, avec une carte bancaire de paiement (qui est matérielle), on peut effectuer de nombreux paiements. Avec un chèque, on ne peut effectuer qu'un seul paiement. On peut, de surcroît, payer des petites sommes de plus en plus faibles.

Cette thèse d'un processus linéaire de dématérialisation de la monnaie peut néanmoins être contestée dans la mesure où la monnaie fiduciaire est apparue en même temps que la monnaie métallique et ses changements de supports. De surcroît, la monnaie scripturale serait apparue au cours de la même période voire avant la monnaie fiduciaire. Ainsi, *« historiquement, les formes initiales de la monnaie scripturale sont apparues avec les premières banques, donc bien avant les billets de banque dont l'invention ne date que du XVII^e siècle »*, comme le rappelle Dominique Plihon dans son ouvrage *La monnaie et ses mécanismes*.

Selon Marie Delaplace, *« des traces d'une forme de comptabilité de dépôt et de prêt ont été trouvées sur des tablettes en terre cuite aux alentours de 1800 avant Jésus Christ. Les Grecs et les Romains, connaissaient le virement entre deux comptes courants ouverts chez le même "trapézite" (changeur de monnaie) ainsi que chez les Arabes dès le IX^e et X^e siècle »* (*Monnaie et financement de l'économie*, Dunod, 2003).

Il ne faut pas oublier la monnaie centrale qui est la monnaie émise par la banque centrale : elle prend la forme de pièces et billets et de monnaie scripturale (solde sur le compte de la banque centrale).

L'observation des économies contemporaines montre que toutes sortes de formes de monnaies ou de mobilisations de celles-ci coexistent. Ce fait souligne la grande adaptabilité des systèmes de paiement pour répondre aux besoins de la circulation des biens et des services.

En matière de monnaie commerciale, deux notions existent : la monnaie au sens étroit et la monnaie au sens large (également appelée quasi-monnaie).

– La monnaie **au sens étroit** est composée de la monnaie manuelle et fiduciaire, avec les pièces métalliques (dites monnaies divisionnaires) et les billets, et de la monnaie scripturale, représentée par les avoirs en comptes à vue des clients auprès des banques commerciales. On doit toutefois noter que le chèque n'est pas de la monnaie, mais un instrument permettant de faire circuler cette monnaie scripturale (au même titre que la carte bancaire).

– La monnaie **au sens large** est composée des placements à vue, c'est-à-dire les avoirs déposés sur les comptes sur livret (livret A, livret B, etc.), des placements à terme constitués pour une durée déterminée et indisponibles avant terme, sauf diminution ou suppression des intérêts, des placements en titres de créances négociables, des encours d'OPCVM (organismes de placement collectif en valeurs mobilières) monétaires, des dépôts à vue libellés en devises.

Selon la Banque de France, en France la monnaie divisionnaire représente 17,1 % de l'ensemble des moyens de paiement en 1980, 12,9 % en 2000 et 18,9 % en 2017. La monnaie scripturale représente 82,9 % en 1980, 87,1 % en 2000 et 81,1 % en 2017.

Pour l'ensemble de la zone euro, la monnaie divisionnaire représente 22,1 % de l'ensemble des moyens de paiement en 1997, 18,8 % en 2000 et 15,2 % en 2017. La monnaie

scripturale représente 77,9 % en 1980, 81,2 % en 2000 et 84,8 % en 2017.

6.1.5 *Les agrégats monétaires*

Les notions de monnaie ou de quasi-monnaie sont reprises dans le concept d'**agrégats monétaires**. Les agrégats monétaires sont les indicateurs de mesure de l'ensemble des liquidités détenues par les agents non financiers.

On distingue :

M1 : Monnaie au sens étroit (billets, pièces et dépôts à vue) ;

M2 : M1 + dépôts à terme d'une durée initiale inférieure ou égale à deux ans

+ dépôts remboursables avec un préavis inférieur ou égal à trois mois ;

M3 : M2 + OPCVM monétaires

+ pensions

+ instruments négociables émis à court terme (durée inférieure ou égale à deux ans).

Selon la Banque de France, en France, en décembre 2017, M3, l'agrégat le plus large, s'élevait à 11 870 milliards d'euros avec $(M3-M2) = 670$ milliards, $(M2-M1) = 3411$, $M1 = 7788$ milliards dont 6676 milliards pour les billets et pièces et 6676 milliards pour les dépôts à vue. La Base Monétaire s'élevant à 2027 milliards.

6.2 MONNAIE ET ACTIVITÉ ÉCONOMIQUE

6.2.1 *L'analyse néoclassique : la neutralité de la monnaie*

Pour la plupart des auteurs « **classiques** »¹⁰, la monnaie n'a qu'une fonction, celle d'intermédiaire des échanges. La présence de la monnaie est un fait incontestable, mais n'est pas fondamentale, dans la mesure où les rapports de valeur entre les objets sont déterminés *avant* l'échange, principalement par les quantités de travail qu'ils incorporent.

L'introduction de la monnaie ne change donc pas fondamentalement les lois de la valeur. Pour ces économistes classiques, la thésaurisation de la monnaie est irrationnelle, car il est absurde de conserver de la monnaie : cela reviendrait pour les agents à se priver de la satisfaction que seule autorise l'utilisation des biens que la monnaie permet d'acheter. Dans cette conception, la monnaie n'est demandée que pour un motif de transaction, toute l'épargne sera soit investie, soit dépensée et la monnaie ne saurait être demandée pour elle-même dans une logique d'épargne. Or, la valeur de la monnaie dépend, toutes choses égales par ailleurs, de la quantité émise, ce qui apporte une justification à la théorie quantitative de la monnaie. Dans l'approche classique et néoclassique, la dichotomie entre la sphère monétaire (celle de la création monétaire) et la sphère réelle (celle de la production réelle de biens et services) implique que la variation de la masse monétaire n'a aucune influence sur l'activité économique. Dès le XVIII^e siècle, David Hume énonce le principe de neutralité à long terme de la monnaie : un doublement de la quantité de monnaie en circulation n'engendre qu'un doublement du niveau général des prix. Dans ses *Principes de l'économie politique et de l'impôt*, David Ricardo montre en 1817 que la masse monétaire et le volume des transactions sont déterminés par des facteurs différents : la quantité de signes monétaires dépend de la production des mines de métaux précieux et des autorités monétaires (émission de papier-monnaie), tandis que les échanges réels sont liés aux besoins des agents économiques et au volume des affaires.

Les auteurs **néoclassiques** ont repris par la suite la théorie quantitative en la reformulant à l'aide d'équations, dont la plus connue est celle de l'économiste mathématicien américain Irving Fisher : dans son ouvrage *Le Pouvoir d'achat de la monnaie* (1911), ce dernier pose l'équation des échanges $M \times V = P \times T$, selon laquelle le produit de la masse monétaire (M) par la vitesse de circulation de la monnaie (V) est égal au produit du volume des transactions (T) par le niveau général

des prix (P). Si certaines hypothèses sont respectées (T ne varie pas lorsque M augmente, la vitesse de circulation V étant constante), la théorie quantitative établit que toute augmentation de la quantité de monnaie entraîne une augmentation du niveau général des prix¹¹. Mais c'est l'analyse de Milton Friedman (prix Nobel d'économie en 1976), appuyée sur d'importantes recherches statistiques, qui réhabilitera la théorie quantitative de la monnaie, sans toutefois accepter l'intégralité de la neutralité de la monnaie : la neutralité monétaire à long terme s'accompagne d'une influence réelle de la monnaie à court terme. Selon les **monétaristes**, l'inflation a essentiellement pour origine un taux d'expansion monétaire trop élevé par rapport aux taux de croissance réel de l'économie, tandis qu'il existe un taux naturel de chômage qui ne dépend que des imperfections du marché du travail, et dont on ne peut s'écarter durablement par une injection de monnaie dans l'économie. La meilleure stratégie des banques centrales en charge de la monnaie consiste alors à neutraliser la politique monétaire en instaurant une norme de croissance de la masse monétaire au même rythme que celle de la production réelle. Les économistes de la **Nouvelle École classique** (NEC)¹² radicaliseront cette proposition, percevant la monnaie comme un épiphénomène et retrouvant en cela la tradition des premiers économistes classiques. Si la théorie quantitative de la monnaie a inspiré la stratégie de nombreuses banques centrales (la Banque Centrale européenne surveille l'évolution des agrégats monétaires comme indicateur avancé de l'inflation) et les politiques de désinflation menées dans les années 1980, la forte croissance de la masse monétaire consécutive aux politiques monétaires de relance décidées après la crise de 2007-2009 inquiète dorénavant certains analystes, face au risque de dérapage des prix ou de bulles spéculatives sur les marchés d'actifs.

6.2.2 *L'analyse keynésienne : la monnaie active*

En rupture avec la théorie néoclassique, John Maynard **Keynes** introduit le **rôle de l'incertitude et des anticipations** dans sa *Théorie générale de l'emploi, de l'intérêt et de la monnaie* (1936). Dans une économie monétaire fondée sur le crédit et l'endettement, les agents économiques, soumis à l'incertitude, peuvent désirer la monnaie pour elle-même. Keynes infirme la loi des débouchés de Say et la théorie de la « monnaie-voile » (qui ne fait que recouvrir les échanges), où la monnaie n'est qu'un simple intermédiaire des échanges : la « **préférence pour la liquidité** » des agents (pour motif de transaction, de précaution ou de spéculation sur le marché des titres) implique une restriction potentielle de la demande et du volume des transactions. L'introduction de facteurs psychologiques (incertitude, crainte de l'avenir) permet à Keynes d'expliquer l'instabilité fondamentale des économies de marché : le taux d'intérêt, variable monétaire, agit sur l'économie réelle en influençant l'incitation à investir des entreprises.

Pour les économistes classiques, la monnaie est neutre, au sens où elle n'a pas d'influence sur l'économie réelle (celle de la production de biens et de services). Dans sa *Théorie générale*, John Maynard Keynes explique au contraire que le taux d'intérêt se fixe sur le marché de la monnaie (et non sur celui du capital, comme l'affirment les néoclassiques) ; la banque centrale peut accroître l'offre de monnaie et faire diminuer le taux d'intérêt, en agissant ainsi sur la variable « motrice » qu'est l'investissement productif, afin de stimuler la production, l'emploi et les revenus distribués. La croissance de la quantité de monnaie en circulation n'accélère pas forcément l'inflation, selon Keynes : elle l'est seulement dans un cas très particulier, lorsque l'économie est au plein emploi. Tant qu'il existe des capacités de production inutilisées dans l'économie, la stimulation monétaire agit positivement sur la production, l'emploi et les prix. Les économistes d'inspiration keynésienne défendent l'idée que la politique monétaire affecte la croissance et l'emploi : en accordant des crédits, les

banques commerciales permettent aux entreprises d'investir et aux ménages de consommer, sur la base d'une anticipation de la production de biens et services à venir. Les prévisions des entrepreneurs (demande effective) et celles des banques doivent donc coïncider pour éviter une création de monnaie excessive (tensions inflationnistes), ou insuffisante par rapport à la production réelle (risque d'aggravation du chômage).

6.2.3 *Les analyses sociologiques de la monnaie*

Les anthropologues et les sociologues ont montré que la monnaie pouvait être profondément imbriquée dans les relations sociales. La monnaie joue un rôle ambivalent quant au lien social : elle contribue à forger une même communauté de paiement et de confiance dans un même étalon de mesure, mais elle peut aussi être génératrice de conflits pour la possession de richesses qui menacent l'équilibre de la société. Dans son livre intitulé *La production des grands hommes*, l'anthropologue Maurice Godelier nous fait découvrir les pratiques d'échanges et d'initiation des Baruyas de Nouvelle-Guinée. Cette tribu, « découverte » seulement en 1951, utilise une monnaie constituée de barres de sel de potassium. Ce sel est produit à partir de cannes à sel cultivées dans ce but. Sa fabrication implique la participation du groupe entier à la culture, à la récolte et à la fabrication. Ce sel est à la fois un bien de consommation pour la pratique cérémonielle et un moyen d'échange (monnaie). C'est également le principal moyen d'échange avec d'autres tribus pour se procurer des moyens de production ou des armes. Dans cette perspective, la monnaie participe à la structuration de la société et à ses relations avec les autres groupes : elle permet la création de liens sociaux.

Dans l'analyse marxiste, la monnaie masque un rapport social inégalitaire de domination du capital sur le travail. Karl **Marx** explique, dans *Le Capital*, que la monnaie est nécessaire à l'échange marchand, qu'elle est un équivalent général, mais que « la monnaie déguise en réalité un rapport

social » (*Le Capital*, livre I). C'est l'argent qui permet au capitaliste l'accumulation du capital à travers un cycle A-M-A ». L'argent A permet l'achat de marchandises M (parmi lesquelles le travail) par le capitaliste. Celui-ci revend des marchandises transformées contre une quantité de monnaie A' supérieure à celle qu'il a consacrée à la production. La monnaie participe ainsi à la production de ce que Marx appelle la plus-value (le profit au sens marxiste). L'exploitation de la force de travail tient à ce qu'une partie de la valeur qu'elle a créée dans le processus de production se voit confisquée par le capitaliste et ne lui est pas payée. La monnaie masque le résultat d'un rapport de force.

Si la monnaie contribue à forger des liens sociaux, elle peut également renforcer l'identification des individus à un même espace politique. Complément du marché unique européen, l'euro a été conçu comme une étape décisive de la construction politique de l'Europe, susceptible d'approfondir l'identité et la citoyenneté européennes. La monnaie constitue traditionnellement un vecteur d'identité et contribue à raffermir le sentiment d'appartenance symbolique à une communauté, une collectivité, tout comme un drapeau, un hymne national ou une Constitution.

6.3 À QUOI SERVENT LES BANQUES ET LES MARCHÉS FINANCIERS ?

6.3.1 *Les fonctions du système financier*

Le **système financier** désigne l'ensemble des institutions (marchés de titres et intermédiaires financiers), des règles et des pratiques qui, au sein d'un espace donné (un pays, une région ou le monde entier), rendent possibles des échanges d'argent. Ces échanges se font entre ceux qui ont accumulé de la monnaie en épargnant chaque année une fraction de leur revenu et ceux qui n'ont pas suffisamment de monnaie pour financer une dépense d'investissement. En l'absence d'un système financier efficace, certains agents devraient

renoncer à des investissements rentables faute de fonds, alors que d'autres disposeraient d'une épargne oisive. Cette mauvaise utilisation des ressources freinerait alors la croissance. Au niveau macroéconomique, on peut distinguer un financement par création monétaire d'un financement rendu possible par la constitution d'une épargne préalable. En effet, lorsque les capacités de financement des uns sont insuffisantes par rapport aux besoins de financement des autres, les banques peuvent créer momentanément, en accordant des crédits, une monnaie additionnelle qui va servir à financer les investissements désirés au-delà de l'épargne réalisée. Ce faisant, elles parient sur le fait que ces financements (dits « monétaires » puisqu'ils sont réalisés par création monétaire) engendreront un supplément de production de biens et de services et de revenus qui permettra de réaliser, au final, une épargne correspondant à cette avance. Cette possibilité doit cependant être contrôlée (le plus souvent par une banque centrale), pour éviter qu'une création de monnaie excessive ne finance des investissements non rentables ou non porteurs de croissance et nourrisse alors l'inflation.

Le système financier permet de mettre en relation les agents à capacité de financement (principalement les ménages) et les agents à besoin de financement (principalement les entreprises et les administrations publiques). Les institutions financières permettent aussi de collecter une multitude d'informations pour aider les agents économiques à prendre leurs décisions : à ce titre, elles contribuent à la bonne affectation des ressources dans l'économie. Comme les agents économiques couvrent généralement leurs dépenses par les revenus qu'ils perçoivent (salaires, recette des ventes, impôts), la contrainte de financement apparaît surtout dans le cadre du financement de leurs investissements.

Pour la comptabilité nationale, qui décompose l'économie en grands secteurs institutionnels, le financement de l'économie désigne à la fois celui des entreprises (les entreprises individuelles, les sociétés non financières et les sociétés

financières) qui cherchent à financer leur investissement productif, celui des administrations publiques qui, le plus souvent, doivent financer leur déficit et celui des ménages qui ont à financer leurs achats de logements. Si chacun de ces agents pouvait financer totalement son investissement par son épargne, il n'y aurait pas besoin d'un système financier très développé. Mais, sans circuits financiers efficaces, les échanges financiers seraient limités, et le potentiel de croissance de l'économie serait considérablement réduit.

La finance est une activité de production et de transferts d'actifs financiers (ceux-ci ayant la particularité d'être toujours à l'actif et au passif d'agents distincts), et donc, fondamentalement, **un réseau de promesses ou d'engagements enchevêtrés**. La finance est censée améliorer l'allocation des ressources dans une économie, à travers l'espace et à travers le temps, sous condition d'incertitude. De ce fait, elle est intrinsèquement instable, puisqu'elle consiste en un commerce d'*anticipations* sur une production future dans un système hyper-connecté.

L'économiste Ross Levine a décrit en 2005 les cinq grandes fonctions économiques du système financier :

- Faciliter l'échange, la couverture et l'assemblage du risque ;
- Produire de l'information pour les entreprises et allouer les ressources dans l'économie ;
- Exercer le contrôle sur les directions des entreprises ;
- Mobiliser efficacement l'épargne disponible ;
- Faciliter l'échange des biens et services par la baisse des coûts de transaction et de collecte de l'information.

6.3.2 *Besoins et capacités de financement*

En règle générale, il y a, chez les agents, un déséquilibre entre les montants à investir et les montants épargnés qui permet de distinguer deux grands profils parmi eux :

– Les **agents à capacité de financement** sont des agents dont les revenus (la monnaie reçue) sont supérieurs aux dépenses (la monnaie utilisée). Ceux-ci financent donc leurs investissements et dégagent, en plus, une épargne financière qui pourra être placée. En comptabilité nationale, cette capacité de financement est le solde du compte de capital et désigne la somme qui reste aux agents lorsqu'on a déduit de leur épargne leurs investissements (leur formation brute de capital fixe – FBCF) soit en logements, soit en capital productifs, ainsi que leurs acquisitions de terrains et d'actifs incorporels comme les brevets. Seuls les ménages présentent une capacité de financement régulièrement positive, ce qui fait de leur épargne la variable stratégique du système financier. Les sociétés financières ont aussi une capacité de financement, mais leur rôle est surtout de recycler celle des ménages.

– Les **agents à besoin de financement** sont dans une situation symétrique. Pour se financer, ils doivent combiner plusieurs types de ressources : des emprunts (sous forme de crédits ou d'émissions de titres comme les obligations), leur épargne antérieure et l'épargne qu'ils réalisent au cours de l'année. Les entreprises sont des agents structurellement en situation de besoin de financement, car leurs projets d'investissement excèdent en général leurs propres ressources. Depuis la fin des années 1970, les administrations publiques sont aussi structurellement emprunteuses et c'est en grande partie pour satisfaire ce nouveau besoin de financement public que les innovations financières sont apparues. Cette distinction est bien sûr une simplification liée à l'approche macroéconomique du phénomène de financement.

En réalité, il peut exister des entreprises ayant une capacité de financement ($\text{investissement} < \text{épargne}$) et des ménages exprimant un besoin de financement ($\text{investissement} > \text{épargne}$), mais, globalement, les ménages

sont en capacité, alors que les entreprises et les administrations publiques sont en besoin.

Le système financier va donc permettre l'ajustement entre capacités et besoins de financement.

6.3.3 *Financement interne et financement externe*

Une autre approche du financement de l'économie peut s'appuyer sur l'origine des fonds vue du côté de l'entreprise. Elle oppose alors le financement interne au financement externe :

- Le **financement interne** vient d'abord des bénéfices non distribués des entreprises. Celles-ci ont en effet deux manières d'utiliser les profits qu'elles dégagent de leurs activités : soit les conserver pour financer leurs investissements nouveaux, soit les distribuer à leurs propriétaires sous forme de dividendes pour récompenser la confiance que ceux-ci leur ont accordée en apportant leur épargne. Ce bénéfice distribué vient alors en déduction des sommes initiales disponibles pour l'investissement. Ce partage interne, qui s'opère dans l'entreprise, peut être source d'une tension importante. Si les bénéfices distribués sont trop faibles, l'entreprise peinera à trouver de nouveaux financeurs ; s'ils sont trop importants, la firme sera obligée de diminuer ses investissements ou d'emprunter des ressources supplémentaires (coûteuses) pour les financer. Ces dernières années, la part des bénéfices distribuée aux actionnaires a fortement progressé dans tous les pays pour dépasser parfois le bénéfice non distribué (voire le bénéfice total), obligeant certaines entreprises à augmenter leur endettement pour compenser ces fuites dans leurs ressources propres. À ces profits s'ajoutent les provisions que constituent les entreprises pour anticiper l'obsolescence du capital qu'elles utilisent. Locaux et machines s'usent en effet au fil du temps et les entreprises

doivent prévoir les ressources financières nécessaires à leur remplacement. Enfin, les entreprises consacrent de plus en plus de dépenses courantes au paiement de leurs activités de recherche-développement (mais aussi de formation ou de développement de leurs marques). Ces dépenses sont souvent qualifiées d'investissements immatériels et, à ce titre, peuvent être considérées comme faisant partie des sommes qui financent l'investissement au sein de l'entreprise. Bénéfices non distribués, provisions pour amortissement et dépenses d'investissement immatériel constituent l'autofinancement des entreprises, c'est-à-dire le financement qu'elles assument à partir de leurs propres recettes.

– Si la forme de financement précédente paraît le meilleur choix pour un agent à la recherche d'une indépendance financière, elle n'est pas en général suffisante, notamment en période de forte croissance. L'agent économique doit faire appel à un **financement externe** qui lui sera fourni par d'autres agents. Les marchés de titres et les intermédiaires financiers se chargent alors de cette nouvelle forme de financement. Le financement externe regroupe ainsi l'ensemble des moyens de financement de l'entreprise autres que l'autofinancement, comme l'appel au marché financier par augmentation de capital (émission de nouvelles actions) ou par emprunt obligataire, ainsi que le recours au crédit bancaire.

6.3.4 *Financement direct et financement indirect*

Lorsqu'un agent souhaite recourir à un financement externe, deux grandes modalités s'offrent de nouveau à lui : la finance directe (« désintermédiée ») et finance indirecte (« intermédiée »).

– Lorsqu'intervient un intermédiaire financier, **le financement est dit « indirect » (ou « intermédié »)**. Cela signifie qu'un intermédiaire va s'intercaler entre les

agents à capacité de financement, auprès desquels il va se charger de collecter des fonds, et les agents à besoin de financement, auxquels il apportera son concours financier. Les intermédiaires les plus connus sont les banques, au travers des opérations de crédit qu'elles réalisent. À côté d'elles se sont développés depuis vingt ans d'autres intermédiaires financiers (organismes de placement collectif en valeurs mobilières, fonds de pension, sociétés d'assurances), qui ont allongé ce circuit de l'intermédiation. Ces « investisseurs institutionnels » achètent en effet des titres directement émis par les entreprises et les administrations, mais peuvent aussi acheter des titres émis par les banques elles-mêmes.

– **Le financement direct (ou « désintermédié »)** désigne, lui, la forme de financement dans laquelle les agents à besoin de financement émettent des titres (titres de propriété comme les actions ou titres de dettes comme les obligations) qui seront acquis directement par les agents ayant une capacité de financement, sur des marchés financiers spécifiques, correspondant chacun à un type de titres vendus. Dans ce cas, un agent économique à besoin de financement s'adresse directement à un agent à capacité de financement. Mais très souvent, cette opération d'appel public à l'épargne (émission d'actions ou d'obligations) se réalise sur un marché qui suppose l'intervention d'un intermédiaire financier dans le cadre de ce qu'on appelle parfois une économie de marché financier (même si l'on doit distinguer plus précisément intermédiation de marché et intermédiation de bilan). La distinction entre ces deux modalités de financement a été à l'origine de plusieurs typologies des systèmes financiers suivant la part occupée par le circuit de la finance directe et de la finance indirecte. Les transformations des systèmes financiers au cours des dernières décennies les ont cependant rendues assez floues, du fait de l'évolution de la finance indirecte.

La finance globalisée a sensiblement complexifié le rôle des intermédiaires financiers : malgré des phénomènes universels à l'échelle mondiale (comme la montée en puissance des investisseurs institutionnels et des fonds spéculatifs), l'opposition traditionnelle entre les systèmes orientés « banques » (économie d'endettement, à l'instar de l'Allemagne) et les systèmes orientés « marchés » (économie de marchés de capitaux, comme aux États-Unis) demeure en partie pertinente pour distinguer des systèmes financiers encore largement différents dans leurs structures, fruits de l'histoire, des règles juridiques, mais aussi de la culture et des mentalités.

Cette mutation financière s'est également accompagnée d'une redéfinition du rôle des banques, désormais largement liées à la finance directe et à l'innovation financière. Cette évolution a été porteuse d'une gestion optimisée de l'épargne, mais également d'une concentration des risques et d'une propagation des chocs, comme l'impact de la crise boursière sur la contraction de l'offre de crédit à l'économie l'a amplement démontré en 2008-2009, jusqu'à menacer d'effondrement le système financier mondial.

6.4 LES BANQUES : DES ACTEURS MAJEURS DU FINANCEMENT DE L'ÉCONOMIE

6.4.1 *Les banques assument le risque de liquidité et le risque de crédit*

Les banques sont des établissements financiers qui collectent les dépôts de leurs clients, disposent du monopole de la gestion des moyens de paiement et assurent la distribution du crédit aux ménages et aux entreprises. Elles facilitent la **circulation de la monnaie** grâce aux différents instruments de paiement (espèces, chèques, cartes de paiement, virements) et concourent ainsi à la croissance du volume des

échanges. Les banques ont diversifié leurs activités et ont désormais d'autres fonctions, telles que le placement de titres (actions, obligations), l'offre de produits d'assurance, les opérations de change, le conseil en gestion de patrimoine, etc. Elles **transfèrent des ressources** et **gèrent les risques** au même titre que les marchés de capitaux : elles collectent l'épargne des agents qui ont une *capacité* de financement pour la distribuer aux agents qui ont un *besoin* de financement (si la banque doit conserver de quoi faire face aux retraits de ses clients, les déposants n'exigent pas *en même temps* la restitution de leurs avoirs). Les banques représentent 70 % du financement externe en Europe, à travers les crédits qu'elles octroient et les titres qu'elles achètent (actions, obligations). Les crédits bancaires sont ainsi essentiels pour les petites et moyennes entreprises (PME), en raison d'un accès aux marchés financiers limité ou inadapté. Même pour les grandes entreprises, qui se financent sur les marchés de titres (finance directe), ce sont souvent les banques qui, au final, détiennent une partie de ces titres émis.

La banque met les ressources des déposants au service du financement des entreprises. Elle « **transforme** » **les dépôts de ses clients en crédits ou titres qui permettent de financer les investissements** des entreprises. Au bilan de la banque, la durée du passif (où sont enregistrés les dépôts) est donc plus courte que celle de son actif (où sont consignés les crédits et achats de titres). On dit généralement que la banque fait « du long avec du court » : c'est ce qu'on appelle la transformation d'échéances.

- Les banques assument un **risque de liquidité** : la fonction spécifique des banques est d'assurer la liquidité en garantissant aux clients la possibilité de récupérer rapidement leurs moyens de paiement afin de financer leurs achats ou de faire face à un imprévu (contrairement à la finance directe). Si le prêteur souhaite pouvoir disposer de ses dépôts rapidement, les emprunteurs ont généralement un horizon plus long : la banque centralise

les moyens de paiement et peut ainsi accorder des crédits à d'autres emprunteurs pour financer des projets d'investissement de grande taille et sur une longue période (on dit que la banque assume un risque de liquidité).

– Les banques sont également soumises au **risque de crédit**, soit la probabilité de défaut de remboursement du client (pertes financières), et doivent ainsi collecter de l'information sur leurs clients (qui sont souvent les déposants). Par ailleurs, la relation entre le prêteur et l'emprunteur est perturbée par le problème des asymétries d'information, car l'emprunteur détient plus d'informations sur la qualité de son projet et sur ses possibilités de remboursement que la banque (s'y ajoute un risque d'aléa moral si l'emprunteur gaspille les ressources prêtées et ne peut au final régler sa dette).

6.4.2 *Globalisation financière et évolution du secteur bancaire*

Avec la globalisation financière, l'activité bancaire s'est davantage orientée vers les marchés financiers. Le mouvement de mondialisation financière à partir des années 1980 s'est accompagné d'un décloisonnement du secteur bancaire : cette ouverture des marchés financiers nationaux a permis aux différents investisseurs (ou emprunteurs) d'intervenir sur les différents marchés (marché monétaire, marché financier, marchés de titres négociables).

La libéralisation des marchés de capitaux a également accru la concurrence au sein du secteur bancaire, même si un certain nombre de règles ont été maintenues, afin de limiter la prise de risques par les banques au niveau européen et international (recommandations de la Banque des règlements internationaux en matière de réglementation prudentielle (« normes prudentielles »)). Les bilans des banques sont devenus nettement plus sensibles à l'évolution des cours boursiers avec l'essor de la finance directe et l'entrée en vigueur des nouvelles normes comptables. Le bilan des

banques s'est transformé : à l'actif, la part des titres financiers a progressé, tandis que la part des crédits a diminué. Au passif, l'endettement a pris des formes nouvelles avec l'émission d'obligations, tandis que la part des dépôts des clients a sensiblement baissé. Nous sommes passés d'une intermédiation de bilan à une intermédiation de marché.

6.5 LE POUVOIR DE CRÉATION MONÉTAIRE DES BANQUES

6.5.1 *La création monétaire par les banques commerciales*

Les banques se sont développées depuis l'époque de la Renaissance, parallèlement au développement de l'échange monétaire : le mot « banque » vient de l'italien *banca*, le banc sur lequel les changeurs étaient installés. Si les opérations de banque sont très anciennes (on trouve déjà la trace de contrats de prêts dans l'Antiquité), leur essor va accompagner celui des échanges et de la croissance capitaliste à l'époque de la révolution industrielle, dont le berceau se situe en Angleterre, au XVIII^e siècle. L'économiste Joseph Schumpeter verra ainsi dans le crédit bancaire l'impulsion indispensable à la mise en œuvre de l'innovation technologique, créatrice de gains de productivité et de croissance. Par la suite, le besoin de capitaux lié aux guerres va relancer la croissance du crédit, tandis que l'émergence du capitalisme industriel au XIX^e siècle va nécessiter la mobilisation d'une épargne plus importante que les simples dépôts des clients pour financer la production d'une économie en profonde transformation (particulièrement en France). Au début du XX^e siècle et à la faveur de la seconde révolution industrielle, les banques combinent progressivement les opérations traditionnelles et les activités de banques d'affaires (opérations sur les marchés boursiers). Le krach boursier de 1929 et l'aggravation de la crise déclenchent un phénomène de panique bancaire : les pertes des banques enregistrées sur

les marchés boursiers poussent les clients à se ruer aux guichets pour réclamer la restitution de leur épargne (*bank run*), et nombre d'établissements sont conduits à la faillite. Le spectre des crises bancaires de la Grande Dépression des années 1930 a justifié par la suite une législation afin de contrôler plus étroitement le secteur bancaire et donné une légitimité à l'intervention rapide et massive de la banque centrale (la « banque des banques »), en tant que « prêteur en dernier ressort », pour sauver les banques. La crise de 2007-2009 a d'ailleurs réactivé les débats de l'époque sur l'organisation du secteur bancaire.

Les banques disposent du **pouvoir de création monétaire** : chaque fois qu'une banque accorde un crédit, la quantité de monnaie en circulation dans l'économie augmente, puisque ce crédit se transforme en dépôt sur le compte du client emprunteur (dont le compte est « crédité », en vertu de la logique selon laquelle « les crédits font les dépôts »), tandis que le remboursement donne lieu à une destruction de monnaie. Le processus de création monétaire est donc essentiellement réalisé sous forme de monnaie scripturale émise par les banques commerciales (ou « banques de second rang »), dans les limites fixées par la banque centrale (ou « banque de premier rang », qui contrôle la quantité de monnaie en circulation dans l'économie afin de préserver le pouvoir d'achat de la monnaie). Les banques doivent donc gérer la croissance du volume du crédit en veillant à maintenir leur marge (profit bancaire), c'est-à-dire la différence entre le coût de leurs ressources et les intérêts reçus des crédits. Cet écart doit permettre de couvrir au moins les frais de fonctionnement et les pertes liées aux défauts de remboursement des clients emprunteurs.

La création monétaire consiste en la transformation de créances sur les agents non bancaires (créances sur l'étranger, créances sur l'économie, créances sur le Trésor public) **en moyens de paiement** immédiatement utilisables pour effectuer des règlements.

Lorsqu'une banque accorde un crédit à un client, elle inscrit à son actif la créance qu'elle détient sur son client et, à son passif, une dette émise sur elle-même représentative du dépôt dont son client dispose grâce à son crédit. De son côté, le client aura, à son passif, une dette vis-à-vis de sa banque, et à son actif, un compte provisionné à hauteur du crédit du montant accordé avant qu'il ne le dépense. La banque crée de la monnaie en monétisant des actifs.

Dans le tableau suivant, en octroyant 100 de crédit, la banque crée 100 de dépôt ; son client a alors une dette de 100 vis-à-vis de la banque et dispose sur son compte de 100 de dépôt à dépenser.

Jeu d'écriture au bilan de la banque		Jeu d'écriture au bilan du client	
Actif	Passif	Actif	Passif
Crédit = 100	Dépôt = 100	Compte de dépôt = 100	Crédit = 100

On voit bien qu'il y a création monétaire, puisque le crédit accordé au client ne s'est pas traduit par une ponction sur les dépôts existants, mais par la création d'une *capacité supplémentaire de paiement*. La banque n'a utilisé aucune ressource préalable pour effectuer cette opération : elle a créé de la monnaie « **ex nihilo** » (à partir de rien). Bien entendu, lorsque l'agent X rembourse le crédit qui lui a été octroyé, il y a destruction de monnaie. La création nette de monnaie suppose donc que les nouveaux crédits accordés au cours d'une période l'emportent sur les crédits remboursés : on dit que « les crédits font les dépôts ».

Dès lors, il y a **accroissement de la masse monétaire lorsque la création de monnaie l'emporte sur la destruction de monnaie**.

L'essentiel de la création monétaire naît donc de la décision d'une banque d'accorder un crédit sur la base d'un projet à un de ses clients. La création monétaire repose donc sur l'activité des banques commerciales : la création monétaire peut être excessive et générer de l'inflation si cette

augmentation de la quantité de monnaie en circulation ne crée pas de richesses.

6.5.2 *Création, destruction de monnaie et variation de la masse monétaire : une synthèse*

Pour savoir s'il y a création monétaire, il suffit de se demander si la masse monétaire augmente. La masse monétaire (MM) correspond à la quantité de monnaie en circulation dans l'économie (QMC), ou encore à la capacité de dépense des agents non financiers (CDANF).

On a donc : $MM = QMC = CDANF$.

$CM \Rightarrow \uparrow MM = \uparrow QMC = \uparrow CDANF$

Étant donné que le crédit bancaire est la principale source de création monétaire, on a :

$\uparrow \text{Crédit} \Rightarrow CM \Rightarrow \uparrow MM = \uparrow QMC = \uparrow CDANF$ et $\Rightarrow \uparrow \text{prix}$ (toutes choses égales par ailleurs).

A contrario, le remboursement de crédits $\Rightarrow DM \Rightarrow \downarrow MM = \downarrow QMC = \downarrow CDANF$.

Dans une économie réelle, globalement et quotidiennement, de nouveaux crédits sont accordés, pendant que d'autres sont en cours de remboursement. On obtient :

$MM \uparrow$ quand $CM > DM$ et $MM \downarrow$ quand $DM > CM$

Autrement dit, quand le montant global des crédits augmente plus vite que le montant global des remboursements, alors $\uparrow MM = \uparrow QMC = \uparrow CDANF$.

A contrario, quand le montant global des crédits augmente moins vite que le montant global des remboursements, alors $\downarrow MM = \downarrow QMC = \downarrow CDANF$.

Bien entendu, compte tenu de l'existence d'autres sources de création monétaire, les achats et ventes de devises et de monnaie domestique peuvent venir modifier l'ampleur des mouvements de la masse monétaire. Enfin, sous l'effet des opérations de transfert de monnaie ou de transformation des formes de monnaie (passage d'une forme à une autre, par

exemple de la forme scripturale à la forme fiduciaire), la masse monétaire reste inchangée.

6.5.3 *Quelle quantité « optimale » de monnaie faut-il créer ?*

La monnaie est créée par le système bancaire : les banques commerciales et la banque centrale sont donc les principaux acteurs de la création monétaire. Pourtant, le réglage de la quantité de monnaie en circulation dans l'économie est difficile, car :

- La quantité de monnaie en circulation dans une économie ne doit pas être trop faible, car les agents économiques seront alors obligés de limiter leurs activités économiques et leurs échanges (consommation, investissement, production, etc.).
- À l'inverse, une quantité de monnaie trop abondante met à la disposition des agents économiques un pouvoir d'achat bien supérieur à la quantité de biens et services disponibles, ce qui peut provoquer une hausse du niveau général des prix (inflation), car la demande progresse plus vite que l'offre.

6.6 LA RÉGULATION DE LA CRÉATION MONÉTAIRE

6.6.1 *Le fonctionnement du système bancaire*

Les banques commerciales et la banque centrale sont les principaux acteurs de cette création monétaire dans le cadre d'un **système bancaire hiérarchisé**. En effet, le système bancaire a un pouvoir de création monétaire à travers la distribution du crédit et il joue un rôle indispensable au financement de l'économie. Le système bancaire évolue en fonction du niveau de développement de l'économie, de l'intensité de la concurrence au niveau national et international, du progrès technique (avec les nouvelles technologies financières, par exemple), les choix de politique

monétaire, l'évolution de la réglementation prudentielle, etc. C'est sur le marché interbancaire, qui est un compartiment du marché monétaire réservé aux banques et à certaines institutions financières, que les banques se procurent les liquidités qui leur sont nécessaires, et à d'autres banques de placer, prêter, à court terme leurs excédents de liquidité. La Banque de France intervient sur ce marché dans le cadre de la politique monétaire définie par la Banque Centrale européenne.

Le rôle principal de la banque centrale est d'assurer la compensation entre les excédents et les déficits de trésorerie quotidiens des banques. Grâce à un compte à la banque centrale, les banques peuvent régler leurs dettes entre elles et obtenir des billets demandés par les agents économiques. La **monnaie centrale** est la monnaie émise par la banque centrale : dans un système bancaire qui est hiérarchisé, chaque banque de second rang émet sa propre monnaie, et cette monnaie ne circule qu'auprès de son propre circuit bancaire. Dès lors, la monnaie centrale est en fait la seule monnaie acceptée par tous les agents économiques et toutes les banques. Chaque banque de second rang dispose d'un compte auprès de la banque centrale. C'est donc par l'intermédiaire de ces comptes que les banques se règlent les dettes entre elles, à la suite des opérations de compensation interbancaire. La monnaie centrale se compose donc de la monnaie fiduciaire et de la monnaie scripturale émises par la banque centrale. Cette monnaie centrale scripturale détenue par les agents financiers (sur le marché interbancaire) n'est pas comptabilisée dans les agrégats monétaires. On dit qu'une banque centrale est « prêteur en dernier ressort », car elle fournit, à des conditions plus ou moins restrictives, selon la politique monétaire, les liquidités dont l'économie a besoin. Dans la zone euro, la Banque Centrale européenne (BCE) définit la politique monétaire des pays appartenant à la zone euro au sein du Système européen des banques centrales (SEBC).

La banque centrale intervient sur le marché monétaire pour prêter de la monnaie banque centrale aux banques commerciales qui ont besoin de se refinancer, moyennant paiement d'un intérêt, et presque toujours en contrepartie d'une créance détenue par les banques (bons du Trésor, effets de commerce, etc.). La banque centrale détermine ensuite le taux d'intérêt auquel elle prête sa monnaie, en fonction des offres et des demandes sur le marché monétaire et, ce faisant, elle joue un rôle d'impulsion pour les autres banques. Elle dirige le marché monétaire et le refinancement des banques commerciales.

Deux situations peuvent alors se présenter sur le marché monétaire :

- Rien n'empêche la banque centrale de prêter sa monnaie à un taux d'intérêt nul ; alors le taux d'intérêt du marché monétaire est également nul, aucune banque ne trouvant d'emprunteur pour un taux positif quand la banque centrale distribue l'argent gratuitement.
- À l'opposé de ce comportement, en théorie, rien n'empêche la banque centrale d'emprunter la monnaie offerte par les banques qui disposent d'excédents en monnaie banque centrale à un taux d'intérêt toujours supérieur à celui offert par les banques emprunteuses. Dans ce cas, tout le monde préfère toujours prêter à la banque centrale, et c'est encore elle qui fixe le taux d'intérêt du marché.

Banque centrale, marché monétaire et refinancement des banques commerciales

Entre ces deux extrêmes, la banque centrale peut faciliter le refinancement des banques, et donc la création monétaire, en offrant beaucoup de liquidités, et en faisant baisser les taux d'intérêt ; ou au contraire, freiner la création monétaire en réduisant son offre de monnaie et en relevant les taux d'intérêt.

La banque centrale est donc bien responsable de la politique monétaire : elle fixe le taux d'intérêt sur le marché monétaire

et agit ainsi sur l'approvisionnement des banques en monnaie banque centrale. Elle contrôle le volume et le prix du refinancement. C'est pour cela d'ailleurs que l'on nomme ce taux d'intérêt manipulé par la banque centrale le taux d'intérêt « directeur ».

En résumé,

- Quand les banques n'ont pu trouver entre elles sur le marché monétaire un financement, la banque centrale *augmente son émission monétaire* afin d'empêcher leur asphyxie.
- En cas d'excès de crédit, la banque centrale *limite la création monétaire* en rendant la monnaie banque centrale plus rare et donc plus chère.

Le taux d'intérêt auquel les banques se refinancent auprès de la banque centrale va influencer sur les taux d'intérêt que ces dernières vont pratiquer pour les crédits à leurs clients (ménages, entreprises). Le taux d'intérêt est le « loyer de l'argent » dans une économie, c'est-à-dire la rémunération du capital prêté : le taux d'intérêt exprime donc un pourcentage de ce capital prêté. Il faut rappeler que le niveau des taux d'intérêt est une variable financière décisive pour l'activité économique. Une hausse des taux directeurs par la banque centrale va pousser les banques commerciales à augmenter le coût du crédit pour rester bénéficiaires, ce qui risque d'inciter les ménages à moins consommer, et les entreprises à moins investir. Au final, l'activité économique sera moins soutenue. À l'inverse, une baisse des taux directeurs par la banque centrale risque d'avoir des effets stimulants, avec pour limite un risque d'accélération de l'inflation.

6.6.2 Les stratégies des banques centrales

La politique monétaire est donc confiée à la banque centrale, généralement indépendante du pouvoir politique. La politique monétaire est l'instrument de politique économique consistant à fournir les liquidités nécessaires au bon fonctionnement et à la croissance de l'économie tout en veillant à la stabilité de la

monnaie. C'est une fonction de régulation de la création monétaire : on peut même appeler cela le **dilemme** de la politique monétaire.

Il existe donc **deux stratégies possibles** pour les banques centrales :

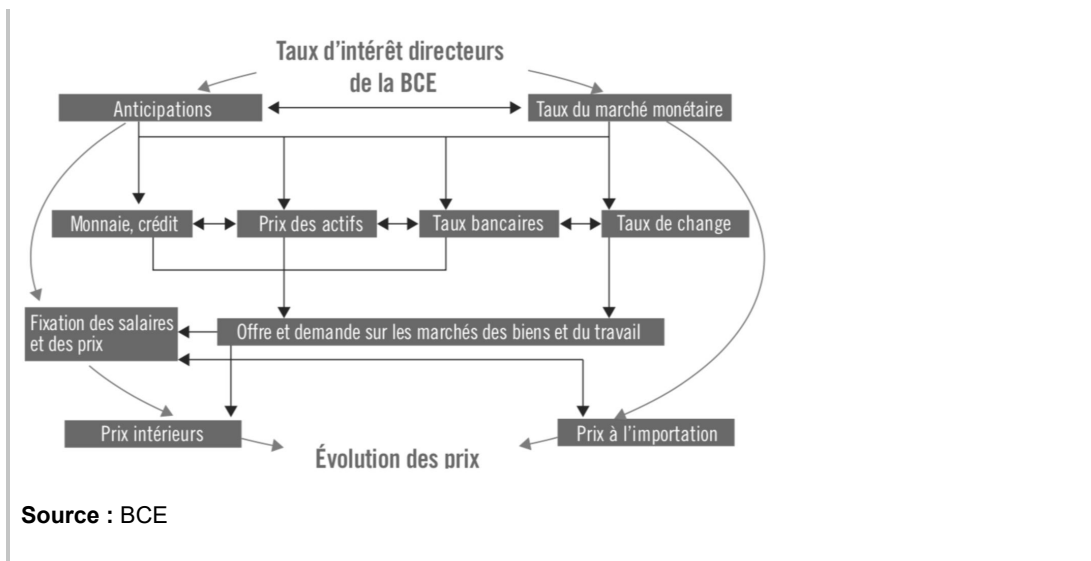
- Lorsque l'économie ralentit, les banques centrales essaient généralement d'abaisser les taux d'intérêt (le coût du capital emprunté). Si les taux sont faibles, les entreprises et les consommateurs seront plus susceptibles d'emprunter et, par conséquent, de dépenser ou d'investir cet argent, ce qui génèrera de l'activité économique.
- À l'inverse, augmenter les taux d'intérêt accroît le coût de l'emprunt, ce qui peut aider à tempérer une économie en surchauffe. Les banques centrales ne fixent pas directement les taux auxquels la plupart d'entre nous empruntent aux banques. Elles déterminent en fait le taux à court terme — au jour le jour — auquel elles prêtent aux autres banques, ce qui influe sur les taux fixés par les autres institutions bancaires.

6.6.3 *Les mécanismes de transmission des effets de la politique monétaire*

Les mécanismes de transmission des effets de la politique monétaire, pour atteindre les objectifs finaux, correspondent aux canaux par lesquels celle-ci va transmettre ses effets à l'ensemble de l'économie en modifiant les comportements économiques monétaires et financiers des agents.

On distingue quatre grands types de canaux de transmission :

- Le canal du taux d'intérêt (canal de la demande de monnaie),
- Le canal du crédit (offre de crédit),
- Le canal du taux de change,
- Le canal du prix des actifs (prise de risque).



GRAPHIQUE 6.1.

La transmission des effets de la politique monétaire : comment les taux d'intérêt influencent-ils les prix ?

En passant par ces divers canaux, la politique monétaire va influencer la demande de biens et services. Au final, ses effets sur l'inflation dépendront aussi des variations de l'offre (capacité de production) et des équilibres/déséquilibres qui en découleront.

6.6.4 L'efficacité de la politique monétaire

L'efficacité de la politique monétaire dépend alors de ces mécanismes de transmission des effets des variations des variables monétaires et financières aux variables de l'économie réelle. On peut distinguer deux types de stratégie : la politique monétaire expansionniste et la politique monétaire restrictive.

- Une **politique monétaire expansionniste** : l'augmentation de la quantité de monnaie se traduit à la fois par une réduction du taux de l'intérêt et par une augmentation du revenu et une dépréciation du taux de change, qui doit provoquer une accélération de l'investissement, un recul de l'épargne et donc une relance de la consommation, soit un accroissement de la demande globale. Cela devrait se traduire par une accélération de

l'inflation, une relance de la croissance et une diminution du chômage.

Toutefois, si le taux d'intérêt ne peut pas baisser (trappe à liquidité) ou si l'investissement n'est pas sensible au taux d'intérêt (inélasticité), le mécanisme de transmission est bloqué et la politique monétaire inefficace.

– Une **politique monétaire restrictive** : une diminution de la quantité de monnaie se traduit à la fois par une hausse du taux de l'intérêt, par une diminution du revenu et une appréciation du taux de change, devrait provoquer un recul de l'investissement, une hausse de l'épargne (et donc un recul de la consommation), soit une réduction de la demande globale. Cela devrait se traduire par une désinflation, un ralentissement de la croissance et une hausse du chômage.

6.6.5 *Le rôle de la politique monétaire en débat*

La régulation de la quantité de monnaie dans l'économie et le rôle de la politique monétaire font l'objet de virulents débats dans le domaine de la théorie macroéconomique, qui restent encore d'actualité. Pour simplifier, on peut ainsi distinguer une **approche « monétariste »** et une **approche « keynésienne »**.

– Dans les années 1950, l'analyse de Milton Friedman a réhabilité la théorie quantitative de la monnaie, sans toutefois accepter l'intégralité de la neutralité de la monnaie : la neutralité monétaire sur le long terme s'accompagne d'une influence réelle de la monnaie à *court terme*, en raison de phénomènes d'illusion monétaire. Mais les économistes de la Nouvelle École classique (NEC) radicaliseront cette proposition, percevant la monnaie comme un épiphénomène, retrouvant en cela la tradition des premiers économistes classiques. La théorie quantitative a inspiré la stratégie de nombreuses banques centrales aujourd'hui (la Banque Centrale européenne surveille l'évolution des agrégats monétaires comme

indicateur avancé de l'inflation) et les politiques de désinflation menées dans les années 1980, en insistant sur les coûts liés à l'inflation (en particulier la détérioration du pouvoir d'achat de la monnaie).

– Selon John Maynard Keynes et les économistes qui s'en inspirent, dans une économie monétaire fondée sur le crédit et l'endettement, les agents économiques, soumis à l'incertitude, peuvent désirer la monnaie pour elle-même comme réserve de valeur. Dans les économies monétaires, le taux d'intérêt se fixe sur le marché de la monnaie (et non sur celui du capital, comme l'affirment les néoclassiques) : la banque centrale peut dès lors relancer l'économie, accroître l'offre de monnaie et faire diminuer le taux d'intérêt en stimulant l'investissement productif, et donc la production, l'emploi et les revenus distribués. Tant qu'il existe des capacités de production inutilisées, la création monétaire augmente la production, l'emploi et les prix.

Les écrits de Keynes sur la théorie monétaire ont été cités pour alimenter la critique des politiques monétaires restrictives des années 1990, fondées sur la désinflation et l'appréciation du taux de change (politique du franc fort et désinflation compétitive en France, controverses sur la parité de l'euro), et accusées de favoriser les placements financiers au détriment des investissements productifs (« tyrannie financière » et « croissance molle »). Les travaux de la Nouvelle Économie keynésienne (NEK) insistent également sur l'efficacité à court terme de la politique monétaire et contestent l'idée d'une neutralité de la monnaie et d'une dichotomie radicale entre le secteur réel et le secteur monétaire, en raison de la viscosité des prix et des salaires.

– **Aujourd'hui, la plupart des banques centrales semblent viser un taux d'inflation faible, mais positif** (situé entre 2 % et 4 %), et les stratégies monétaires de sortie de crise menées en 2007-2009 semblent avoir validé la pertinence du diagnostic keynésien. Toutefois, les

statuts de la Banque Centrale européenne (BCE) restent largement inspirés des enseignements de l'école monétariste et de la Nouvelle École classique (NEC). Selon l'analyse monétariste, la politique de relance monétaire est inefficace, car, même si elle peut permettre une diminution du chômage à court terme, elle n'entraîne qu'une accélération de l'inflation à long terme, tandis que le chômage revient systématiquement à son niveau naturel (déterminé par les variables réelles de l'économie, comme les institutions du marché du travail ou la productivité). La réduction du chômage, selon la Banque Centrale européenne, passe donc prioritairement par des réformes structurelles du marché du travail et des biens. Pour la majorité des économistes aujourd'hui, la relance monétaire est en effet inefficace à long terme, car toute modification prévisible de la masse monétaire en circulation dans l'économie est rationnellement anticipée et la monnaie demeure neutre sur longue période.

– Les stratégies des banques centrales peuvent enfin se heurter à l'écueil de la **trappe à liquidités** : en dessous d'un certain niveau de taux d'intérêt, l'élasticité de la demande de monnaie au taux d'intérêt est infinie et l'élasticité de l'investissement est dans le même temps nulle (tous les agents anticipent une future hausse des taux d'intérêt et donc une baisse de la valeur des titres). Lorsque l'économie se trouve dans la trappe à liquidités, seule une politique budgétaire peut alors permettre un retour vers le plein emploi, car la politique monétaire de relance perd toute efficacité (la baisse des taux d'intérêt ne stimule plus l'investissement). Cette notion retrouve une certaine actualité aujourd'hui face à la contrainte actuelle des taux zéro à laquelle la politique monétaire se heurte. Les banques centrales doivent alors déployer d'autres instruments pour lutter contre la récession (politiques monétaires non conventionnelles).

10 C'est-à-dire les « pères fondateurs » de l'économie comme David Ricardo, Jean-Baptiste Say et John Stuart Mill. Selon Jean-Baptiste Say, la loi des débouchés s'applique dans une économie de marché et « les produits s'échangent contre les produits ». Dans une économie qui fonctionne essentiellement suivant les principes d'une économie de troc, « l'argent n'est que la voiture de la valeur des produits ». La monnaie n'influence donc pas l'activité économique.

11 C'est Irving Fisher qui a précisé les liens entre les variations de la masse monétaire, les taux d'intérêt réels et les fluctuations cycliques : pour un taux d'intérêt nominal donné, la croissance de la masse monétaire diminue les taux d'intérêt réels, stimule l'investissement, la production et provoque une augmentation proportionnelle du niveau général des prix. Malgré des phases transitoires à court terme, la théorie quantitative reste pertinente à long terme. L'école de Cambridge, avec Alfred Marshall et Arthur Cecil Pigou, prolongera les travaux de la théorie quantitative de la monnaie en analysant les effets temporaires à court terme d'une variation de la demande d'encaisses monétaires, de la vitesse de circulation de la monnaie et des taux d'intérêt (avec ce que l'on appelle « l'effet d'encaisses réelles »).

12 Lucas, Sargent, Wallace.

LES ROUAGES FONDAMENTAUX DU MARCHÉ ET SES LIMITES

« L'offre et la demande sont comme les deux lames d'une paire de ciseaux dont on ne sait jamais laquelle des deux coupe le tissu. » - Alfred Marshall, 1890

« Fixez les prix, et les problèmes se multiplieront ; laissez les prix se déterminer d'eux-mêmes, et les problèmes disparaîtront comme par enchantement. » - Milton Friedman, 1975

« Adam Smith s'était demandé dans la Richesse des nations (1776) pourquoi un grand nombre d'agents motivés par leur propre intérêt et prenant des décisions indépendantes ne créaient pas un chaos social dans une économie de propriété privée. Smith lui-même avait acquis une compréhension profonde de la coordination impersonnelle de ces décisions par les marchés des marchandises. Cependant, seul un modèle mathématique pouvait prendre totalement en compte l'interdépendance des variables concernées. En construisant un tel modèle, Walras fonda la théorie de l'équilibre économique général. Walras et ses successeurs avaient conscience que sa théorie serait vide en l'absence d'un argument soutenant l'existence de son concept central. Mais, pendant plus d'un demi-siècle, cet argument n'allait pas au-delà du comptage des équations et des inconnues et de l'égalisation de leur nombre. » - Gérard Debreu, 1987

7.1 Pourquoi échanger sur différents marchés ?

7.2 les principes de base du fonctionnement des marchés concurrentiels

7.3 L'équilibre de l'offre et de la demande : la formation du prix d'équilibre

7.4 Déplacements de l'offre et de la demande : changements d'équilibre

7.5 Les conditions d'efficacité des mécanismes de l'offre et de la demande

7.6 Le fonctionnement des marchés réels : la concurrence imparfaite

7.7 Les politiques de la concurrence

7.8 Quelles sont les principales situations de défaillances du marché ?

On peut considérer que l'organisation des économies contemporaines oscille entre les mécanismes du marché et les principes de l'intervention des pouvoirs publics. Étant donné que la plupart des économies contemporaines sont des économies mixtes, dans lesquelles les principes de fonctionnement des marchés dominent, nous devons nous interroger sur la nature des marchés, les modalités de leur fonctionnement, les situations d'imperfection et de faillance des mécanismes de l'offre et de la demande.

7.1 POURQUOI ÉCHANGER SUR DIFFÉRENTS MARCHÉS ?

Les individus se spécialisent, car ils sont plus efficaces dans certaines activités que dans d'autres. Cette division du travail conduit logiquement à la nécessité d'échanger. Ils échangent, car ils y gagnent : c'est donc le gain à l'échange qui les incite à se spécialiser et les pousse à échanger.

Qu'est-ce qu'un marché ?

Dans les premiers chapitres de ce livre, nous avons présenté l'échange et ses justifications économiques : rareté des ressources, obligation de faire des choix, coût d'opportunité, gains à l'échange... Nous avons distingué les productions marchandes des productions non marchandes. On en déduit donc qu'il y aura des échanges marchands et des échanges non marchands. Mais comment sont organisés ces échanges ? Selon quels mécanismes ? Logiquement, les échanges marchands ont lieu sur des marchés. Ce qui nous conduit à nous poser les questions suivantes : « Qu'est-ce qu'un marché ? », « Comment un marché fonctionne-t-il ? ».

Les multiples dimensions d'un marché

Au-delà de la définition grand public qui en est donnée quotidiennement, pour l'économiste, un marché présente plusieurs dimensions inséparables :

Un marché est un **lieu d'échange, réel ou fictif**. Ce qu'on appelle « marché » existe depuis très longtemps. On sait aussi qu'on peut en rencontrer une très grande diversité. Il y a autant de marchés que de biens et services marchands (pétrole, biens immobiliers, nourriture, marché boursier...). Dans ces conditions, puisqu'il existe une grande variété de marchés, nous avons besoin d'une définition générique, qui s'applique à tous les marchés. D'autant plus que les innovations conduisent à de nouveaux marchés. Un marché est un lieu d'échange, réel ou fictif, entre offreurs et demandeurs. Les offreurs sont les producteurs-vendeurs et les demandeurs sont les acheteurs. Un très grand nombre de nos échanges quotidiens ont lieu sur des marchés. Sur un marché, quel qu'il soit, se fixe un prix auquel se réalisent des transactions. Nous préciserons plus loin les mécanismes de fixation des prix sur différents types de marché.

Un marché est une **forme d'organisation ordonnée des échanges économiques**.

Le marché permet l'apparition d'un ordre économique. L'organisation économique dans une économie de marché permet d'éviter le chaos. Un système de prix et de marchés répond à une logique interne favorisant les échanges de manière ordonnée. Un marché est donc aussi un moyen de communication, de circulation de l'information rendant possible la confrontation entre l'offre et la demande. Le marché fonctionne bien dans la plupart des cas, et les situations de guerre sont là pour nous montrer que l'absence de marché, plus largement la désorganisation économique, est propice à l'avènement de pénuries (sans oublier les effets des catastrophes naturelles, la destruction du capital).

Pourtant, nous verrons que, derrière cette célèbre « main invisible » (selon la métaphore d'Adam Smith formulée en 1759 et 1776) se cache, selon les théoriciens de l'économie de marché, un agent qui joue un rôle central, un personnage qui est le pivot du marché et qui incarne cette main invisible, que personne n'a jamais vue. En théorie, pour que cet ordre spontané existe, il doit exister un « secrétaire de marché » ou un « commissaire-priseur », chargé de coordonner les offres et les demandes individuelles.

Un marché est un **ensemble de mécanismes d'ajustement** entre les **décisions d'offre** des uns et les **décisions de demande** des autres, qui permet les échanges et, *in fine*, la satisfaction des besoins. Le marché correspond à un mécanisme permettant aux demandeurs-acheteurs et aux offreurs-vendeurs de se rencontrer pour échanger des biens et services. Ils sont tous en interaction les uns avec les autres pour participer à la détermination des prix des échanges et des quantités échangées. Les prix représentent les conditions dans lesquelles les individus et les firmes acceptent d'échanger sur les marchés de biens et services, sur les marchés de facteurs de production... Le marché permet de satisfaire les besoins des demandeurs solvables exclusivement. C'est le principe de l'exclusion par le prix, une sorte de suffrage censitaire, que nous retrouverons plus loin.

Le système de prix participe, en outre, à la résolution des trois problèmes économiques fondamentaux. Pour comprendre comment une économie de marché résout les trois problèmes économiques fondamentaux, on peut se fonder sur l'idée des bulletins de vote en euros. Ces bulletins correspondent à des décisions d'achats solvables. Les firmes ayant des perspectives de profit doivent anticiper le montant à venir de ces bulletins de vote en euros, c'est-à-dire le montant anticipé de cette demande solvable... afin de mettre en œuvre la combinaison productive permettant de répondre à cette demande. Le commerce entre nations, ou commerce international, dépend aussi des coûts relatifs des différents biens échangés et des facteurs de production. C'est ce système de prix et de marchés qui va permettre de décider ce qui sera produit et ce qui ne le sera pas : c'est la réponse au « **que produire ?** ». La concurrence, en stimulant l'amélioration de la compétitivité, va permettre de déterminer la réponse au « **comment produire ?** ». En effet, pour les entreprises, le meilleur moyen de rester concurrentielles et d'améliorer la compétitivité prix, afin de maximiser le profit, est de conserver des coûts de production les plus bas en adoptant les méthodes de production les plus efficaces. Enfin, la confrontation de l'offre et de la demande qui va en résulter sur les différents marchés permettra de déterminer pour qui les biens et services sont produits. C'est la réponse au « **pour qui produire ?** ».

Les marchés de facteurs déterminent les niveaux de salaire, de taux d'intérêt et de profits, de revenus fonciers... Ce sont les prix des facteurs. En faisant la somme de tous ces revenus, on peut calculer le revenu global. La répartition des revenus entre les individus est ainsi déterminée par les quantités de facteurs disponibles et les prix de ces facteurs. Nous avons présenté plus tôt cette répartition primaire et la redistribution dont elle fait l'objet pour de nombreuses raisons.

Un marché est un **ensemble de mécanismes de coordination de décisions décentralisées** qui va permettre

la **convergence** entre le **bien-être individuel**, fondé sur des décisions plus ou moins rationnellement égoïstes, et le **bien-être collectif**, non intentionnellement recherché.

Un marché est un **ensemble d'institutions**, au sein desquelles les droits de propriété jouent un rôle fondamental. C'est une institution sociale. Un marché ne peut fonctionner sans institutions. Dans la réalité, et cela est observé et confirmé historiquement, un marché a besoin d'institutions multiples pour fonctionner. Les institutions économiques désignent un ensemble de règles qui s'imposent aux individus dans le cadre du marché et qui définissent les conditions des transactions. Pour qu'un marché existe, il faut donc des règles qui organisent les échanges. Il n'y a pas de marché sans règles. Ces règles sont souvent des lois ou des règlements. Elles constituent le cadre préalable des transactions marchandes. Donc, le libre-échange n'est pas aussi libre que cela. Le libre-échange existe au sens où on peut faire des choix, mais ces choix sont d'abord encadrés, contraints par des règles. Les règles du jeu favorisent l'ordre et le bon déroulement du jeu économique marchand pour l'ensemble des joueurs.

Parmi ces institutions, les droits de propriété sont essentiels. Un marché a d'abord besoin de droits de propriété car, très concrètement, sur un marché, on échange des droits de propriété. Aucun marché ne peut fonctionner sans **droits de propriété**. Ces derniers instaurent les bases de la **confiance** dans les échanges marchands. Or, sans confiance, il n'y a pas d'échanges ! En effet, il n'est pas possible de vendre ce qui ne nous appartient pas, ni d'acheter à autrui ce qui ne lui appartient pas. L'absence de droits de propriété interdit tout marché, tout échange marchand. Ces droits de propriété ne sont pas des données de la nature, il faut les instituer. Ainsi, avec le développement des technologies de l'information et de la communication (TIC), les internautes se sont mis à échanger spontanément de la musique, des films en « peer to peer », et ce n'était pas illégal. Il n'existait pas de lois pour empêcher ou encadrer ce nouveau type d'échanges. La

définition des droits de propriété sur ces produits n'était plus valable. On peut retenir de cet exemple que l'innovation peut conduire à redéfinir les droits de propriété, les conditions d'échange, et donc à mieux circonscrire certains marchés. On peut encore comprendre le caractère impératif des droits de propriété pour que les échanges puissent avoir lieu dans une économie de marché quand on aborde les échanges d'organes, l'esclavage, la prostitution, des biens communs tels que les quotas d'émission de CO₂, le bois de la forêt amazonienne qui appartient au Brésil, la propriété de la terre (de terrains) qui étaient là bien avant que l'homme ne peuple la planète... **La monnaie** est une autre institution centrale des économies de marché (voir chapitre 6). Elle facilite les échanges et son institution doit fonder la confiance que les agents économiques placent en elle pour accepter de l'utiliser. La disparition d'une monnaie par perte de confiance de la part de ses utilisateurs peut détruire un marché en rendant les échanges compliqués, voire impossibles. Ce point sera développé dans le chapitre suivant.

Il ne faut pas confondre marché et capitalisme, économie de marché et économie capitaliste. Le marché est un système d'échanges. Le capitalisme est un système de propriété privée des moyens de production qui permet l'accumulation de richesse par quelques-uns au détriment du plus grand nombre et qui fonctionne d'autant mieux que, sur les marchés, la concurrence est imparfaite.

7.2 LES PRINCIPES DE BASE DU FONCTIONNEMENT DES MARCHÉS CONCURRENTIELS

Pour comprendre le fonctionnement d'une économie de marché, de manière générale, et d'un marché en particulier, nous devons nous appuyer sur un des modèles les plus connus de la science économique : le **modèle de l'offre et de la demande**. Dans cette partie, nous en précisons les grands principes de fonctionnement, sous certaines conditions. Plus tard, nous présenterons les conditions précises, strictes et indispensables pour que les mécanismes de marché soient complètement efficaces, afin de résoudre

les trois problèmes économiques fondamentaux et tous les autres. Ensuite, nous présenterons les principaux types d'imperfections et de défaillances. Ceci nous permettra enfin de justifier certaines interventions des pouvoirs publics et de l'État, au sens large, en particulier.

Ce modèle que nous allons étudier ne permet pas d'expliquer tous les échanges, étant donné qu'il peut exister des échanges non marchands. Il ne prétend pas non plus expliquer parfaitement le fonctionnement des marchés qui gouvernent nos échanges quotidiens. Pour autant, il permet de comprendre, au moins partiellement, le fonctionnement de nombreux marchés et une large part des mécanismes économiques, financiers... de nos économies de marché encadrées, régulées par l'État.

Dans le cadre de ce modèle, l'approche est résolument microéconomique. Sur chaque marché, il existe des comportements d'offre et des comportements de demande. Il convient donc d'en présenter d'abord les principes de base avant d'aborder la question des échanges et de leurs conditions sur les marchés.

7.2.1 *La courbe ou fonction de demande*

Une **courbe ou fonction de demande** représente, toutes choses égales par ailleurs, la relation entre prix et quantité demandée. Sur une courbe de demande, la quantité demandée d'une marchandise s'élève quand son prix baisse et diminue quand son prix augmente. On peut observer cette relation décroissante sur la courbe D_1 dans le graphique 7.1.

Autrement dit, lorsqu'on se déplace le long d'une courbe de demande, la quantité demandée d'une marchandise varie en sens contraire de la variation de son prix.

Ici, on cherche à isoler l'effet des variations du prix sur la quantité demandée d'un bien.

Cela ne veut aucunement dire qu'on pense que seule la variation du prix est déterminante de la quantité demandée. C'est pour cela que, pour être rigoureux, on devrait toujours

préciser que l'on raisonne « *toutes choses égales par ailleurs* » ou *ceteris paribus*.

En revanche, lorsque la courbe de demande se déplace, la courbe d'offre étant inchangée, la quantité demandée d'une marchandise varie quel que soit le prix de la marchandise.

Notez bien les termes utilisés. Nous parlons de demande pour désigner la courbe ou la fonction. En revanche, nous relierons le prix à la quantité demandée en chaque point de la courbe. Demande et quantité demandée ne désignent pas la même chose. Les confondre peut conduire à des raisonnements complètement erronés et à des conclusions fausses !

7.2.1.1 La loi de décroissance de la demande et les déplacements le long de la courbe de demande

Toutes choses égales par ailleurs, lorsque le prix d'un bien augmente, la quantité demandée diminue. Toutes choses égales par ailleurs, lorsque le prix d'un bien diminue, la quantité demandée augmente. Toutes choses égales par ailleurs, le long de la courbe de demande, prix et quantité demandée varient en sens contraire. Ou encore, lorsqu'on se déplace le long d'une courbe de demande, la quantité demandée d'une marchandise varie en fonction inverse (sens inverse) de son prix. Notez bien que le long de la courbe, donc toutes choses égales par ailleurs, la causalité va de la variation du prix à la variation de la quantité offerte.

7.2.1.2 Les autres déterminants de la demande : les facteurs significatifs

Quand nous levons l'hypothèse « toutes choses égales par ailleurs », cela veut dire qu'on souhaite prendre en considération tous les autres facteurs qui peuvent influencer les décisions des demandeurs et qui peuvent donc faire varier la quantité demandée au-delà des seuls effets des variations du prix du bien considéré.

On peut distinguer des **facteurs objectifs** et des **facteurs subjectifs** de la demande des agents économiques. Ces

facteurs correspondent à des variables économiques qui vont se traduire par des déplacements de la courbe de demande et non des déplacements le long de la courbe de demande, en fonction des seules variations du prix du bien.

De nombreuses **variables** sont à l'**origine du déplacement de la courbe de demande**. Nous allons présenter les plus importantes.

Ici, la quantité demandée d'un bien ne dépend plus seulement du prix. De nombreuses variables peuvent aussi déterminer la demande, comme le revenu, les préférences, les anticipations et le nombre d'acheteurs. Dans ce cas-là, on sort du raisonnement « toutes choses égales par ailleurs » : on ne se déplace plus le long de la courbe, mais c'est la courbe elle-même qui se déplace.

– **Le revenu** : un revenu plus faible réduit la quantité demandée d'un bien normal. Ceci est surtout vrai pour un bien dit normal, c'est-à-dire un bien pour lequel, toutes choses égales par ailleurs, une augmentation du revenu conduit à une augmentation de la demande, quel que soit le niveau de prix. Mais tous les biens ne sont pas des biens normaux, et l'effet des variations du revenu peut changer selon le type de bien. Il peut exister des biens dits inférieurs pour lesquels la demande augmente, ou ne diminue pas, malgré la hausse des prix, même si le revenu diminue. Un bien inférieur est un bien pour lequel, toutes choses égales par ailleurs, une augmentation du revenu provoque une baisse de la demande.

– **Les prix des biens proches** : les prix des biens proches peuvent affecter les quantités de biens que nous demandons, soit parce qu'ils sont substituables (autrement dit l'un peut remplacer l'autre), soit parce qu'ils sont complémentaires (autrement dit, l'un ne va pas sans l'autre). Les biens sont **substituables** quand la baisse du prix d'un bien réduit la quantité demandée d'un autre bien ; ces biens sont considérés comme des substituts. Pour des biens substituables, l'augmentation du prix de l'un

engendre une augmentation de la quantité demandée de l'autre. Les substituts sont souvent des biens qui sont alternativement comme les pâtes ou le riz, les vestes ou les blousons, les tickets de cinéma ou les abonnements à des chaînes payantes (*pay per view*) comme Netflix, OCS, Prime Vidéo, ou encore pour la musique, Spotify, Deezer...

Les biens **complémentaires** sont des biens tels que la hausse du prix de l'un entraîne une baisse de la demande de l'autre, ou encore lorsque la baisse du prix d'un bien fait augmenter la demande d'un autre bien. Les compléments sont souvent des paires de biens qui sont utilisés ensemble, comme le pétrole et les voitures, les ordinateurs et les logiciels, le pain et le fromage, les fraises et la crème, le lard et les œufs, les crêpes et le Nutella.

– **Les préférences et les goûts des consommateurs** sont des déterminants subjectifs, mais incontournables, pour expliquer les variations de la demande. Peu importe le prix, si nous n'aimons vraiment pas un bien, la demande sera nulle. De plus, la hausse du prix n'affectera pas ou peu la quantité demandée.

– **Les anticipations** que nous pouvons formuler sur l'avenir en ce qui concerne notre emploi, nos revenus... peuvent influencer notre demande de biens ou de services aujourd'hui. L'incertitude inhérente à l'avenir peut peser plus ou moins lourd dans nos choix, et peu important les variations de prix. Par exemple, si vous vous attendez à percevoir un revenu plus élevé le mois prochain, vous pourriez être tenté de dépenser une partie de votre épargne actuelle. Pour prendre un autre exemple, si vous anticipez que le prix de la glace baissera demain, vous seriez peut-être moins tenté d'en acheter un cornet au prix d'aujourd'hui.

– **Le nombre d'acheteurs** va déterminer la demande globale (soit la somme des demandes individuelles) sur un marché. Or, quand, pour diverses raisons, le nombre

d'acheteurs augmente, le bien peut devenir rare et son prix peut augmenter.

La demande du marché ou demande agrégée est égale à la somme des demandes individuelles.

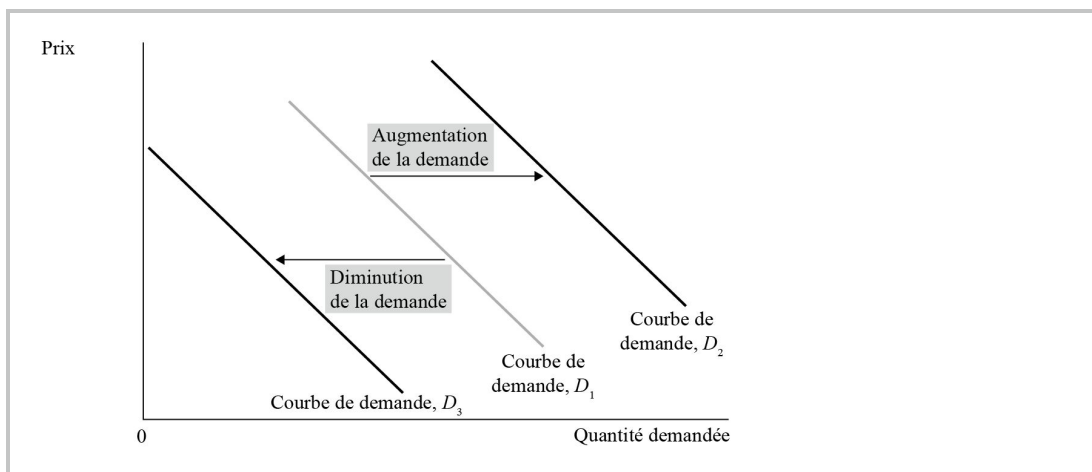
L'élasticité de la demande d'un bien par rapport aux variations du prix de celui-ci désigne la sensibilité de la quantité demandée aux variations du prix.

Quand une faible variation du prix entraîne une forte variation de la quantité demandée, on dit que la demande est très élastique, sensible. La courbe sera peu pentue et plus plate.

En revanche, quand une forte variation du prix entraîne une faible variation de la quantité demandée, on dit que la demande est peu élastique. La courbe de demande sera très pentue et plutôt verticale. Si une faible variation du prix entraîne une forte variation de la quantité de demande, alors la demande est très élastique. Par exemple, pendant les soldes, la quantité de demande est élevée et le prix est faible. La demande est alors élastique.

7.2.1.3 Déplacements de la courbe de demande

Lorsqu'on déplace une courbe de demande, la courbe d'offre étant inchangée, la quantité demandée varie, quel que soit le prix de la marchandise considérée. Lorsque la courbe de demande se déplace, on sort du raisonnement « toutes choses égales par ailleurs ». Désormais, la quantité demandée d'un bien ne dépend plus seulement du prix de ce bien. Pour un même niveau de prix d'un bien donné, la quantité demandée peut varier sous l'effet d'un autre déterminant de la demande. Lorsque la courbe de demande se déplace vers la gauche, cela signifie qu'il y a diminution de la demande et diminution du prix et lorsqu'elle se déplace vers la droite, il y a augmentation de la demande et augmentation du prix.



GRAPHIQUE 7.1.

La courbe de demande et les déplacements de la courbe de demande

7.2.2 La courbe ou fonction d'offre

La courbe d'offre ou fonction d'offre représente, toutes choses égales par ailleurs, la relation entre prix et quantité offerte. Sur une courbe d'offre, la quantité offerte d'une marchandise s'élève quand son prix augmente et diminue quand son prix baisse. On peut observer cette relation décroissante sur la courbe O_1 dans le graphique 7.2.

Autrement dit, lorsqu'on se déplace le long d'une courbe d'offre, la quantité offerte d'une marchandise varie dans le même sens que la variation de son prix.

Ici, on cherche à isoler l'effet des variations du prix sur la quantité offerte d'un bien.

Cela ne veut aucunement dire qu'on pense que seule la variation du prix est déterminante de la quantité offerte d'un bien donné. C'est pour cela que, pour être rigoureux, on devrait toujours préciser que l'on raisonne « toutes choses égales par ailleurs ».

En revanche, lorsque la courbe d'offre se déplace, la courbe de demande étant inchangée, la quantité offerte varie quel que soit le prix de la marchandise.

Notez bien les termes utilisés. Nous parlons d'offre pour désigner la courbe ou la fonction. En revanche, nous relierons le prix à la quantité offerte en chaque point de la courbe. Offre et quantité offerte ne désignent pas la même chose. Les confondre peut conduire à des raisonnements complètement erronés et donc à des conclusions fausses !

7.2.2.1 La loi de croissance de la courbe d'offre et les déplacements le long de la courbe d'offre

Toutes choses égales par ailleurs, afin de maximiser le profit, lorsque le prix d'un bien augmente, la quantité offerte augmente et lorsque le prix d'un bien diminue, la quantité

offerte diminue. Ceteris paribus, le long de la courbe d'offre, prix et quantité offerte varient dans le même sens.

Notez bien que le long de la courbe, donc ceteris paribus, la causalité va de la variation du prix à la variation de la quantité offerte !

7.2.2.2 Les autres déterminants de l'offre et les déplacements de la courbe d'offre

Quand nous levons l'hypothèse « toutes choses égales par ailleurs », cela veut dire qu'on souhaite prendre maintenant en considération tous les autres facteurs qui peuvent influencer les décisions des offreurs et qui peuvent donc faire varier la quantité offerte au-delà des variations du prix du bien considéré. Ces facteurs correspondent à des variables économiques qui vont se traduire par des déplacements de la courbe d'offre et non des déplacements le long de la courbe d'offre, en fonction des seules variations du prix du bien considéré. De nombreuses **variables** sont **à l'origine du déplacement de la courbe d'offre**. Voici les plus importantes : le prix des facteurs de production, la technologie, les anticipations et le nombre de vendeurs.

– Les prix des facteurs de production

Pour produire, nous le savons, nous devons utiliser différents moyens de production. Du capital circulant faisant l'objet de consommations intermédiaires et du capital fixe. Toutes choses égales par ailleurs, la hausse du coût de l'un de ces facteurs doit se traduire par une hausse du coût de production. Or, lorsque le prix de l'un ou plus de ces facteurs augmente, l'entreprise devient moins compétitive et produire risque de devenir moins rentable (pour un prix inchangé). Si la hausse du prix d'un facteur est importante, elle peut même conduire l'entreprise à la faillite. Ainsi, on peut observer une relation négative entre le prix des facteurs et l'offre de biens. La courbe d'offre se déplace vers la droite en cas de baisse des coûts et vers la gauche en cas de hausse des coûts.

– **Le progrès technique**, tel que nous l'avons défini, peut modifier radicalement la capacité de production et donc l'offre d'une entreprise. Par la réduction des coûts de production qu'il peut entraîner, le progrès technique peut se traduire par une forte hausse de la production, car l'entreprise devient plus compétitive. La courbe d'offre se déplace vers la droite.

– **Les anticipations** vont aussi influencer les offreurs et donc les quantités offertes, quels que soient les niveaux des prix. Si les dirigeants d'une entreprise sont optimistes, pour des raisons objectives ou subjectives, ils vont augmenter leur offre (avec hausse simultanée des stocks). La courbe d'offre se déplace vers la droite. Cette abondance pourrait faire baisser les prix. *A contrario*, s'ils sont pessimistes, ils vont déstocker au lieu de produire davantage. La courbe d'offre se déplace vers la gauche. La rareté va augmenter et pourrait produire une hausse des prix.

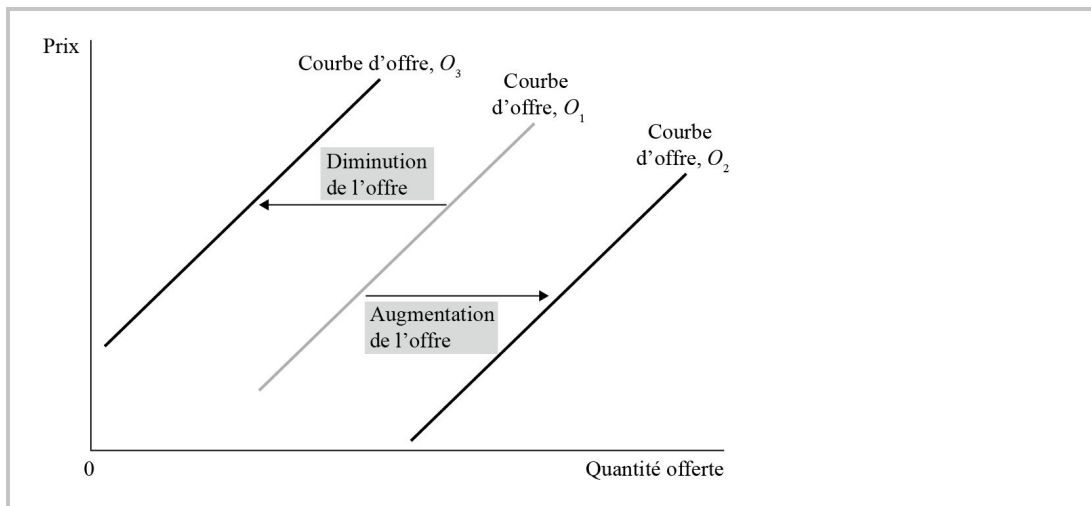
– **Le nombre de vendeurs** sur un marché va avoir un effet sur la quantité globale produite et offerte. Si le nombre de vendeurs augmente et se traduit par une hausse de l'offre, l'effet d'abondance qui en découle peut se traduire par une baisse des prix, en raison de la concurrence. La courbe d'offre se déplace vers la droite. Si le nombre de vendeurs baisse et se traduit par une baisse de l'offre, l'effet de rareté qui en découle peut se traduire par une hausse des prix, en raison de la plus faible concurrence. La courbe d'offre se déplace vers la gauche. L'offre du marché ou offre agrégée est égale à la somme des offres individuelles.

ÉLASTICITÉ DE L'OFFRE PAR RAPPORT AU PRIX

L'élasticité de l'offre d'un bien par rapport aux variations du prix de celui — ci désigne la sensibilité de la quantité offerte aux variations du prix. Quand une faible variation du prix entraîne une forte variation de la quantité offerte, on dit que l'offre est très élastique, sensible. La courbe sera peu pentue et plus plate. En revanche, quand une forte variation du prix entraîne une faible variation de la quantité offerte, on dit que l'offre est peu élastique, sensible. La courbe d'offre sera très pentue et plutôt verticale.

7.2.2.3 Déplacements de la courbe d'offre

Lorsque la courbe d'offre se déplace, la courbe de demande étant inchangée, la quantité offerte varie, quel que soit le prix de la marchandise considérée. Lorsque la courbe d'offre se déplace, on sort du raisonnement « toutes choses égales par ailleurs ». Désormais, la quantité offerte d'un bien ne dépend plus seulement du prix de ce bien. Pour un même niveau de prix d'un bien donné, la quantité offerte peut varier sous l'effet d'un autre déterminant de l'offre. Donc, lorsque l'offre augmente, la courbe se déplace vers la droite et le prix du bien baisse. À l'inverse, lorsque l'offre diminue, la courbe se déplace vers la gauche et le prix du bien augmente.



GRAPHIQUE 7.2.

La courbe d'offre et les déplacements de la courbe d'offre

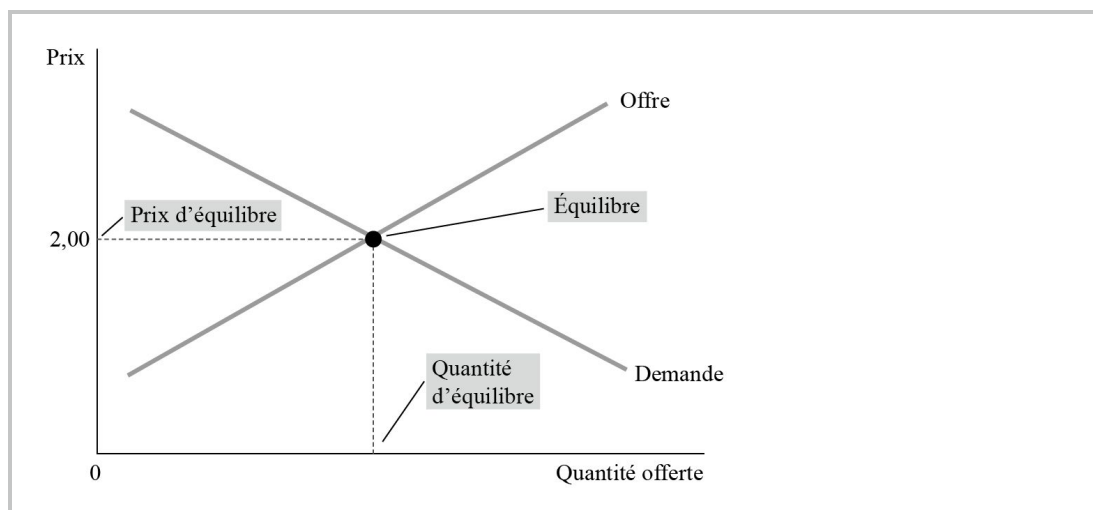
7.3 L'ÉQUILIBRE DE L'OFFRE ET DE LA DEMANDE : LA FORMATION DU PRIX D'ÉQUILIBRE

7.3.1 *L'équilibre du marché*

L'équilibre du marché s'observe à l'intersection entre la courbe de demande et la courbe d'offre : il est nommé E, il nous donne un prix d'équilibre, ainsi qu'une quantité

d'équilibre (QE, telle que $Q_d = Q_o$). L'intersection entre les courbes d'offre et de demande représente l'équilibre du marché. Au point d'équilibre, il y a égalité entre les quantités effectivement échangées, mais cette égalité correspond rarement à une situation où tous les offreurs et tous les demandeurs sont satisfaits. Seuls ceux qui acceptent d'échanger à ce prix obtiennent satisfaction. Mais ce n'est pas parce que tous ceux qui échangent sont satisfaits, que tous ceux qui auraient voulu échanger le sont également. Au niveau du prix d'équilibre, seuls ceux qui acceptent d'échanger à ce prix vont échanger, les autres, auxquels ce prix ne conviendra pas, resteront insatisfaits.

Dans le cadre de cet équilibre, tous les agents économiques ne sont pas forcément satisfaits. Par exemple, si le prix sur le marché d'un bien est de huit euros, des offreurs souhaitaient le vendre plus cher que le prix du marché, à quinze euros. Ils décident de ne pas le mettre sur le marché. De même, si le prix d'un bien est de dix euros, des demandeurs souhaitaient l'acquérir moins cher, pour cinq euros. Ils décident donc de ne pas l'acheter.



GRAPHIQUE 7.3.

La représentation graphique de l'équilibre sur le marché

On peut donc en conclure que « toutes choses égales par ailleurs », les seuls qui sont satisfaits sont ceux qui ont

échangé et accepté ce prix de marché. Il **ne faut donc pas confondre l'équilibre graphique comptable ($Q_o = Q_d$) et l'équilibre économique**, qui désigne ici une situation où tous les agents économiques qui voudraient échanger pour satisfaire leurs besoins le pourraient ! Or, ce n'est pas le cas.

7.3.2 Deux types de déséquilibre

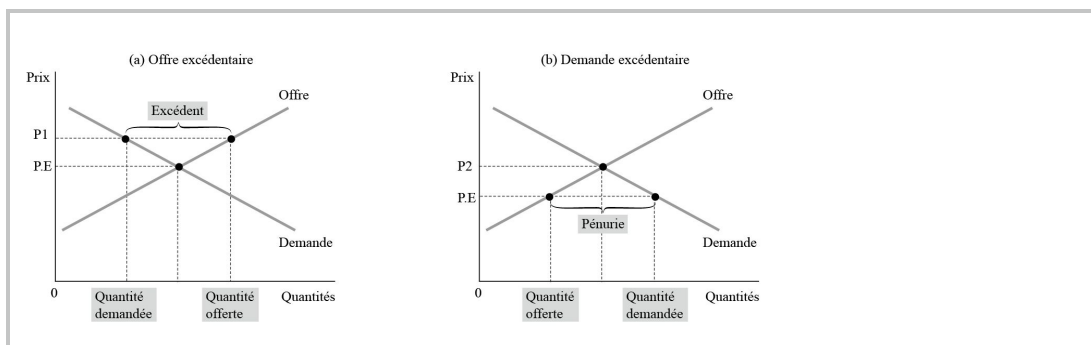
Dans la réalité, le prix d'équilibre est rarement atteint. On observe deux situations typiques de déséquilibre entre Q_o et Q_d : l'abondance et la pénurie.

– Lorsque l'offre est supérieure à la demande, le marché est en situation d'abondance, d'excédent. L'offre est rationnée, car elle ne peut pas vendre tout son stock. L'offre va alors être obligée de baisser son prix pour s'approcher du prix d'équilibre et ainsi augmenter la demande.

– Lorsque la demande est supérieure à l'offre, le marché est en situation de pénurie. La demande est rationnée, car les acheteurs sont trop nombreux et ne peuvent pas tous acheter. L'offre va alors augmenter les prix pour faire baisser la demande et ainsi retrouver un prix d'équilibre.

– $P < P_E \Rightarrow Q_d > Q_o$: Dans ce cas, quand le prix est trop faible, il y a rareté ou excédent de la quantité demandée, donc déficit de la quantité offerte.

– $P > P_E \Rightarrow Q_o > Q_d$: Dans ce cas, quand le prix est trop élevé, il y a abondance ou excédent de la quantité offerte, donc déficit de la quantité demandée.



GRAPHIQUE 7.4.

Les marchés qui ne sont pas à l'équilibre et les équilibres avec rationnement

Sur un marché en déséquilibre, on dit que c'est la **règle du côté court** qui s'applique. L'échange se réalise logiquement au niveau de la plus petite des deux quantités ! Quand la règle du côté court s'applique, le côté long est donc rationné.

– Quand $P < P_E \Rightarrow Q_d > Q_o$, alors les demandeurs sont rationnés. En effet, une partie d'entre eux ne sont pas satisfaits.

– Quand $P > P_E \Rightarrow Q_o > Q_d$, alors les offreurs sont rationnés. En effet, une partie d'entre eux ne sont pas satisfaits.

7.3.3 Les déséquilibres et la dynamique du retour à l'équilibre : le mécanisme de tâtonnement

Les situations de rationnement entraînent un retour à l'équilibre via un **mécanisme de tâtonnement**.

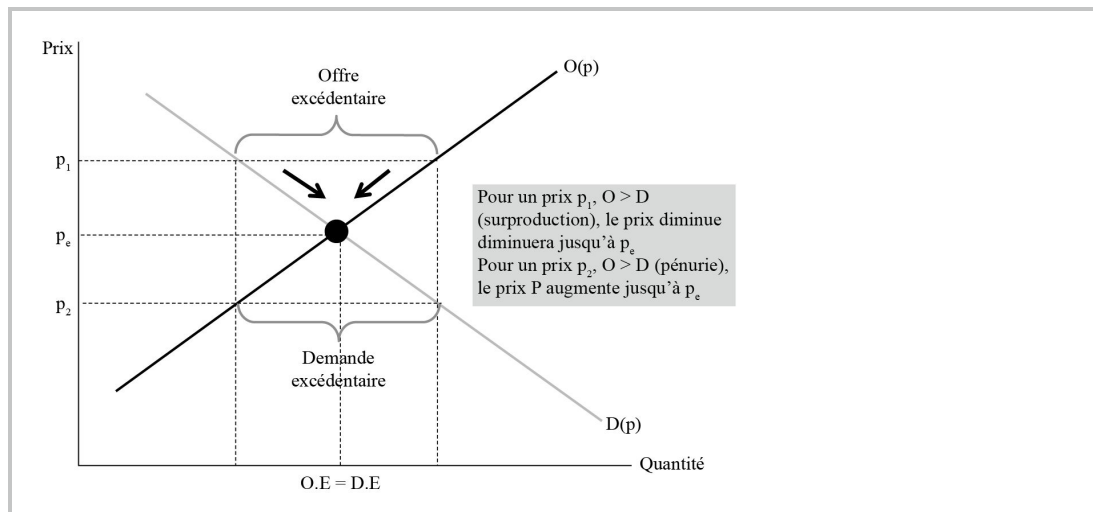
– **Cas n° 1.** Que se passe-t-il, sur un marché quelconque, lorsque la quantité offerte est supérieure à la quantité demandée ($Q_o > Q_d$) ? Niveau de prix, écarts de quantités, sens de l'ajustement ?

Lorsque la quantité offerte est supérieure à la quantité demandée, il y a **abondance**, parce que le niveau des prix est élevé et supérieur au niveau de prix d'équilibre. À ce prix élevé et supérieur au prix d'équilibre, les offreurs sont incités à produire davantage et les demandeurs sont incités à réduire la quantité demandée.

Dans ces conditions, le niveau de prix va baisser, afin de permettre son retour au niveau d'équilibre où la quantité demandée est égale à la quantité offerte.

– **Cas n° 2.** Que se passe-t-il, sur un marché quelconque, lorsque la quantité demandée est supérieure à la quantité offerte ($Q_d > Q_o$) ? Niveau de prix, écarts de quantités, sens de l'ajustement ? Lorsque la quantité demandée est supérieure à la quantité offerte, il y a **rareté**, parce que le niveau des prix est bas et inférieur au niveau de prix

d'équilibre. À ce prix élevé et supérieur au prix d'équilibre, les offreurs sont incités à produire moins et les demandeurs sont incités à augmenter la quantité demandée.



GRAPHIQUE 7.5.

Du déséquilibre à l'équilibre : le tâtonnement et le secrétaire de marché

Dans ces conditions, le niveau de prix va augmenter, afin de permettre son retour au niveau d'équilibre où la quantité demandée est égale à la quantité offerte.

On parle de **tâtonnement**, dans la mesure où le retour à l'équilibre se fait progressivement, car les agents économiques ne connaissent pas le prix d'équilibre.

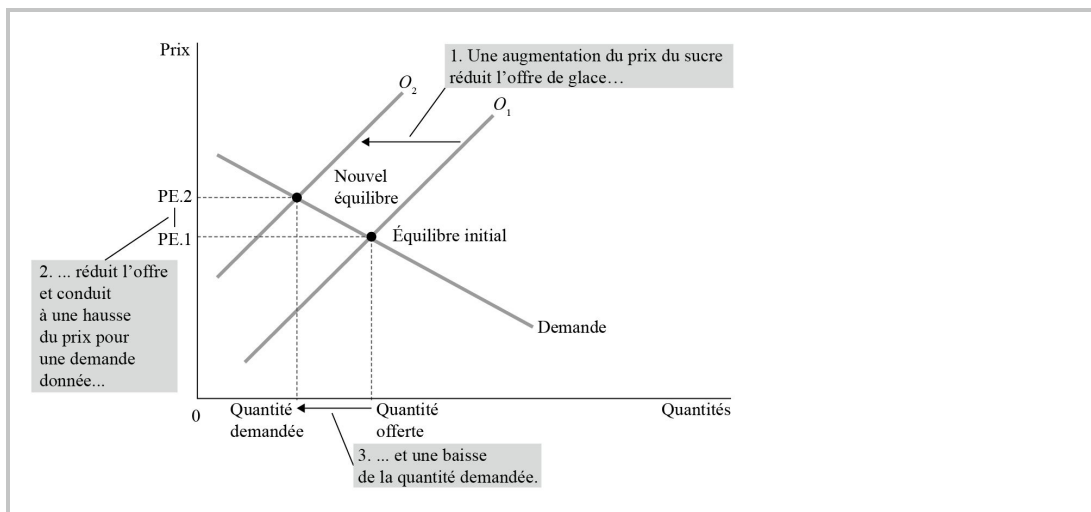
DE LA NÉCESSITÉ D'UN « COMMISSAIRE-PRISEUR »

Un « secrétaire de marché » est une institution nécessaire à la réalisation de l'ajustement. Les théoriciens de l'économie de marché ont expliqué que pour qu'un marché concurrentiel fonctionne parfaitement, il aurait besoin d'un « commissaire-priseur », un « secrétaire de marché », c'est-à-dire un agent qui annonce les prix aux potentiels offreurs-vendeurs et demandeurs-acheteurs, un peu comme dans une vente aux enchères.

7.4 DÉPLACEMENTS DE L'OFFRE ET DE LA DEMANDE : CHANGEMENTS D'ÉQUILIBRE

Une fois que l'équilibre est atteint, cela ne signifie pas qu'il restera inchangé. Les marchés évoluent, les situations ne

sont pas immobiles. Divers facteurs peuvent venir affecter l'offre et la demande simultanément. Un nouvel équilibre devra se former, modifiant ainsi le prix et la quantité échangée (quantité offerte/quantité demandée). Une modification de l'offre ou de la demande se traduit graphiquement par un déplacement de la courbe concernée (courbe d'offre/courbe de la demande). La loi de l'offre et de la demande désigne le mécanisme d'ajustement (confrontation et équilibrage par tâtonnement) des quantités offertes et demandées en fonction des variations de prix, uniquement.



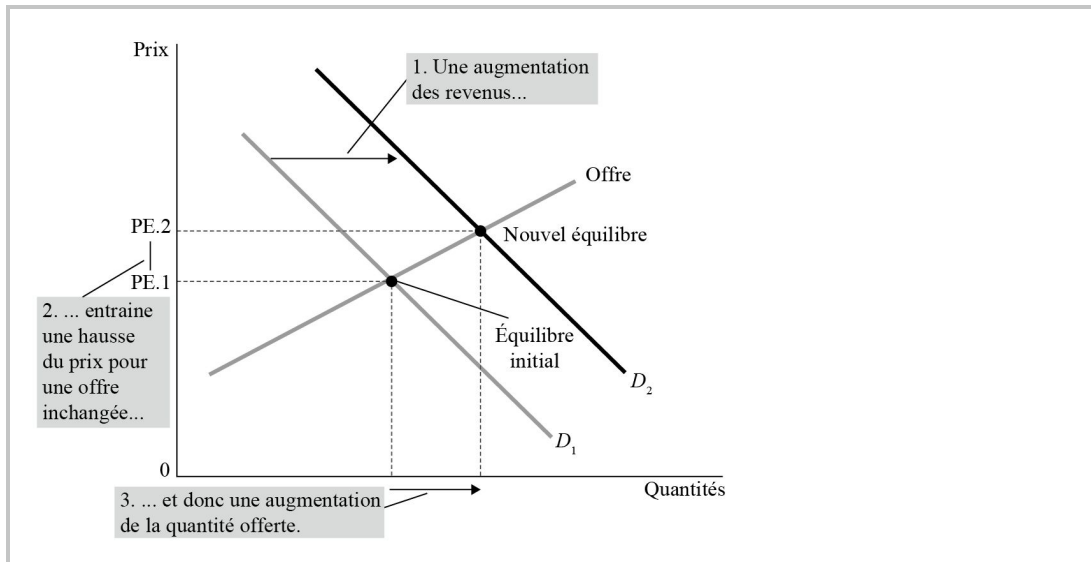
GRAPHIQUE 7.6.

Baisse de l'offre et changement d'équilibre

Généralement, un accroissement de l'offre diminue le prix et augmente la quantité offerte (et augmente la quantité demandée), car la courbe d'offre se déplace, la courbe de demande étant inchangée. *A contrario*, une diminution de l'offre augmente le prix et diminue la quantité demandée (et diminue la quantité offerte), car la courbe d'offre se déplace, la courbe de demande étant inchangée.

Généralement, un accroissement de la demande augmente le prix et augmente la quantité demandée (et augmente la quantité offerte), car la courbe de demande se déplace, la courbe d'offre étant inchangée. *A contrario*, une diminution de

la demande diminue le prix et diminue la quantité demandée (et diminue la quantité offerte), car la courbe de demande se déplace, la courbe d'offre étant inchangée.



GRAPHIQUE 7.7.

Hausse de la demande et changement d'équilibre

7.5 LES CONDITIONS D'EFFICACITÉ DES MÉCANISMES DE L'OFFRE ET DE LA DEMANDE

7.5.1 *La concurrence et les marchés*

La concurrence désigne une situation de compétition entre entreprises qui va les conduire à chercher à conquérir des parts de marché au détriment des concurrents effectifs et à rendre plus difficile l'entrée de concurrents potentiels.

Afin de mieux comprendre les conditions de fonctionnement des différents marchés, la théorie économique nous propose le modèle de marché de concurrence pure et parfaite (CPP). À l'aune de celui-ci, tous les autres modèles de marché de concurrence imparfaite peuvent être définis. Il est donc inutile de dire que ce modèle n'est pas réaliste. Par définition, il faut rappeler qu'aucun modèle n'est réaliste. Ce modèle est en

revanche fort utile pour appréhender de manière rigoureuse le fonctionnement des marchés réels imparfaits. Le mécanisme de l'offre et de la demande tel que nous l'avons présenté précédemment ne peut fonctionner de manière optimale qu'en situation de CPP. Dans toutes les autres situations de concurrence imparfaite, ce modèle d'offre et de demande fonctionne moins bien, voire pas du tout. Pour qu'un marché fonctionne selon les mécanismes de l'offre et de la demande présentés dans le chapitre précédent, il faut qu'il soit concurrentiel, c'est-à-dire conforme aux principes de la CPP (concurrence pure et parfaite). Moins il le sera, moins les mécanismes du marché (ceux de l'O et de la D) fonctionneront.

Pour que les mécanismes du marché fonctionnent parfaitement, il est nécessaire que les cinq conditions de CPP soient réunies simultanément. Dans tous les autres cas, dès qu'une seule des conditions n'est pas remplie, les mécanismes du marché fonctionnent imparfaitement ou pas du tout. Le modèle de concurrence pure et parfaite est une référence à l'aune de laquelle on va pouvoir présenter et expliquer toutes les situations de marché concrètes, réelles. Il s'agit donc d'un modèle fondamental pour comprendre l'économie de marché. Cependant, comme tout modèle, il est bon d'en connaître les fondements et les conditions de validité pour éviter de dire n'importe quoi et de tirer des conclusions erronées.

7.5.2 Les conditions de concurrence pure et parfaite (CPP)

Ce sont les économistes néoclassiques qui ont proposé ce modèle. Les pères fondateurs de cette analyse microéconomique sont le Français Léon Walras (1834-1910), l'Anglais Stanley Jevons (1835-1882) et l'Autrichien Carl Menger (1840-1921). Dans leur modèle, les consommateurs et producteurs sont supposés rationnels et ils cherchent à maximiser leurs intérêts propres. Pour des agents

économiques supposés parfaitement rationnels, les conditions suivantes doivent être respectées simultanément. En 1954, Kenneth Arrow et Gérard Debreu ont publié un article dans lequel ils ont finalisé le modèle de marché de concurrence pure et parfaite.

– **Atomicité du marché :**

Un marché « atomistique » est un marché sur lequel les offreurs et demandeurs n'influencent pas le prix. Il y a atomicité quand il y a un très grand nombre d'offreurs et de demandeurs qui sont tous très petits par rapport au marché global, et qui ne peuvent donc pas influencer le niveau du prix du marché, ils sont price takers ou preneurs de prix. Sur un marché parfaitement concurrentiel, aucun demandeur ou offreur isolé n'a le pouvoir de faire changer le prix. De la même manière qu'une goutte d'eau de l'océan ne peut pas changer le sens du courant... Le poids relatif de chaque offreur/demandeur dans l'ensemble de l'offre/demande agrégée est négligeable.

– **Homogénéité du marché :**

Il y a homogénéité sur un marché lorsque les produits échangés sont identiques et donc substituables. Seul le prix peut les distinguer ! Il n'y a aucune différenciation par la qualité, la présentation, le design, l'emballage....

– **Libre entrée sur le marché :**

Il y a libre entrée sur un marché lorsque n'importe quel concurrent potentiel peut entrer facilement. Les coûts d'entrée et de sortie ne sont pas des barrières infranchissables, il n'y a pas d'obstacles. Quiconque souhaite créer une entreprise sur un marché peut le faire sans être confronté à des coûts prohibitifs, ou à des délais. Les firmes déjà présentes ne peuvent s'opposer à l'arrivée de nouveaux entrants en instaurant des barrières tarifaires, administratives, techniques.

– **Transparence de l'information :**

Sur un marché, il y a transparence de l'information lorsque tous les offreurs et tous les demandeurs sont parfaitement

informés sur les conditions du marché (prix, homogénéité des produits...) et sont capables de traiter l'information parfaitement. La transparence de l'information suppose la perfection de l'accès à l'information et la capacité à la traiter pour l'utiliser et prendre des décisions présentes qui peuvent engager le futur.

La complétude de l'information est une situation plus contraignante que l'information parfaite, situation dans laquelle un agent connaît les valeurs passées, présentes et futures de différentes variables économiques. En économie, dans le cadre de la théorie des jeux, il y a complétude de l'information dans les situations où la connaissance du modèle (des fonctions d'utilité et des stratégies des autres agents) et de l'information passée, présente et future est commune à tous les agents.

– Mobilité des facteurs de production :

La mobilité des facteurs de production est complète quand les facteurs de production peuvent s'orienter vers d'autres marchés, d'autres activités sans freins divers. Les facteurs de production sont parfaitement mobiles et peuvent se déplacer d'une industrie à l'autre. La main-d'œuvre et le capital productif peuvent se diriger spontanément vers d'autres marchés sans délais ni coûts.

Le modèle de la CPP constitue une référence pour estimer le degré de concurrence/non-concurrence. Le mécanisme de l'offre et de la demande fonctionne très bien en situation de CPP, quand ces cinq conditions sont réunies simultanément. Dès l'instant où une seule condition n'est pas respectée, nous sommes en situation de concurrence imparfaite et le mécanisme de l'offre et de la demande fonctionne moins bien ou pas du tout.

Ce modèle, très abstrait, va servir d'étalon pour mieux estimer le degré de concurrence ou de manque de concurrence sur un marché. C'est seulement sur la base de ce constat que des politiques de concurrence adaptées pourront être définies et mises en œuvre.

De la même manière qu'on ne peut qualifier une vitesse d'excessive s'il n'y a pas de limite de vitesse indiquée clairement, on ne peut affirmer qu'il n'y a pas assez de concurrence sans référence précise pour en juger. À la lumière de cette dernière, il deviendra possible de dénoncer un manque de concurrence et de mener des politiques de concurrence adéquates.

7.5.3 Les marchés contestables : de la concurrence effective à la concurrence potentielle

Certains économistes, au début des années 1980 (Baumol, Panzar, Willig), ont approfondi la réflexion et ont montré que, dans certaines situations, la concurrence n'avait pas besoin d'être effective pour que ses bienfaits s'exercent. En effet, selon la théorie des « **marchés contestables** »

(**discutables**), il suffit que la concurrence soit potentielle pour que les concurrents en place se comportent comme si la concurrence était effective.

La concurrence potentielle est liée à la possibilité d'entrée de nouveaux concurrents sur le marché considéré, concurrents qui pourraient venir prendre des clients et réduire les parts de marché des entreprises présentes. La concurrence potentielle va obliger les concurrents effectifs à faire des efforts sur les prix et la qualité pour réduire les incitations des concurrents potentiels à entrer sur le marché. Ainsi, la simple possibilité d'entrée, sur le marché, de nouveaux concurrents oblige les producteurs en place à se comporter comme si les concurrents étaient déjà là. Selon cette logique, si un concurrent effectif propose des biens de qualité faible à des prix élevés, il risque de faciliter l'entrée de nouveaux concurrents sur le marché, car les perspectives de profits sont, dans un tel cas, plus élevées.

Cette théorie nous montre que même si l'atomicité n'est pas réalisée, il suffit que la libre entrée soit possible pour que les entreprises présentes ne se comportent pas comme si elles étaient à l'abri de toute concurrence. On comprend intuitivement le rôle majeur des autorités de la concurrence en matière de contrôle des barrières à l'entrée.

Par exemple, avant que Free n'obtienne sa licence de téléphonie mobile, les trois grands opérateurs en place (Orange, SFR, Bouygues) pensaient que Free ne l'aurait pas et ils n'ont donc pas baissé leurs prix ni vraiment amélioré la qualité des services proposés. Le marché des fournisseurs d'accès à Internet a évolué selon une logique assez proche.

En résumé, la théorie des marchés contestables nous explique donc que la concurrence d'un futur nouvel entrant potentiel peut agir sur les concurrents en place, comme la concurrence effective. Tant que la condition de libre entrée est respectée, la menace d'un nouveau concurrent suffit à obliger les concurrents déjà en place à faire des efforts de prix et de qualité (fidélisation des clients, promotions...).

En concurrence pure et parfaite, le niveau de profit tend logiquement vers zéro (règle d'épuisement du produit et profit nul). Seule la concurrence imparfaite permet l'existence durable de profits : « *Sur les marchés concurrentiels, la demande de facteurs de production est déterminée par leur productivité marginale. Dans le cas simplifié où les facteurs sont rémunérés en termes du seul produit, nous avons : salaire = productivité marginale du travail et rente = productivité marginale de la terre et ainsi de suite pour tous les facteurs. De cette manière, 100 % du produit est distribué, ni plus ni moins, entre les facteurs de production.* » Source : Economics de Samuelson et Nordhaus, fin du chapitre 12 sur la « détermination des revenus par les marchés », p. 222)

7.6 LE FONCTIONNEMENT DES MARCHÉS RÉELS : LA CONCURRENCE IMPARFAITE

Les **imperfections** de la concurrence sur les marchés désignent des situations où les mécanismes de l'offre et de la demande fonctionnent mal et aboutissent à des équilibres non optimaux. Des politiques de concurrence sont indispensables pour rétablir, renforcer et maintenir un degré suffisant de concurrence afin de rendre l'économie plus efficace.

Nous savons que le modèle de concurrence pure et parfaite (CPP) est une référence théorique qui nous permet de faire des comparaisons pour mieux comprendre les situations réelles de concurrence sur les différents marchés. Le modèle de CPP, aussi abstrait et aussi irréaliste soit-il, nous aide à mieux aborder le fonctionnement de tous les marchés réels sur lesquels nous échangeons chaque jour. En effet, on peut mieux appréhender le degré éventuel de défaut de concurrence. On peut mieux diagnostiquer les problèmes de fonctionnement des marchés. On peut mieux déterminer ce qui doit être fait afin que cela fonctionne mieux, pour que ceux qui échangent puissent le faire dans de meilleures conditions. En résumé, le modèle de concurrence pure et parfaite sert à mieux expliquer toutes les formes de marché en situation de concurrence imparfaite et indique la direction à suivre pour les politiques de concurrence.

7.6.1 *Les structures de marché en situation de CI en fonction de la condition d'atomicité*

Les économistes proposent une typologie des situations de concurrence imparfaite (CI). Nous pouvons commencer à distinguer les différentes catégories de marché de concurrence imparfaite en fonction de la condition d'atomicité. Heinrich Von Stackelberg (1905-1946) a ainsi mis en évidence neuf structures possibles de marché en fonction du nombre plus ou moins élevé d'offreurs ou de demandeurs. Cependant, **deux cas typiques et extrêmes** vont retenir davantage notre attention : le **monopole**, avec un offreur qui fait face à une multitude de demandeurs et le **monopsone**, avec un demandeur qui est confronté à une multitude d'offreurs. Ces deux types de structures de marché remettent en question l'atomicité et favorisent l'émergence d'un pouvoir de marché, des situations de « price makers » aux antipodes de la situation de « price taker », caractéristique fondamentale d'un marché en CPP.

Il existe **deux autres cas intermédiaires** et plus répandus dans l'économie réelle : l'**oligopole**, où quelques offreur font face à une multitude de demandeurs. Les quelques offreur pourraient se faire concurrence, mais ils ne le font pas, car une guerre des prix serait un jeu perdant-perdant. Ce choix leur confère un pouvoir partagé de marché. Symétriquement, l'**oligopsone** est un marché sur lequel un petit nombre de demandeurs sont confrontés à une multitude d'offreur. Dans ce cas, ce sont les demandeurs qui ont un pouvoir partagé de marché. Il est clair que la composition de l'offre et de la demande sur un marché dépend ici du nombre et de la taille relative des offreur et des demandeurs. Chaque structure de marché va influencer le degré de concurrence.

Définitions et exemples pour ces neuf types de structures de marché

Le **monopole** est une situation dans laquelle un seul offreur est en position d'exclusivité pour la production d'un bien ou service offert à un grand nombre d'acheteurs. C'est le cas de la Poste (uniquement pour les plis de moins de 30 grammes), d'EDF, du réseau ferroviaire... En Afrique du Sud, BEERS contrôle 80 % de la production de diamants et dispose donc

d'un grand pouvoir de marché. Sa position de « price maker » peut l'aider à éliminer ses petits concurrents.

Le **monopole bilatéral** est un marché sur lequel il y a un vendeur et un acheteur. C'est le cas, en France, de la construction des réacteurs Nucléaire par Areva pour EDF.

Le **monopole contrarié** est un marché sur lequel un offreur fait face à quelques demandeurs. C'était le cas de l'avion Concorde.

Le **monopsonne** est un marché sur lequel une infinité d'offres échangent avec un seul demandeur. C'est le cas du marché du tabac avant transformation en cigarettes, avec la Seita. C'est encore le cas dans le domaine des composants industriels, quand une usine spécialisée (voitures...) dispose d'un pouvoir de monopsonne face aux nombreux offres de composants industriels. Dans le domaine des fonctions régaliennes de l'État, celui-ci est en situation de monopsonne pour ses achats d'armes, mais aussi pour l'achat de matériel nécessaire à la sécurité intérieure et la justice. On peut aussi retenir l'exemple des commandes de l'État chinois face aux constructeurs aéronautiques mondiaux. EDF achète son électricité à Areva.

Le **monopsonne contrarié** est un marché caractérisé par quelques offres et un seul demandeur. En France, l'Armée achète ses armes à une seule entreprise publique Nexter. C'est aussi le cas de l'industrie de l'armement avec quelques offres (Dassault, Thalès...), l'État étant le seul acheteur.

L'**oligopole** désigne un marché où quelques offres font face une infinité de demandeurs. Les cas sont très nombreux. Il y a les marchés des sodas (aux États-Unis : Coca-Cola avec 50 % du marché, Pepsi-Cola avec 20 %...), des eaux minérales, des eaux pétillantes, des opérateurs de services de téléphonie mobile, des systèmes d'exploitation, de la distribution de carburant, des constructeurs automobiles, fabricants de smartphones (Apple, Samsung, Huawei...), les fabricants de Système d'exploitation : en 2016, Windows (79 %), OS X (17 %), Linux (4 %).

L'**oligopole bilatéral** correspond à un marché sur lequel quelques offreurs et quelques demandeurs échangent. Le marché de la haute couture en est un exemple typique. Le marché du luxe (haute couture et sur-mesure) au sens large également.

L'**oligopsone** désigne un marché sur lequel une infinité d'offeurs va échanger avec quelques demandeurs. L'exemple type est celui de la grande distribution, où de nombreux producteurs font face aux quelques centrales d'achats des grands distributeurs (Auchan, Carrefour, Leclerc...).

Le marché du travail où un nombre réduit d'employeurs (demandeurs de travail) échange avec un nombre élevé de travailleurs potentiels (offeurs de travail).

7.6.2 Les structures de marché en situation de CI en fonction de la condition d'homogénéité

La concurrence monopolistique : innovation et différenciation

La « concurrence monopolistique » (Edward Hastings Chamberlin) correspond à une situation de quasi-monopole et doit permettre d'augmenter le pouvoir de marché de l'entreprise en s'appuyant sur une stratégie de différenciation des produits. La concurrence est donc imparfaite, relativement à l'hypothèse d'homogénéité sur un marché, car la différenciation va affaiblir le signal prix, l'information donnée par le niveau de prix. Il en est ainsi sur de très nombreux marchés : automobile, ordinateur, télévision, smartphone, réfrigérateur... Il devient très difficile pour le demandeur de disposer d'une information suffisante et claire pour établir un rapport qualité/prix pertinent, afin de faciliter les arbitrages en réduisant les coûts d'opportunité.

7.6.3 Le monopole et la fixation des prix

En situation de monopole (ou de monopsone), le pouvoir de marché est total. Le monopoleur fixe le prix qui lui convient,

car, pour les acheteurs, c'est « à prendre ou à laisser ». Les monopoleurs sont « price makers » et les demandeurs sont « price takers ». Une entreprise en situation de monopole a peut-être pu éliminer, grâce à diverses stratégies, ses concurrents ou a pu dresser des barrières à l'entrée de nouveaux concurrents.

Mais, en amont de ces stratégies, les économistes distinguent trois catégories de monopoles selon le type d'explication, voire de justification : monopole naturel, monopole légal et monopole d'innovation.

Les différents types de monopoles

Le **monopole naturel** est un monopole justifié par des conditions techniques de production qui sont spécifiques, en raison de la présence de coûts fixes élevés. Par exemple, dans les industries en réseau comme le transport ferroviaire, l'électricité, l'eau, l'investissement de départ est très élevé et dissuasif. Il s'agit d'une sorte de barrière naturelle à l'entrée. Du fait de l'existence de rendements d'échelles croissants, la présence de concurrents rendrait la production peu rentable.

Le **monopole légal** est fixé par la loi ou par un organisme réglementaire. Il permet de restreindre la concurrence afin d'atteindre des objectifs politiques comme la sécurité, l'aménagement du territoire, la gestion d'un bien stratégique, la solidarité. Les services publics comme la poste et la distribution de courrier en font partie. Tout le monde doit pouvoir avoir accès au service, dans l'intérêt de tous. Il s'agit donc d'une forme de monopole protégée par des règlements administratifs, dans l'intérêt de tous.

Le **monopole d'innovation** est une situation temporaire. Tout innovateur, en particulier si l'innovation est majeure, radicale, est en situation de monopole transitoire, car il dispose transitoirement d'une position d'exclusivité (souvent, un brevet viendra favoriser cette protection temporaire). Ainsi, Apple avec le premier smartphone, l'iPhone, a proposé une innovation majeure dans la téléphonie et il a fallu attendre quasiment deux ans avant l'arrivée de réels concurrents sur

ce marché. Cette situation de protection est nécessaire pour la persistance des innovations, elle fournit une incitation à innover.

Pour Philippe Aghion et Peter Howitt, les entreprises innovantes bénéficient d'un surplus de rentabilité en vertu de la réduction du degré de concurrence : c'est la théorie des rentes de monopole liées à l'innovation. Cela s'explique surtout par la protection temporaire octroyée par le dépôt d'un brevet. Ce droit d'exclusivité transitoire incite les entreprises à innover.

Cependant, bien qu'il y ait des raisons logiques à l'existence de ces situations de marché, les monopoles suscitent de nombreux débats entre les économistes, les producteurs et les consommateurs, car, en situation de monopole, une entreprise dispose d'un **pouvoir de marché** qui peut conduire à une réduction de la qualité de la production et à une hausse des prix de vente, afin de réduire les coûts de production tout en améliorant la rentabilité, sans trop de risques de perte de clientèle.

7.6.4 Entreprises et pouvoir de marché

Si la concurrence peut être un bienfait pour le consommateur, ou plus largement l'acheteur final, en se traduisant par un meilleur rapport qualité/prix, et si la concurrence peut inciter les entreprises à faire des efforts dans ce sens afin de devenir plus compétitives, cette concurrence peut être destructrice pour les entreprises. En effet, le concurrent est par définition celui qui va prendre des clients et qui verra se réduire comme une peau de chagrin son chiffre d'affaires et ses bénéfices. Un client ne peut être en même temps client de deux entreprises en concurrence pour le même produit. L'arbitrage aboutit inévitablement à un choix !

7.6.4.1 Les dangers d'une concurrence potentiellement destructrice

Conscientes de ces dangers, les entreprises peuvent mettre au point de nombreuses stratégies pour réduire les chances

d'entrée et de rester sur le marché de concurrents potentiels ou encore pour éliminer des concurrents déjà présents.

Logiquement, dans une compétition, éliminer les concurrents est un moyen d'accroître ses chances de gagner. Ainsi en est-il en économie entre entreprises en concurrence, c'est-à-dire en compétition pour capter le maximum de clients effectifs et potentiels par rapport aux concurrents.

7.6.4.2 Les pratiques anticoncurrentielles sont nombreuses

De manière générale, elles consistent à accroître le pouvoir de marché des entreprises, même si c'est en réduisant les bienfaits potentiels de la concurrence, parmi lesquels on trouve des prix plus faibles et une meilleure qualité pour les consommateurs. Elles sont pourtant interdites par le droit de la concurrence et sanctionnées, mais les gains potentiels sont si élevés que contourner la concurrence est un jeu qui peut rapporter plus qu'il ne peut coûter.

7.6.4.3 Les entreprises peuvent réduire le nombre de concurrents via des ententes et fusions

Les ententes illicites ou cartels permettent de réduire l'intensité de la concurrence.

Un oligopole n'est pas une nuisance économique en soi. Il ne pose pas automatiquement de problèmes de concurrence. La position dominante n'est pas le problème, c'est l'abus éventuel de la position dominante qui est nuisible. C'est donc le comportement de concurrents qui abusent de leur pouvoir de marché qui peut nuire aux demandeurs et à l'économie.

Sur un oligopole, qui est la situation de marché la plus répandue, les quelques offreurs vont vouloir et pouvoir s'entendre pour éviter de se livrer une concurrence destructrice. Quand les concurrents d'un oligopole s'entendent, ils forment un cartel. Souvent, derrière les oligopoles se cachent des cartels. Les ententes et cartels sont interdits par la loi, car si les membres de l'oligopole s'entendent, ils peuvent agir comme un monopole et *ipso*

facto le prix augmente et la qualité baisse. L'existence d'un oligopole peut donc se traduire par la constitution d'un cartel et des abus de position dominante néfastes pour l'acheteur final et l'économie dans son ensemble.

De surcroît, plus faible est le nombre d'entreprises sur un marché d'oligopole, plus les ententes sont faciles et rentables. Ainsi, le duopole dans la construction d'avions Airbus/Boeing était une situation de marché très favorable à une entente afin d'éviter une concurrence pouvant être mutuellement destructrice.

Il existe de nombreux autres exemples de cartels. Le marché mondial de l'automobile est un exemple majeur d'oligopole et d'ententes. Environ dix constructeurs influencent le marché mondial. Les membres de l'oligopole peuvent s'entendre tacitement, mais ils ne peuvent pas le faire légalement. Il existe de nombreux autres exemples d'oligopoles cachant des cartels : les eaux minérales, les eaux pétillantes, les lessives, les opérateurs de téléphonie mobile, en France, avant l'arrivée de Free...

Le cartel des constructeurs de camions est un autre exemple significatif : les constructeurs de camions se sont entendus pendant 14 ans dans le cadre d'un cartel pour acquérir un pouvoir collectif de marché et devenir « *price maker* ». Les entreprises se sont mis d'accord pour pratiquer des prix élevés pour les camions, ceux-ci étant répercutés sur les acheteurs de camion et donc *in fine* sur les consommateurs. De cette manière, ces entreprises ont pu amortir les coûts élevés de mise en conformité aux normes environnementales en matière d'émission de CO₂.

7.6.4.4 Les opérations de fusions-acquisitions (FA ou fuscaqs)

**contribuent aussi à faire baisser l'intensité de la concurrence,
car elles renforcent le degré de concentration
via les regroupements d'entreprises**

Les opérations de fusions-acquisitions (FA) aboutissent à moins de concurrence, car le nombre d'entreprises sur le marché diminue et celles qui restent augmentent en taille. Potentiellement, leur pouvoir de marché est plus grand et leur capacité à former un cartel efficace est également plus grande. Les opérations de FA se traduisent par une concentration plus grande, dans la mesure où des entreprises se sont regroupées et qu'*in fine*, celles qui restent sont moins nombreuses, mais plus grosses.

On a pu observer des mouvements contraires en termes de concentration sur le marché de la téléphonie mobile en France par exemple. Avec l'entrée de Free aux côtés d'Orange, SFR et Bouygues, la concentration a été réduite et la concurrence accrue. Aujourd'hui, un mouvement contraire semble en cours, car de nouvelles fusions sont envisagées, et donc une concentration accrue, même si à ce jour la fusion entre Orange et Bouygues semble reportée *sine die*.

COMMENT EXPLIQUER ET JUSTIFIER LES ENTENTES ?

La **théorie des jeux** permet d'expliquer pourquoi et comment les concurrents ont intérêt à s'entendre plutôt qu'à se faire concurrence. Cette théorie permet d'expliquer pourquoi la coopération est préférable à la logique de la main invisible. D'une certaine manière, la théorie des jeux nous propose un mécanisme qui a une logique contraire à celle de la main invisible de Smith. Ce n'est pas en recherchant rationnellement son intérêt égoïste qu'un agent économique participe à l'intérêt collectif, mais c'est plutôt en recherchant, de manière partiellement altruiste, l'intérêt collectif qu'on peut améliorer son intérêt propre *in fine*. La coopération peut être préférable à la concurrence sauvage.

7.6.4.5 Les stratégies de prix et les barrières techniques sont d'autres moyens d'éliminer les concurrents

Les membres d'un oligopole peuvent aussi s'entendre et agir de concert pour empêcher de nouveaux concurrents d'entrer sur le marché en s'appuyant sur diverses stratégies de prix : les **prix prédateurs (dumping)**, les **prix limites** (un prix limite est un prix fixé, par une société en situation de monopole, en dessous du prix que sa situation de monopole lui permettrait théoriquement de pratiquer) et les **prix dissuasifs liés à des capacités excédentaires**.

Dans le cadre d'une stratégie de **prix prédateur**, du *dumping*, une entreprise peut pratiquer transitoirement des prix de vente très bas (en dessous du coût de production) en vue de se protéger et/ou d'étendre sa part de marché au détriment des concurrents effectifs ou potentiels. Cette stratégie revient à réduire la concurrence dans un premier temps, afin de profiter ensuite de profits plus importants dans le cadre d'une position concurrentielle plus favorable.

Les comportements de prix prédateurs peuvent être considérés comme de la concurrence déloyale, car, pendant un temps, les entreprises peuvent vendre à perte (vendre à un prix plus faible que le coût unitaire) afin de provoquer la faillite de leurs concurrents sur le marché. L'entreprise peut vendre jusqu'au moment où le nouvel entrant est en faillite. Cette pratique désigne le dumping.

Notons que les monopoleurs ou oligopoleurs sont de très grosses entreprises qui peuvent, *de facto*, réaliser des économies d'échelle. Quand l'échelle de production augmente, le coût unitaire de production baisse. Or, si chaque unité coûte moins, le prix peut baisser plus facilement...et cette stratégie est facile à mettre en place.

Cette stratégie de prix démontre la pression concurrentielle exercée par les entreprises qui la pratiquent, alors même qu'elles ne détiennent qu'une part de marché limitée, et comment les entreprises installées sur le marché concerné sont amenées à mettre en œuvre des stratégies variées pour y faire face, qu'il s'agisse du transport aérien ou du hard discount. Les entreprises qui ne s'adaptent pas à cette pression concurrentielle perdent des parts de marché et risquent de voir leur rentabilité diminuer et les pertes se creuser ; sans réaction de leur part, et si le processus se poursuit, elles sont vouées à être éliminées.

En situation de **capacité excédentaire**, les **prix** peuvent être maintenus à un niveau artificiellement et temporairement bas. En accumulant plus d'équipements productifs que nécessaire compte tenu de la demande du moment, une entreprise

envoie un signal aux concurrents potentiels concernant leurs capacités respectives de production et la possibilité d'engager une compétition dure en matière de prix, en augmentant la production pour bénéficier d'économies d'échelle.

Les stratégies de prix des monopoles discriminants

Les monopoles discriminants sont des entreprises qui vont proposer un bien ou service homogène à des prix différents à des populations différenciées en fonction de multiples critères. La discrimination par les prix consiste à identifier des groupes différents de clients (selon l'âge, le niveau de revenu, la fonction professionnelle, le statut [chômeur, étudiant...], la période de l'année...) qui n'ont pas la même disposition à payer et à leur proposer un prix plus ou moins élevé pour le même produit.

Cette discrimination en matière de prix doit permettre de capter le maximum de clients et de maximiser les profits en jouant sur les écarts de sensibilité au niveau des prix (élasticités prix de la demande) afin de capter le surplus de chaque consommateur (disposition à payer).

Par exemple, pour un smartphone ou des abonnements mobiles, il peut exister des forfaits spécifiques pour les jeunes ou les professionnels ; il peut exister des prix différents pour les places de cinéma (tarif moins de 18 ans, tarif chômeur), pour les billets de train (moins de 25 ans, plus de 60 ans), d'avion...

Cette forme de discrimination par le prix en fonction de la clientèle à laquelle le produit est vendu peut être positive dans la mesure où certains consommateurs peuvent accéder à des produits qui leur seraient inaccessibles dans le cas contraire.

Les **barrières techniques** sont susceptibles de réduire la concurrence en réduisant les chances d'entrée sur le marché de potentiels concurrents. Ainsi, sur le marché de l'automobile, des entreprises peuvent dresser diverses barrières à l'entrée de potentiels concurrents pour réduire les pertes de clients, de chiffre d'affaire et de bénéfices.

Par exemple, dans le secteur de l'après-vente automobile, on peut distinguer un marché « primaire », de la vente de véhicules, très concurrentiel, et un marché « secondaire », celui de l'après-vente, fort peu concurrentiel, car les constructeurs mettent en place de nombreuses barrières. Ils peuvent empêcher de potentiels concurrents de devenir des concurrents effectifs afin de garder le contrôle de la réparation et de l'entretien des véhicules. Ils utilisent pour cela des barrières techniques liées à l'électronique, à la visibilité-accessibilité des pièces... en se protégeant encore par des brevets ou droits d'auteur. Ces droits de propriété intellectuelle interdisent la fabrication et la commercialisation sur le territoire des pièces concurrentes de celles vendues par le constructeur. L'ensemble de ces protections confèrent au constructeur un monopole de droit et le protège de pertes potentielles.

Au fond, ces exemples de marché nous montrent que les entreprises n'aiment pas vraiment la concurrence et cherchent à l'éviter par tous les moyens possibles. Cela peut se comprendre aisément, puisque la concurrence imparfaite est indispensable à l'existence du profit. En effet, en situation de CPP, les opportunités de profits disparaissent progressivement, car tant qu'il existera des possibilités de profits sur certains marchés, il existera des entreprises qui souhaiteront y entrer. Le profit ne peut durablement exister qu'en situation de concurrence imparfaite. On ne peut donc pas être étonné d'observer que les entreprises tentent de maintenir et d'accroître leur pouvoir de marché. C'est logique, mais cela peut se faire au détriment de la demande, de la dynamique économique globale et du progrès.

Compte tenu de ces comportements et stratégies, l'État et les autorités de la concurrence sont donc légitimes pour essayer de veiller à ce qu'il y ait un degré de concurrence minimum.

7.7 LES POLITIQUES DE LA CONCURRENCE

7.7.1 Les objectifs de la politique de la concurrence

Les politiques de concurrence visent à agir sur les conditions de l'offre sur le marché, sur les conditions de concurrence par les prix et par la qualité.

Étant donné l'existence de ces comportements anticoncurrentiels, de ces stratégies et pratiques interdites par le droit de la concurrence, la politique de la concurrence a pour objectif final de sanctionner afin de dissuader ce type d'ententes illicites et nuisibles. La politique de la concurrence est une forme de politique structurelle dont l'objectif est d'établir, de maintenir, de renforcer cette concurrence. Les entreprises devront dès lors investir et innover dans le cadre d'une concurrence effective aux effets potentiellement vertueux. Pour faire face à la concurrence, les entreprises ne doivent donc pas pouvoir s'appuyer sur des ententes et profiter de la position dominante qui en découle afin d'imposer des prix plus élevés aux acheteurs, tout en réduisant la qualité, dans le but d'obtenir une augmentation de leurs profits en tirant avantage d'une rente de situation.

En aval, les autorités de la concurrence sont donc indispensables pour contrôler ce type de pratiques et, les fautes étant constatées, la politique de la concurrence va se traduire par des sanctions pour ce type d'abus de position dominante.

Ceci étant, de quels moyens disposent concrètement les autorités de la concurrence pour établir, maintenir et renforcer le degré de concurrence nécessaire pour sauvegarder les bienfaits de la concurrence ? De manière générale, les autorités chargées de veiller au respect de la concurrence interviennent en amont par le contrôle et la réglementation puis, en aval, via des sanctions.

7.7.2 Les politiques de la concurrence : réglementations et contrôles en amont

Diverses **réglementations et interdictions** ont pour finalité de favoriser et à maintenir la concurrence entre les producteurs. Elle agit principalement sur les pratiques des entreprises. En effet, celles-ci peuvent nuire à la concurrence, comme nous l'avons vu ci-dessus. Par exemple, les réglementations relatives aux **brevets et droits d'auteur** permettent de réguler la concurrence. Il s'agit de protéger temporairement les innovateurs de la concurrence pour maintenir le caractère incitatif de ces opérations plus ou moins risquées. L'objectif est aussi d'éviter la concurrence déloyale, la copie, l'imitation.

Contrôle de la concentration, des fusions

La politique de la concurrence permet de lutter en amont, c'est-à-dire avant que les fautes soient commises, contre les abus potentiels et probables de position dominante liés à des concentrations excessives sur différents marchés via des fusions-acquisitions. La politique de la concurrence peut conduire à autoriser ou interdire des fusions selon les cas. Cela dépendra des effets attendus de la concentration sur le degré de concurrence et ses conséquences. Les opérations de fusions-acquisitions (FA) aboutissent à moins de concurrence, car le nombre d'entreprises sur le marché diminue et celles qui restent augmentent en taille. Potentiellement, leur pouvoir de marché est plus grand et leur capacité à former un cartel (entente) efficace est également plus forte.

Réglementations et interdictions pour réduire les barrières à l'entrée et favoriser la concurrence

Toujours en amont, par le biais des diverses réglementations et interdictions, la politique de la concurrence peut agir principalement sur les pratiques des entreprises. En effet, celles-ci peuvent nuire à la concurrence en élaborant diverses stratégies pour construire des barrières à l'entrée.

Nous avons vu plus haut le cas du marché de l'automobile, qui permet à des entreprises de dresser diverses barrières à l'entrée de potentiels concurrents pour réduire les pertes de

clients, de chiffre d'affaire et de bénéfices (marchés primaire et secondaire).

L'intervention des autorités de la concurrence se révèle logiquement indispensable dans ce type de cas.

7.7.3 Les politiques de la concurrence : sanctions en aval

La politique de la concurrence permet aussi de sanctionner les abus de position dominante constatés. Ils peuvent être liés à des situations de monopole et à la constitution de cartels, qui désignent une entente entre plusieurs firmes rivales, sur un marché d'oligopole, afin de limiter la concurrence et de maintenir des prix élevés pour une qualité limitée.

Les politiques de la concurrence face aux monopoles et aux abus de position dominante

De fait, une entreprise en situation de monopole dispose d'un pouvoir de marché maximum si elle n'est pas indirectement menacée par l'arrivée de nouveaux entrants potentiels (marché de monopole non contestable). La théorie des marchés contestable proposée au début des années 1980 a permis de comprendre qu'il suffisait que la concurrence soit potentielle pour que les entreprises présentes sur un marché se comportent comme si cette concurrence était effective.

Si elle n'a pas à craindre la concurrence, l'entreprise en situation de monopole peut profiter de sa position pour imposer des prix plus élevés aux consommateurs, tout en réduisant la qualité et augmenter ainsi ses profits en tirant avantage de sa rente de situation.

En revanche, si elle se sent potentiellement menacée par l'entrée d'un concurrent, elle réduira ses prix et ses bénéfices pour désinciter et rendre plus difficile l'entrée de concurrents. En aval, le régulateur est donc indispensable pour contrôler ce type de pratiques et, les fautes étant constatées, la politique de la concurrence va se traduire par des sanctions pour ce type d'abus de position dominante.

Les politiques de la concurrence face aux ententes, aux abus de position dominante et nuisances des cartels

La régulation de la concurrence passe donc par des sanctions en aval, c'est-à-dire une fois les pratiques interdites constatées. Ce qui est très difficile, car les ententes sont tacites. Face à ces ententes illégales et avérées, les autorités de la concurrence peuvent imposer de lourdes amendes pour non seulement sanctionner les tricheurs, mais aussi pour dissuader ceux qui pourraient être tentés.

Chaque année, le Conseil de la concurrence publie un rapport sur ce thème. En 2004, Noblet (importateur exclusif de Casio) et Texas dominant le marché des calculatrices (89 %) et se sont entendus pour supprimer en même temps leurs modèles les moins chers, et augmenter les prix des autres. Puis ils se sont mis d'accord avec les distributeurs Carrefour, Majuscule et Distributeurs associés pour que tous pratiquent le même prix exorbitant. En punition de cette entente et de cet abus de position dominante, l'amende s'est élevée à 3,8 millions d'euros.

En 2006, les trois opérateurs de téléphonie mobile Orange, SFR et Bouygues Telecom ont été condamnés pour entente illicite à verser une amende record de 534 millions d'euros (256 millions d'euros pour Orange, 220 pour SFR et 58 pour Bouygues Telecom). Il leur était reproché de s'être entendus, entre 2000 et 2002, pour geler leurs parts de marché en maintenant leurs prix à un niveau artificiellement élevé. Ils auraient également échangé, entre 1997 et 2003, des informations confidentielles et stratégiques sur le nombre de nouveaux abonnements et de résiliations.

En 2016, dans le cas du cartel des constructeurs de camions, le moyen de sanctionner les membres d'un cartel pour entente illicite fut de leur faire payer une amende élevée. L'autorité de la concurrence compétente, la Commission européenne en l'occurrence, a infligé une amende record d'un montant de 2 926 499 000 d'euros aux constructeurs de

camions concernés par cette entente sur une période de 14 ans.

En avril 2015, la Commission européenne a officiellement accusé le moteur de recherche Google d'abus de position dominante et de violation des règles de la concurrence. Bruxelles reproche à l'entreprise d'avoir injustement avantage son propre service de comparaison de prix, en violation des règles de l'Union européenne en matière d'ententes et d'abus de position dominante (Google disposant déjà de 90 % des parts de marché). Dans l'attente d'une décision de l'Europe, il est possible que la firme Google soit condamnée à payer une amende qui pourrait être équivalente à 10 % de son chiffre d'affaires mondial, soit 6 milliards de dollars. Le géant américain des moteurs de recherche a été condamné en juin 2017 à régler une amende de 2,43 milliards d'euros. Puis, en juillet 2018, Google a été condamné à une amende historique de 4,34 milliards d'euros par la Commission européenne pour abus de position dominante via Android, son système d'exploitation pour smartphones, lui permettant de favoriser ses propres applications et son moteur de recherche.

Sans prétendre à l'exhaustivité, nous avons pu montrer comment, et même pourquoi, la politique de la concurrence intervient en amont et en aval pour établir, maintenir, voire renforcer la concurrence pour protéger l'utilisateur final sans pour autant nuire aux entreprises. Cependant, il convient de mentionner que la politique de la concurrence s'applique aussi aux interventions étatiques nationales qui peuvent biaiser la concurrence par des subventions et autres aides aux entreprises. L'Union européenne lutte contre les monopoles publics et les aides publiques qui instaurent des conditions de concurrence déloyale.

En situation de concurrence imparfaite, les mécanismes du marché fonctionnent, mais on observe des distorsions, des équilibres non optimaux dès qu'une seule des conditions de la CPP n'est pas respectée.

En revanche, dans le cas de défaillances des marchés, les mécanismes de l'offre et de la demande, l'ajustement des quantités et des prix, ne fonctionnent plus du tout.

7.8 QUELLES SONT LES PRINCIPALES SITUATIONS DE DÉFAILLANCES DU MARCHÉ ?

Les **défaillances** de marché désignent des situations où les mécanismes de l'offre et de la demande ne permettent pas une allocation efficace des ressources. Dans ces conditions, la main visible des pouvoirs publics doit compenser l'absence de main invisible du marché pour améliorer l'allocation des ressources (rares) et favoriser la convergence des intérêts individuels et collectifs.

On peut distinguer trois grandes catégories de défaillances : les asymétries d'information, l'existence d'externalités et l'existence de biens collectifs.

7.8.1 *Les asymétries d'information*

Quotidiennement, les agents économiques qui prennent des décisions, qui font des choix, se trouvent le plus souvent en situation d'asymétrie d'information, au sens où ils ne disposent pas tous des mêmes informations ni de la même capacité à traiter ces informations pour prendre leurs décisions. Ces asymétries peuvent induire des comportements non vertueux, des comportements où l'intérêt des uns ne rejoint pas celui des autres. En situation d'asymétrie d'information, les mécanismes de marché sont défaillants, c'est-à-dire qu'ils ne permettent pas aux agents économiques de faire le meilleur choix. Aussi, une hypothèse de CPP n'est plus respectée, celle de la perfection de l'information, c'est-à-dire que tous les agents économiques n'ont pas un égal accès à l'information et qu'ils n'ont pas tous la même capacité d'analyse de l'information.

L'asymétrie d'information est une situation dans laquelle les offreurs ou les demandeurs ne disposent pas de la même information sur le marché. Le vendeur ou le demandeur a une meilleure information sur le marché et, donc, il peut l'utiliser à son profit.

Les asymétries d'information sont des situations au cours desquelles certains agents sont plus ou moins bien informés que d'autres. Ils peuvent donc avoir accès à l'information sur des produits échangés. Ils peuvent avoir des problèmes à traiter l'information afin de faire leurs choix.

Par exemple :

- Sur le marché des voitures d'occasion, le vendeur dispose de plus d'informations sur le véhicule qu'il souhaite vendre que l'acheteur, qui cherche à obtenir le maximum d'information afin d'arrêter son choix. Il y a bien une asymétrie d'information en faveur de l'offreur.
- Sur le marché des crédits, il y a les demandeurs de crédit (ceux qui veulent emprunter) et les offreurs de crédit (ceux qui prêtent, comme les banques). Les demandeurs ont davantage d'information sur leur solvabilité que les offreurs.

Il y a bien une asymétrie d'information en faveur des demandeurs.

L'existence d'asymétries d'information, et donc d'imperfection de l'information, va se traduire par deux types de comportements de la part des agents les moins informés :

- des comportements d'anti-sélection ou sélection adverse ;
- des comportements d'aléa moral.

L'existence de ces comportements menace l'existence même des marchés concernés, car les échanges peuvent s'arrêter. Or, s'il n'y a plus d'échanges, il n'y a plus de marché !

7.8.1.1 L'anti-sélection ou sélection adverse

Il y a anti-sélection quand les agents économiques font collectivement des choix qui vont conduire à la défaillance du marché, au sens où leurs décisions agrégées vont se traduire par des effets négatifs sur ceux qui souhaitent échanger.

G. Akerlof a mis en évidence, dans un célèbre article publié en 1970, « The Market for Lemons », les conséquences des comportements d'anti-sélection en prenant l'exemple des marchés de voitures d'occasion. Sur ce marché, Laurent le vendeur dispose de beaucoup plus d'informations sur les voitures qu'il souhaite vendre que David l'acheteur. David, n'étant pas né de la dernière pluie, n'est pas naïf et se doute bien que Laurent peut vouloir masquer les défauts majeurs des voitures. David sait que chez Laurent Autodoccase, les vices cachés sont légion. Dans ces conditions, David et les acheteurs vont exiger des baisses de prix parce qu'ils manquent de confiance. Du coup, les vendeurs honnêtes de voitures d'occasion qui passent par le garage Laurent Autodoccase vont se retirer du marché. Si ce comportement est individuellement rationnel, il devient destructeur collectivement. En effet, peu à peu, le marché des voitures d'occasion va devenir un marché de tacots.

C'est le comportement individuel rationnel des acheteurs, qui doutent à juste titre, et du vendeur, qui rationnellement tente de profiter de l'asymétrie d'information, qui va conduire à un marché de voitures de faible qualité, voire à la disparition du marché.

On pourrait montrer, de la même manière, que les comportements de sélection adverse situés du côté du demandeur, sur le marché du crédit bancaire, ont toutes les chances de conduire à un marché sur lequel seuls les emprunteurs de mauvaise qualité restent et acceptent de payer des taux d'intérêt élevés, du fait du haut niveau des primes de risque. En effet, les banques qui prêtent vont vouloir se protéger des emprunteurs risqués en imposant des taux d'intérêt élevés. Du fait de ces taux d'intérêt élevés, les emprunteurs peu ou pas risqués vont se retirer du marché du crédit. En situation d'asymétrie d'information, le prix ne reflète

plus la qualité. Il envoie un mauvais signal. La sélection adverse se traduit par la disparition des produits de haute qualité sur le marché, ce qui, *in fine*, peut faire disparaître le marché.

7.8.1.2 L'aléa moral

L'aléa moral désigne une situation dans laquelle une personne prend davantage de risques, car elle sait qu'elle sera sauvée.

Prenons quelques exemples simples, en dehors de l'économie : « *Les pompiers ne sont pas contents. En effet, ceux qui ont pris des risques pour secourir récemment un spéléologue, coincé dans une grotte dans la Drôme, estiment que le contribuable n'a pas à payer pour sa négligence. Et pour la première fois, un service départemental d'incendie et de secours a déposé plainte contre le spéléologue secouru, lui reprochant d'avoir "mis en danger" les sauveteurs. Dans le cas présent, le sauvetage d'un spéléologue expérimenté a nécessité l'intervention de dix-sept sapeurs-pompiers et 55 sauveteurs, dont quatre plongeurs. Le spéléologue, en situation d'aléa moral, est incité à prendre des risques inconsidérés puisqu'il ne sera pas le seul à être sanctionné par son inconscience et qu'il a même toutes les chances d'être sauvé. Il peut s'attendre à ce que les pompiers viennent à son secours. De plus, il ne payera pas la facture* » (« Sauvetage de spéléologues, économie et aléa moral », David Mourey, Pontault-Combault, 2007).

Le spéléologue prendra davantage le risque de finir coincé dans une grotte, car il sait qu'il a toutes les chances d'être sauvé. Il mettra donc en danger la vie des sauveteurs. Mais il ne s'en inquiètera pas vraiment si, de surcroît, il n'a rien à payer *in fine*. Logiquement, si les spéléologues savaient qu'en cas de sauvetage, ils devraient payer une amende pour mise en danger de la vie d'autrui, ils réfléchiraient autrement. Il en est de même pour les baigneurs sur les plages de l'océan Atlantique en été qui se baignent en dehors des zones surveillées, car ils savent que les MNS ou les CRS

viendront les sauver ; mais aussi des alpinistes qui prennent le risque d'aller en montagne malgré les mauvaises prévisions météorologiques et mettent en danger les secouristes.

Par ailleurs, sur le marché de l'assurance, on peut aussi observer des comportements d'aléa moral de la part des assurés. En effet, un assureur auto doit assurer une personne, mais il ne sait pas si l'assuré sera complètement responsable au volant, du fait qu'il est assuré.

On peut encore penser aux grosses banques (« too big to fail », ou banques dites « systémiques ») qui peuvent prêter sans vérifier suffisamment la solvabilité de leurs clients, ou qui peuvent faire des placements risqués, car elles savent qu'en cas de risque avéré, elles ont toutes les chances d'être sauvées par la banque centrale, qui voudra éviter un risque systémique. La crise financière internationale de 2007 nous donne une illustration du problème de l'aléa moral. Pendant des années, de grands groupes bancaires ont pris d'énormes risques sur les marchés financiers. Or, ces banques sont très dépendantes les unes des autres. Si l'une chute, elle peut entraîner en cascade la chute des autres dans le cadre de ce qu'on appelle un « effet domino ». Une banque peut acquérir la certitude que les pouvoirs publics lui viendront en aide et elle multipliera les prises de risques. Lutter contre l'aléa moral est un exercice difficile, car il s'agit de ne pas sauver ceux qui ont pris trop de risques. En septembre 2008, l'État américain et les banques ont décidé que la banque d'affaires Lehmann Brothers ferait faillite pour faire un exemple et sanctionner l'aléa moral... ce qui a entraîné une grave crise boursière et financière. Par ailleurs, sur le marché du crédit, on sait qu'il y a une asymétrie d'information qui peut entraîner la sélection adverse, ce qui risque de faire disparaître le marché. En effet c'est la banque qui a le moins d'informations, et elle va alors augmenter les taux d'intérêt. Les agents économiques qui empruntent ne vont plus demander de crédit, et seuls les entrepreneurs dont les projets sont les plus risqués accepteront ce taux d'intérêt.

7.8.2 Les externalités

Les externalités ou effets externes désignent les effets des activités économiques et des échanges des uns sur d'autres agents économiques, qui ne participent pas à ces échanges. On parle d'externalités lorsque les actions d'un agent économique ont un impact positif ou négatif sur le bien-être et le comportement d'autres agents, et que cet impact n'est pas pris en compte dans les calculs de l'agent qui le génère. En toute logique, sur un marché de concurrence pure et parfaite, un échange ne devrait avoir aucun effet sur ceux qui n'y participent pas. Dès que ce n'est pas le cas et qu'il y a des effets sur d'autres personnes, il y a des externalités et donc une défaillance du marché.

On doit distinguer deux types d'externalités¹³ :

Les **externalités positives** désignent les situations où un agent économique est favorisé par les opérations économiques d'un ou plusieurs autres, sans qu'il ait à payer quoi que ce soit. Imaginons par exemple une entreprise qui va former un travailleur pour qu'il acquière de meilleures compétences ; on peut parler d'effets externes positifs dans la mesure où, lors du départ de l'entreprise de ce travailleur, il emportera avec lui son capital humain amélioré et en fera bénéficier plus ou moins une autre entreprise, qui n'en aura pas supporté les coûts. L'implantation d'une activité au voisinage d'une autre qui bénéficie des effets de synergie ou autres effets induits par cette nouvelle proximité est aussi une forme d'externalité positive.

Les **externalités négatives** désignent les situations où un agent économique est défavorisé par les opérations économiques d'un ou plusieurs autres, sans qu'il en soit dédommagé par ceux-ci. Une entreprise qui produit un bien, mais qui pollue et dégrade l'environnement naturel lors de son processus de production, va ainsi échanger ses produits avec des acheteurs. Cependant, les nuisances occasionnées à autrui à cause de la pollution n'entreront pas dans sa structure de coût et les personnes victimes de ses nuisances

devront supporter seules les réparations et protections liées à des échanges auxquels elles n'ont pas participé. Un aéroport provoque des nuisances sonores et olfactives et détruit la qualité de l'air pour les habitants à proximité. Ceux qui ne prennent jamais l'avion devront pourtant supporter les coûts d'installation de doubles vitrages, de soins liés à la dégradation de leur santé physique et morale... On peut encore penser aux coûts pour la société et la collectivité du travail dangereux de certaines personnes, sans que ceux qui bénéficient des biens produits versent une compensation (comprise dans le prix) à ceux qui participent à ces productions.

En conclusion, les auteurs des externalités négatives n'en subissent pas le coût économique, car le pollueur (offreur ou demandeur) ne paye pas les dégâts liés à sa production. Plus encore, c'est la collectivité qui a toutes les chances de devoir en supporter les coûts.

7.8.3 *Les biens collectifs*

Dans le cadre d'une typologie limitée, les économistes nous proposent de distinguer quatre catégories de biens en fonction de deux critères, la rivalité et l'excluabilité. En résumé, cela donne ceci :

	Exclusion	Non-Exclusion
Rivalité	Biens privés Vêtements, pain, carburant, smartphone, le présent ouvrage, un ordinateur portable...	Biens communs Les ressources halieutiques (poissons), le gibier pour les chasseurs, les pâturages (l'herbe mangée par un troupeau)...
Non-Rivalité	Biens de club Film au cinéma, autoroutes, programmes télévisés payants : Canal+, CanalSat, beIN Sports, Netflix, Prime Video, Spotify, Deezer...	Biens collectifs Éclairages publics, Jardin du Luxembourg, Défense nationale, Sécurité intérieure, Éducation nationale...

Les **biens privatifs** sont ceux qui s'échangent de manière marchande sur les différents marchés. Les **biens collectifs** ont les caractéristiques contraires et ne peuvent donner lieu à un échange marchand.

Enfin, il existe deux catégories intermédiaires : les **biens de club**, qui peuvent avoir un prix de marché, mais sont non rivaux et les **biens communs**, qui sont rivaux, mais ne peuvent avoir de prix.

Un bien est qualifié de **rival** dans la mesure où la consommation par l'un empêche complètement celle de l'autre. Si David s'achète une crêpe, Laurent ne pourra pas en profiter, sauf s'il la vole...

Un bien est dit **excluable** quand son accès peut être limité en vertu de l'exclusion par le prix.

Celui qui paye le prix peut satisfaire son besoin, celui qui ne paye pas ne le peut pas.

Un **bien collectif** respecte la non-rivalité (la consommation de ce bien ne peut empêcher la consommation de ce même bien par un autre individu) et la non-exclusion (il est impossible d'exclure par les prix l'usage de ce bien). Pour ces biens, les mécanismes de marché sont complètement défaillants. On ne peut exclure par le prix, car tout le monde a intérêt à ce que tout le monde y ait accès et il n'y aucune rivalité.

Il ne faut pas confondre les biens collectifs et les biens publics :

- Un bien collectif n'est pas forcément un bien public et réciproquement.
- Un bien public existe sur la base d'un choix des pouvoirs publics, de l'État.
- Un bien collectif existe en raison d'une défaillance de marché.

Les agents privés peuvent tenter de résoudre les problèmes de production et d'accès à un bien collectif. C'est le cas des

fondations privées qui gèrent des espaces naturels ouverts à tous.

13 Il est possible de distinguer les externalités de consommation (exemple) des externalités de production (exemple) et, au sein de chaque catégorie, des externalités positives ou négatives.

LES SOURCES DE LA CROISSANCE ÉCONOMIQUE À LONG TERME

« La difficulté de déterminer les causes de la croissance tient avant tout à ce que la croissance se mesure, qu'elle est une quantité, mais les phénomènes qui la déterminent sont essentiellement qualitatifs. » - Raymond Aron (1905-1983)

« La croissance est associée d'abord à une idée biologique : on parle de la croissance d'une plante. Elle implique une augmentation de dimension, de volume, de poids, de taille. Un développement est un déroulement hors de soi. Une croissance est quantitative, mesurable, alors qu'un développement postule quelque idée de qualité, quelque chose qui échappe à la mesure. » - Henri Guitton, 1976

« Les sociétés modernes sont avides de croissance, davantage que de richesse. Mieux vaut vivre dans un pays pauvre qui s'enrichit (vite) que dans un pays (déjà) riche et qui stagne. Les Français ont follement apprécié les Trente Glorieuses, car tout était neuf. Mais au bout du compte, la page reste toujours blanche du bonheur à conquérir. » - Daniel Cohen, 2009

SOMMAIRE

[8.1 Croissance économique, développement et bien-être](#)

[8.2 Calculer et mesurer la croissance économique](#)

8.3 Les tendances de la croissance économique ou la croissance tendancielle

8.4 Les principales explications de la croissance économique à long terme

8.5 Les origines du progrès technique

8.6 Les théories de la croissance endogène et le rôle des institutions

8.1 CROISSANCE ÉCONOMIQUE, DÉVELOPPEMENT ET BIEN-ÊTRE

8.1.1 *La croissance économique*

Dans une économie, la croissance économique peut se définir comme une hausse soutenue et durable du PIB (cf. chapitre 4) destinés à satisfaire les besoins de tous les agents de l'économie. C'est un phénomène durable. Sur une courte période (moins de deux ans), on préfère utiliser le terme d'expansion (cf. chapitre 9). C'est un phénomène quantitatif, mais qui a également des effets qualitatifs.

De manière analogue, nous devons éviter de confondre les notions de récession et de dépression :

- La récession correspond à un ralentissement de la croissance. Autrement dit, il y a toujours croissance de la production globale.
- En revanche, en situation de dépression, le PIB baisse fortement et durablement. Le taux de croissance est négatif. Mais le PIB ne peut jamais être négatif, cela n'a aucun sens, contrairement à ce que l'on entend parfois de la part de commentateurs pressés.

Comme pour la production, nous verrons que ce qui compte, c'est la croissance du volume (Q) de production, bien davantage que la croissance de la valeur (pxQ) de la production. Selon la même logique, la distinction entre croissance effective et croissance potentielle restera cruciale, à court terme comme à long terme (cf. la distinction entre PIB effectif et PIB potentiel).

8.1.2 *La croissance n'est pas le développement.*

En effet, le PIB ne permet pas de mesurer ce que nous appelons le développement dans ses diverses acceptions.

Lato sensu, le développement désigne l'ensemble des transformations techniques, sociales, démographiques, structurelles, culturelles, politiques... accompagnant la croissance de la production. Le développement est donc une notion bien plus large qui permet de prendre en compte les aspects économiques, sociaux, techniques, démographiques, institutionnels, structurels, qualitatifs, culturels... Le développement peut être associé à l'idée de progrès économique et social, estimé par l'amélioration des niveaux de vie, d'instruction, et, *in fine*, de bien-être pour l'ensemble de la population. Il se manifeste par des changements dans les différentes dimensions qui accompagnent la croissance économique.

Selon François Perroux, économiste français (1903-1987), le développement désigne « *la combinaison des changements mentaux et sociaux d'une population qui la rendent apte à faire croître cumulativement et durablement son produit réel global* ».

Croissance économique et développement sont donc étroitement liés : si la croissance peut favoriser le développement, elle n'est pas forcément suffisante.

8.2 CALCULER ET MESURER LA CROISSANCE ÉCONOMIQUE

8.2.1 *De la croissance en valeur (nominale) à la croissance en volume (réelle)*

C'est le simple calcul du taux de variation du PIB qui permet de mesurer la croissance économique effective. Le taux de croissance correspond donc au taux de variation du PIB. La croissance est le phénomène, le taux de croissance est sa mesure. Pour mesurer la **croissance réelle**, en volume ou encore en monnaie constante, il suffit d'éliminer l'effet de l'augmentation des prix. On va donc « déflater » le PIB

nominal afin d'obtenir le PIB en volume. Ensuite, on peut calculer le taux de croissance du PIB réel pour obtenir la croissance en volume. Notons que pour obtenir le PIB réel et calculer le taux de croissance du PIB en volume, on peut utiliser l'IPC ou le déflateur du PIB.

IPC ET DÉFLATEUR DU PIB

L'indice des prix à la consommation (IPC) est l'instrument de mesure de l'inflation. Il permet d'estimer, entre deux périodes données, la variation moyenne des prix des produits consommés par les ménages. C'est une mesure synthétique de l'évolution de prix des produits, à qualité constante. L'indice des prix hors tabac sert à indexer de nombreux contrats privés, des pensions alimentaires, des rentes viagères et aussi à revaloriser le SMIC. L'indice retenu pour le SMIC est celui des « ménages du 1^{er} quintile de la distribution des niveaux de vie, hors tabac ». Depuis la diffusion de l'IPC de janvier 2016, l'Insee publie un nouvel indice en base 2015 = 100 en lieu et place de l'indice en base 1998 = 100. Cet indice rénové constitue la huitième génération de l'indice depuis 1914.

Il est essentiel de rappeler que l'indice des prix à la consommation n'est pas un indice du coût de la vie. En effet, l'indice des prix à la consommation cherche à mesurer les effets des variations de prix sur le coût d'achat des produits consommés par les ménages. L'indice du coût de la vie cherche à mesurer les variations des coûts d'achat pour maintenir le niveau de vie des ménages à un niveau spécifié.

Le déflateur du PIB mesure les variations de prix dans leur domaine respectif de l'économie. Il est utilisé pour corriger le PIB des effets de l'inflation. De manière générale, un déflateur implicite mesure les variations de prix dans un domaine de l'économie en divisant la grandeur en valeur par cette même grandeur en volume.

Les déflateurs implicites sont nommés d'après l'agrégat utilisé. Les déflateurs du PIB, de la dépense de consommation finale, de la formation de capital brute, des exportations et des importations mesurent les variations de prix dans leur domaine respectif de l'économie.

Le déflateur du PIB s'écarte de l'IPC, en fonction notamment, de l'évolution des prix des importations, des exportations et de la FBCF.

Source : Insee

Ainsi, lorsque la croissance économique est calculée à partir du PIB en euros courants, on obtient la **croissance nominale**. En revanche, la croissance réelle ou en volume correspond à la croissance du PIB en euros constants. On peut retenir le tableau de correspondance suivant :

Croissance nominale	Croissance réelle
Croissance en valeur	Croissance en volume
Croissance à prix courants	Croissance en prix constants
Croissance en euros courants	Croissance en euros constants

8.2.2 Croissance tendancielle et taux de croissance annuel moyen

Le taux de croissance étant instable, et ce point sera développé dans le prochain chapitre, on peut aussi calculer la

croissance tendancielle à partir du TCAM. En effet, le taux de croissance annuel du PIB varie plus ou moins d'une année à l'autre. La tendance va représenter, par rapport aux taux annuels, ce que représente la moyenne des notes par rapport aux notes successives d'un élève. Le TCAM permet de déterminer le trend de la croissance sur une période donnée. Il nous permet de mesurer l'ampleur de la hausse du PIB en moyenne chaque année sur la période. En tant que moyenne, le TCAM est donc constant. Pour mieux comprendre l'intérêt de calculer ce TCAM, surtout sur longue, voire très longue période, on peut se demander combien d'années il faut au PIB pour doubler selon le taux de croissance annuel moyen choisi.

Taux de croissance annuel moyen	Période de doublement (CM = 2)
1%	70 ans
1,5%	47 ans
2%	35 ans
3%	23 ans
4%	17 ans
5%	15 ans
7%	10 ans
10%	7 ans
15%	4,66 ans

On peut alors établir une correspondance simple entre TCAM et période de doublement d'une valeur (CM = 2). On peut utiliser ce qu'il est convenu d'appeler la règle des 70 (ans) pour déterminer le nombre d'années nécessaire pour que le PIB double : période de doublement = 70 / TCAM.

8.2.3 Les indicateurs de variation de la production par habitant ou du revenu par tête

Étant donné que le PIB peut être envisagé comme un gros gâteau à partager, *in fine*, entre tous les agents économiques de l'économie, on ne peut se contenter de calculer la taille et la variation du PIB, en tant que gâteau global. On doit aussi calculer la part moyenne de gâteau-PIB que produit et reçoit

chaque habitant et le taux de croissance du PIB par tête. Certes, il ne s'agit que d'une moyenne, et les écarts de répartition entre les habitants ne sont pas pris en compte (les inégalités économiques).

À population donnée ou avec variation de la population, on obtient :

$$\text{PIB/hab} = \frac{\text{PIB}}{\text{Population totale}}$$

C'est une estimation de la richesse produite (ou de la production réalisée) en moyenne par chaque habitant pendant un an (puisque le PIB est calculé sur un an). En particulier, en dynamique, le rapport entre croissance démographique et croissance économique va devenir crucial, même s'il ne s'agit que d'une moyenne.

$$\text{Variation PIB/hab} = \frac{\text{Variation du PIB}}{\text{Variation de la population totale}}$$

8.2.4 Comparaisons internationales et parités de pouvoir d'achat (PPA)

Pourtant, lors de comparaisons internationales ou dans le temps, le niveau du PIB par habitant donne une indication utile et essentielle à condition de faire quelques corrections relatives aux niveaux des prix et à leurs évolutions et de convertir les grandeurs économiques au taux de change PPA (les parités de pouvoir d'achat [PPA] sont des taux permettant de convertir les prix dans une monnaie commune tout en éliminant les différences de pouvoir d'achat entre monnaies). En d'autres termes, leur utilisation permet d'éliminer l'effet, lors de la conversion, des différences de niveau des prix entre pays. En pratique, les taux de change PPA sont difficiles à calculer. L'indice Big Mac permet d'obtenir une approximation correcte d'un taux PPA, là où est vendu le big mac.

8.2.5 Du concept de croissance potentielle à sa mesure : un exercice compliqué

Enfin, pour prolonger notre sous-partie sur la distinction entre le PIB effectif et le PIB potentiel, il est utile de savoir que si les concepts de PIB potentiel et de croissance potentielle sont faciles à comprendre, en revanche, mesurer ces deux grandeurs est un exercice très délicat. Nous parlerons, de préférence, d'estimation pour bien marquer notre prudence et nous préciserons ces points dans la partie sur les facteurs de la croissance économique.

En effet, la croissance potentielle va dépendre notamment de notre capacité à améliorer la quantité, la qualité, la combinaison des facteurs de production...

Les problèmes méthodologiques qui rendent complexes la définition et surtout la mesure de la croissance potentielle ne doivent pas nous conduire à ignorer le concept et à sous-estimer son intérêt et l'importance de son estimation.

8.2.6 *Des limites du PIB aux limites de la mesure de la croissance économique*

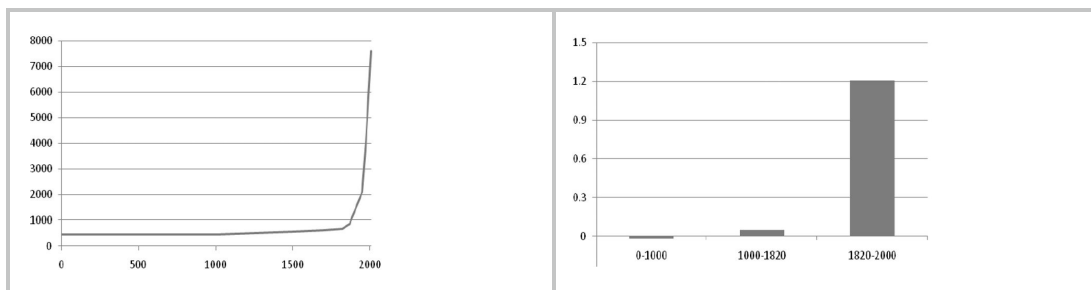
Les limites du PIB en tant qu'indicateur de croissance économique découlent logiquement des limites du PIB en tant qu'indicateur de production globale. Quand le PIB augmente, c'est-à-dire quand il y a croissance économique, nous savons que celle-ci peut être sous-estimée ou surestimée. Ainsi, la production globale de biens et services pour satisfaire nos besoins sera sous-estimée du fait de la croissance des productions domestiques légales ou de toutes les productions illégales déjà mentionnées par ailleurs. En revanche, la croissance économique sera surestimée quand le PIB augmentera davantage et plus vite sous l'impact de productions entraînant des externalités positives ou négatives. Dans le cas d'externalités négatives, la croissance sera plus forte grâce à des productions qui permettent de satisfaire les besoins des uns, mais en dégradant la satisfaction des autres, voire des mêmes, d'ailleurs. De manière analogue, la croissance du PIB sera plus forte en

raison de l'augmentation des activités de réparation des externalités négatives. Ici, le PIB augmente quand on répare ce qui a été dégradé du fait d'autres productions comptabilisées, qui devaient améliorer la satisfaction des besoins. En résumé, plus on produit et plus on croît en dégradant et en nuisant d'une part, et plus on doit compenser, réparer, d'autre part. Mais, dans les deux cas, le PIB augmente et la croissance peut être plus forte.

Maintenant que nous connaissons les indicateurs de croissance et les limites de leur capacité à mesurer ce qu'ils sont censés mesurer, alors nous pouvons utiliser ces indicateurs pour observer les évolutions de la croissance économique sur le long, voire très long terme.

8.3 LES TENDANCES DE LA CROISSANCE ÉCONOMIQUE OU LA CROISSANCE TENDANCIELLE

Le calcul du TCAM sur longue période nous permet de mettre en évidence la croissance tendancielle, c'est-à-dire le taux de croissance qu'aurait connu l'économie s'il était resté strictement stable à long terme.



GRAPHIQUE 8.1.

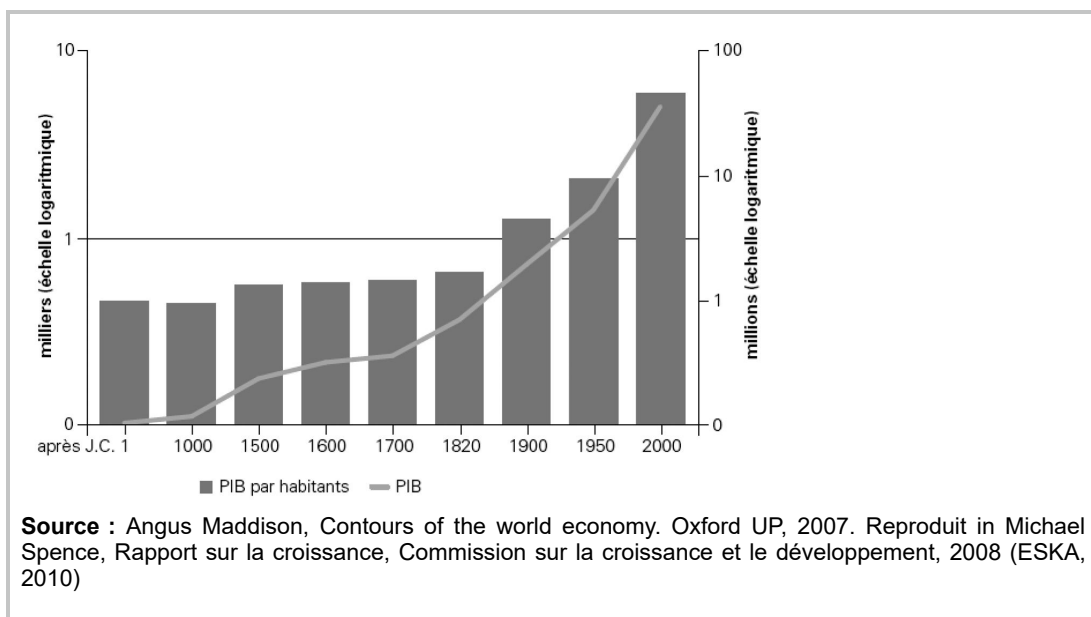
Le PIB par habitant de l'an 1 à l'an 2000 (en \$ 1990, aux parités de pouvoir d'achat)
 "Oded Galor, From Stagnation to Growth", Brown Lecture, 7 décembre 2011 (Données : Angus Maddison).

GRAPHIQUE 8.2.

Le taux de croissance annuel moyen du PIB par habitant de l'an 1 à l'an 2000
 "Oded Galor, From Stagnation to Growth", Brown Lecture, 7 décembre 2011 (Données : Angus Maddison).

L'observation des données mondiales montre une croissance incontestable du PIB mondial, mais une croissance inégale selon les pays et les régions du monde. En nous appuyant sur les travaux d'Angus Maddison, le grand spécialiste de l'histoire quantitative de la croissance à très long terme, nous constatons que la croissance était quasi nulle jusqu'en 1500, qu'il y a eu un frémissement entre 1500 et 1820 et que le vrai départ de la croissance économique du monde remonte au début du XIX^e siècle. Ainsi, la hausse du revenu par habitant entre 1500 et 1820 serait environ trente fois inférieure à celle que nous observons sur la période 1820-2000.

On peut aisément observer une accélération de la croissance mondiale depuis deux siècles, avec la Révolution industrielle et la diffusion progressive de ses effets. Après 1820, même si on ne peut dater aussi précisément le début de la croissance, la Révolution industrielle produit ses effets au-delà de l'Angleterre (1750-80). Une incontestable dynamique de croissance se met en place. En effet, de 1500 à 1700, la croissance du produit par tête est devenue positive, mais faible, de l'ordre de 0,1% par an, passant à 0,2% par an entre 1700 et 1820. Même pendant la Révolution industrielle, la croissance n'était pas très élevée par rapport aux niveaux actuels.



GRAPHIQUE 8.3.

Évolution du PIB et du PIB par habitant dans le monde de l'an 1 à l'an 2000
(en \$ 1990, aux parités de pouvoir d'achat)

La croissance économique est donc un phénomène récent à l'échelle de l'histoire humaine. Mais elle n'a pas commencé partout au même moment, et n'a pas connu le même rythme selon les aires géographiques. On peut considérer qu'avant 1820, l'économie stationnaire dominait les économies nationales. Dans ces économies qualifiées de « préindustrielles » ou « traditionnelles » (dominées par l'agriculture), la croissance est inexistante ou presque. On peut observer un frémissement, entre 1500 et 1820, dans la mesure où le taux de croissance annuel moyen devient positif à la suite de grandes découvertes. À l'aune de ce qu'on a pu observer sur les deux derniers millénaires, les périodes de croissance soutenue et durable des années glorieuses pour les pays d'Europe de l'Ouest font plutôt figure d'exception, comme le seront probablement celles de la Chine des années 2000 dans les décennies à venir. Notons qu'on pourrait observer des différences au niveau de la croissance démographique simultanément. Cela permettra d'expliquer une partie des inégalités internationales en matière de croissance et de niveau de PIB par tête, aujourd'hui et à différentes époques du passé. Les grandes tendances de la croissance économique étant présentées, nous devons désormais tenter d'en expliquer les principales raisons.

L'observation de la croissance dans la longue durée a conduit les économistes à mettre en évidence plusieurs faits stylisés, des régularités empiriques et généralement, l'analyse de Kaldor sert de référence.

En effet, en 1961, Kaldor a essayé de mettre en évidence, à partir d'observations empiriques, des caractéristiques de la croissance économique à long terme dans les pays occidentaux. Il a énoncé six constats, six faits stylisés :

1. Un taux de croissance continu et relativement régulier de la production par tête

2. Un taux de croissance continu et relativement régulier du stock de capital par tête
3. Un taux de rendement du capital (taux de profit) approximativement constant
4. Un ratio capital/produit constant : la productivité moyenne du capital (produit/capital) est constante
5. Les parts du capital et du travail dans le revenu national sont constantes. Quasi-stabilité de la répartition du revenu national entre travail et capital
6. La persistance de fortes disparités internationales des taux de croissance par tête : les pays ont des taux de croissance de la productivité différents.

Ces six faits stylisés sont interdépendants. Rigoureusement, les faits 1 à 5 caractérisent le processus de croissance d'une économie considérée isolément, mais ce processus est commun à toutes les économies analysées par Kaldor. En revanche, le fait 6 est de nature comparative, car il permet de mettre en évidence une différence dans les processus de croissance au niveau international.

Plusieurs décennies après l'énoncé de ces faits stylisés, on admet que, de manière générale, ils correspondent toujours aux observations empiriques dont nous disposons.

À long terme, la croissance de la production (revenu) par tête est continue et régulière, ce qui confirme le **fait 1**. C'est le trait majeur de la croissance économique moderne. Il y a similitude de la croissance par tête dans des économies qui ont atteint un même stade de développement. Les observations confirment le **fait 4** concernant la constance du ratio capital/produit (coefficient de capital) c'est-à-dire de la productivité moyenne du capital (produit/capital). Le taux de croissance du stock de capital (constant) et approximativement égal à celui de l'activité (production).

La vérification simultanée des faits stylisés 1 et 4 implique logiquement celle du **fait 2** : taux de croissance continu et régulier du stock de capital par tête. Ce fait confirme la

prédiction de Ricardo et de Marx pour qui l'accumulation du capital par tête peut croître de manière ininterrompue.

Le **fait 3** indique que le taux de profit de longue période est constant, ce qui implique la stabilité de la part du revenu des facteurs (travail et capital) dans le revenu national (stabilité de la répartition du revenu, **fait 5**).

Remarquons que les faits 3 et 5 illustrent l'erreur de prédiction ricardienne (état stationnaire), celle du marxisme et celle du marginalisme (les néoclassiques font l'hypothèse que les parts du travail et du capital sont constantes). Conformément au fait 3, il n'y a pas de baisse tendancielle du taux de profit (comme l'affirmait Marx) et ce fait est contradictoire avec le principe des rendements marginaux décroissants. De même, l'absence de baisse tendancielle du taux de profit est confirmée par la constance des parts du capital et du travail dans la répartition du revenu.

À long terme, le taux de profit est approximativement égal à la somme du taux d'intérêt réel et du taux de dépréciation du capital.

Les faits stylisés de Kaldor sont donc approximativement vérifiés. À long terme, il existe des régularités empiriques de la croissance économique.

Cela justifie la recherche de modèles théoriques susceptibles de reproduire et d'expliquer ces régularités.

Cependant, ces faits stylisés n'indiquent rien sur les causes de la croissance. Ils se contentent de caractériser son apparence, non de l'expliquer.

8.4 LES PRINCIPALES EXPLICATIONS DE LA CROISSANCE ÉCONOMIQUE À LONG TERME

8.4.1 *Les causes de la richesse des nations*

Les différentes conceptions de l'enrichissement des nations précèdent les théories de la croissance. En effet, la

« croissance économique » restera longtemps une énigme dans l'histoire de la pensée économique. La question « pourquoi certains pays sont-ils riches et certains pays sont-ils pauvres ? » est l'une des plus classiques dès les premières réflexions en économie politique. Les principales causes de l'enrichissement des nations recensées par les différentes écoles de pensée étaient :

- La conquête de l'or par les découvertes, l'industrie et l'exportation, le commerce pour les mercantilistes.
- Seule la terre est capable de donner plus qu'elle ne coûte, est source de valeur, pour les physiocrates. Une civilisation, ou plus modestement, un royaume, est riche s'il peut dégager un surplus agricole.
- Le monde est fini et la « *La rente, loin de mesurer la générosité divine, en montre au contraire l'avarice* » selon les classiques. L'avarice de la nature est donc liée à la loi des rendements marginaux décroissants. La croissance démographique sera un obstacle à la croissance de la production par habitant selon les prévisions alarmistes de Malthus et Ricardo. En précurseur, Cantillon affirmait que les hommes, lorsqu'ils ne sont pas contraints par le manque de nourriture, se reproduisent « comme les souris dans une grange ». C'est la loi de Malthus qu'on retrouvera chez Ricardo.
- Seul le travail qui a la propriété de donner plus qu'il ne coûte à entretenir est source de valeur, selon Marx et les marxistes.
- La croissance de la population et celle du stock de capital sont possibles, mais, en raison de la loi des rendements marginaux décroissants, pas celle des niveaux de vie selon le marginalisme néoclassique (voir 2.2.2).

Ces approches de « l'enrichissement des nations », pour reprendre la célèbre formule d'Adam Smith dans son ouvrage classique de 1776, ne proposent pas de véritables théories de la croissance. Plus tard, et progressivement, l'amélioration

des connaissances en économie permettra la mise en évidence des déterminants durables de la croissance économique en prenant appui sur des catégories de facteurs de production dont il s'agira d'estimer les apports relatifs.

8.4.2 *La théorie moderne de la croissance : la décomposition des effets des facteurs de croissance*

L'approche par les facteurs de production comme explication de la croissance est microéconomique. Au niveau macroéconomique, on préférera parler de moteurs de la croissance tels que la consommation des ménages, l'investissement des entreprises, les exportations...

Dans les modèles de croissance les plus simples, fondés sur des fonctions de production agrégées, les économistes distinguent le facteur travail, le capital productif physique et le progrès technique, dans un sens large.

Logiquement, on peut obtenir plus de production et de croissance à partir d'une augmentation quantitative de la population active combinée à l'amélioration de la qualité du travail, à l'accumulation du stock d'un capital productif et au progrès technique, au sens d'un ensemble d'innovations (sous toutes leurs formes) qui viendront faire progresser la qualité de ces facteurs de production et leur combinaison. De fait, nous sommes conduits à distinguer une croissance dite « extensive », lorsqu'elle est portée par la hausse des quantités de facteurs (toutes choses égales par ailleurs, davantage de travailleurs et d'équipements conduisent à plus de croissance), d'une croissance dite « intensive », qui sera fondée sur une utilisation-combinaison plus efficace de facteurs de production plus performants. La croissance intensive sera portée par les gains de productivité, c'est-à-dire que la production augmentera plus vite que les quantités de facteurs.

8.4.3 Les fondements de la théorie de la croissance selon Robert Solow

Dans le cadre théorique proposé par Robert Solow en 1956, les principaux facteurs de croissance sont, pour une quantité de travail donnée, l'accumulation de capital productif et le progrès technique (PT). L'ampleur de l'impact du PT sera estimée par la productivité globale des facteurs (PGF). On observera qu'à long terme, la croissance économique dépend essentiellement de l'augmentation de la PGF. Dans ces conditions, il sera indispensable de se poser la question des origines de cette grandeur et donc, de la source majeure de croissance.

8.4.3.1 Le modèle de base de la croissance

Au niveau agrégé, la fonction de production est donnée par $Y = F(K, L)$ et la production par travailleur ou par tête ou $y = f(k)$ avec $y = Y/L$ et $k = K/L$.

Cette fonction de production permet d'expliquer le niveau du PIB potentiel de long-terme.

Dans ce cadre de base, il n'existe **pas de progrès technique**. La quantité de biens qu'il est possible de produire (Y) dépend des facteurs primaires de production : le nombre de machines (K , capital productif physique) et du nombre de personnes qui travaillent et produisent (L , facteur travail). La croissance de la production globale dépend de l'augmentation des quantités de facteurs travail et capital.

Deux hypothèses conditionnent les enseignements de ce modèle : la loi des rendements d'échelle constants et la loi des rendements marginaux décroissants.

Si **les rendements d'échelle sont constants**, la production augmente à la même vitesse que la quantité de facteurs. Pour un nombre z donné, on a $zY = F(zK, zL)$.

Si on combine 10 machines à 10 hommes pour produire 100 unités d'un, alors 20 machines et 20 hommes permettront d'en produire 200. On aura $2Y = F(2K, 2L)$.

Les fonctions de production avec des rendements d'échelle constants ont des implications intéressantes. Ainsi, en posant, $z=1/L$. L'équation $[zY = F(zK, zL)]$ devient : $Y/L=F(K/L,1)$.

Y/L , la production par travailleur (mesure de la productivité), est une fonction directe du capital par travailleur.

Selon la **loi des rendements marginaux décroissants** (LRMD), lorsque la quantité d'un des facteurs est constante, alors, l'augmentation de la production qu'il est possible d'obtenir en accroissant d'une unité la quantité de l'autre facteur est de plus en plus faible. La productivité marginale des facteurs décroît.

8.4.3.2 La croissance économique s'explique par l'accumulation du capital par travailleur

Dans ces conditions, toutes choses égales par ailleurs, l'accumulation du capital physique, via l'investissement, va déterminer l'intensité de la croissance économique. En l'absence de progrès technique, l'accumulation de capital physique va permettre d'entretenir, au moins temporairement, la croissance. Une plus grande quantité de capital augmente la capacité de production et soutient la croissance, directement, car la hausse du stock de capital se traduit par une hausse de la capacité de production et indirectement par l'augmentation du capital par tête.

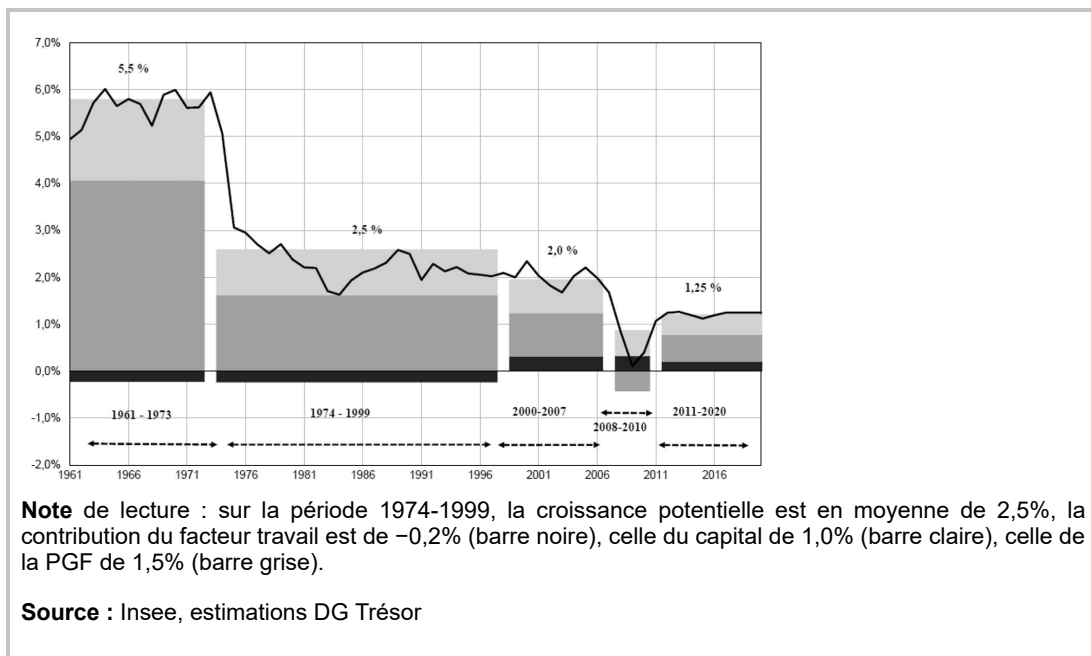
Dans cette acception, l'accumulation du capital aura un impact sur la croissance économique, mais ses effets seront de plus en plus faibles, en raison de la LRMD : la productivité marginale du capital est décroissante, car chaque unité de capital supplémentaire entraîne une hausse plus faible de la productivité, donc de la capacité de production, que la précédente. Il sera possible d'accélérer transitoirement la croissance économique, mais, à long terme, le taux de croissance rejoindra un taux régulier et l'économie convergera vers un état stationnaire.

Dans le modèle de croissance de Solow sans progrès technique, on ne peut donc pas faire augmenter durablement la production et la croissance de cette production en faisant seulement croître les quantités de facteurs primaires de production, c'est à dire, en faisant croître la population en âge de travailler qui travaille (et produit) et en accumulant du capital productif physique.

8.4.4 La « comptabilité » de la croissance, le résidu et la PGF

Comment expliquer les divergences des PIB par tête (mesurés dans la même unité), dans l'espace et dans le temps ? La réponse à cette question est un défi pour la théorie et la comptabilité de la croissance à long terme.

Ainsi, pour comprendre d'où vient la croissance, les économistes commencent par tenter de mesurer la contribution des différents facteurs primaires de production.



GRAPHIQUE 8.4.

Croissance potentielle et contributions moyennes des facteurs en France depuis 1961

La comptabilité de la croissance est une méthode qui doit permettre de mesurer ce qui, dans la croissance du PIB, peut s'expliquer par la croissance des quantités de facteurs de production. Cette comptabilité s'est développée à la fin des années cinquante à la suite des travaux de Solow (1957) et surtout de Denison (1962), (1967). Il ne s'agit en aucun cas d'une théorie de la croissance, mais d'une méthode d'analyse quantitative. Le point de départ de la comptabilité de la croissance est la fonction de production agrégée.

Le résidu de la croissance correspond à la fraction du taux de croissance de la production agrégée qui n'est pas expliquée par le taux de croissance des quantités de facteurs (i.e. au « progrès technique »). Devant l'importance de ce résidu (Solow l'estime aux 7/8èmes de la croissance des États-Unis sur la période 1909-1949), on a cherché à prendre en compte non plus seulement l'augmentation quantitative des facteurs, mais aussi leurs améliorations qualitatives (ainsi, pour le facteur travail, l'amélioration du niveau moyen de formation).

Supérieure à 5 % dans les années 1960, la croissance potentielle aurait fléchi à partir de la fin des Trente Glorieuses (1945-1973) pour atteindre environ 2% dans la période précédant la crise.

Cette comparaison EU/UE, sur la période 1966-2008, montre un peu plus l'ampleur du poids de la PGF dans la croissance à long terme aux EU et dans l'UE considérée comme une seule économie. On peut également observer, depuis les années 60, une tendance à la réduction du taux de croissance de la PGF et, simultanément, du taux de croissance du PIB.

8.4.5 *Seul le progrès technique permet durablement une croissance soutenue*

L'introduction du progrès technique dans la fonction de production (désormais à rendements d'échelle croissants) et les modèles de croissance permettent d'expliquer le fait que

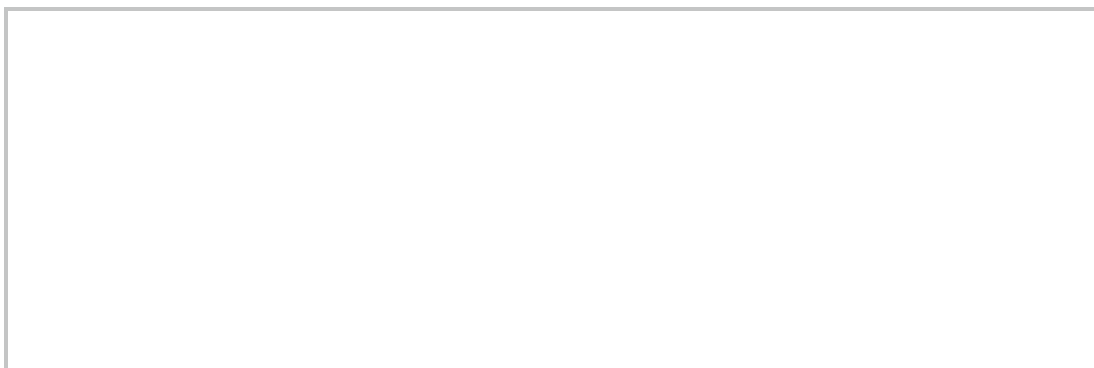
la production augmente plus rapidement que les quantités de facteurs utilisées. Autrement dit, que le produit ou revenu par tête puisse augmenter de manière soutenue et durablement.

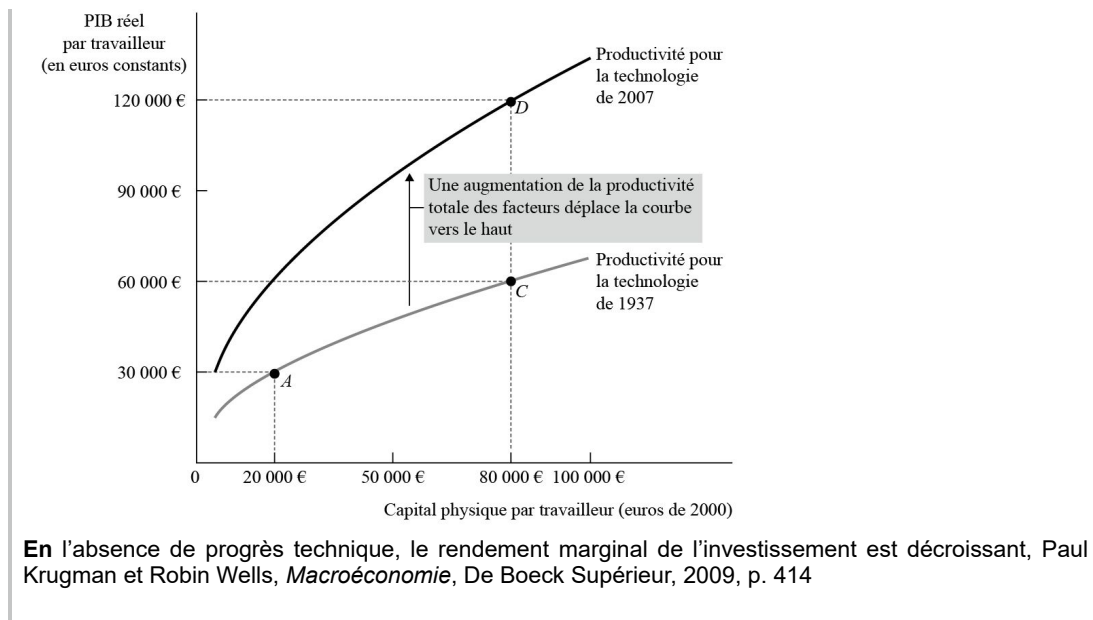
Au niveau agrégé, cette fonction de production devient par $Y = A \times F(K, L)$ et la production par travailleur ou par tête ou $y = A \times f(k)$ avec $y = Y/L$ et $k = K/L$.

L'existence du **progrès technique** se traduit par une **multiplication de la capacité de production des facteurs primaires**, travail et capital physique. Le progrès technique a pour effet d'accroître la capacité de production des travailleurs parce qu'il en augmente l'efficacité. Si on considère qu'un travailleur produit aujourd'hui près dix fois ce que le travailleur au début du xx^e siècle produisait, c'est comme si le nombre de travailleurs avait été multiplié par 10.

Le progrès technique permet de lutter contre les limites imposées par la loi des rendements marginaux décroissants. Dans la « comptabilité » de la croissance, le progrès technique va se traduire par une hausse de la productivité globale des facteurs (PGF), du résidu.

L'augmentation de la PGF est souvent associée au progrès technique, mais, plus largement, elle englobe d'autres facteurs tels que les infrastructures et les institutions. Cette hausse de la PGF pourra donc s'expliquer par les effets des différents types d'innovation, de leurs combinaisons, de la diffusion progressive de leurs effets, mais aussi par les infrastructures et arrangements institutionnels qui encadrent l'activité des agents économiques et déterminent leur degré d'initiative.





GRAPHIQUE 8.5.

PIB par travailleur et capital par travailleur

Concrètement, ce sera grâce à la combinaison de nombreuses innovations et à l'accumulation de différentes formes de capital qu'il sera possible de maintenir durablement une croissance soutenue. On pourra investir dans du capital humain (connaissances, compétences), dans la recherche-développement, la technologie, les infrastructures, les institutions...

Le graphique ci-dessus permet de visualiser d'une part les effets de la productivité marginale décroissante et d'autre part les effets de la PGF. Sur la courbe du bas, on observe clairement la décroissance de la productivité du capital par travailleur. La courbe du haut résulte de l'augmentation de la PGF qui entraîne une hausse du niveau de productivité du capital par travailleur et une accélération de la production par travailleur. La PGF, qui est l'expression des effets du progrès technique, permet de repousser le mur des rendements marginaux décroissants.

8.4.6 Ralentissement de la croissance de la PGF et « stagnation séculaire »

*Pour l'ensemble des pays de l'OCDE, on peut observer, depuis les années 70, une réduction du taux de croissance de la productivité globale des facteurs (PGF). La concomitance de 1975 à 1995 de ce phénomène et de celui du fort accroissement des dépenses de R&D a donné naissance au fameux « **paradoxe de la productivité** ». Il n'empêche que la part de la croissance imputable à la variation de la PGF demeure élevée dans certains pays du fait de la baisse parallèle de la croissance expliquée par les facteurs traditionnels de production (travail et capital).*

Le « **paradoxe de la productivité** » est également connu sous le nom de paradoxe de Solow. « On voit des ordinateurs partout sauf dans les statistiques de la productivité », R. Solow, "We'd Better Watch Out", New York Times Book Review, 12 juillet 1987. Malgré le développement des nouvelles technologies de l'information et des télécommunications, la croissance de la productivité des années 1980 et 1990 a été plus faible qu'au cours des deux décennies précédentes.

En raison du ralentissement de la croissance de la **productivité globale des facteurs** (PGF), c'est-à-dire la partie de la croissance de la production qui n'est expliquée ni par la croissance de l'emploi ni par celle du stock de capital productif, il se pourrait bien d'ailleurs, selon de nombreux experts internationaux, que nous devions à l'avenir nous habituer à des taux de croissance du PIB durablement faibles.

Cette hypothèse d'une « **stagnation séculaire** », contenue dans les travaux récents de l'économiste américain Robert Gordon, professeur à l'université Northwestern, a déclenché un vif débat académique aux États-Unis à partir de l'été 2012. Cette intuition, qui contient l'idée que la croissance rapide que le monde a connue depuis 250 ans constitue un épisode unique dans l'histoire de l'humanité, est en effet capitale, car elle générerait une baisse tendancielle de notre revenu par tête à long terme.

Cette **décélération de la productivité et de la croissance**, qui pourrait bien se poursuivre durant le siècle prochain, impliquerait des tensions sociales grandissantes sur le partage des richesses, dans le cadre d'un jeu à somme nulle entre les groupes sociaux (ce que l'un gagne, l'autre le perd).

En effet, de nombreux facteurs pourraient agir comme des « vents contraires », selon la formule de Robert Gordon, et freiner l'expansion du PIB jusqu'à la ramener à des niveaux très faibles (autour de 0,2%) : le vieillissement démographique (hausse des dépenses de retraites, baisse de l'incitation à investir), l'inefficacité croissante des systèmes éducatifs (les rendements des investissements dans l'éducation diminuent), la compression du pouvoir d'achat sous l'impact de la mondialisation, les effets du désendettement public et privé, le coût de la lutte contre le réchauffement climatique ou encore la montée des inégalités agiraient comme de puissants freins à la croissance.

8.4.7 L'accumulation de capital par tête et la PGF vont expliquer les gains de productivité du travail.

$$\Delta (K/L) + \Delta \text{PGF} \rightarrow \Delta (Y/L)$$

En effet, on peut considérer à ce stade que les gains de productivité du travail ont deux origines : l'accumulation du capital par tête et la productivité globale des facteurs.

C'est pour ces raisons que les économistes insistent sur l'importance cruciale des gains de productivité du travail. Pour une quantité de travail donnée, on retrouve l'impact du capital physique, mais également du capital humain, de la technologie, des connaissances technologiques... sur la productivité du travail. Mais les deux facteurs ont des impacts très différents sur le processus de croissance.

L'accumulation du capital, au sens le plus étroit, a des effets limités, compte tenu de la loi des rendements marginaux décroissants. Dans le modèle de Solow, une hausse du taux d'épargne dans l'économie pourra permettre d'accroître le niveau de l'investissement et de la production, mais n'aura pas d'effet durable sur le taux de croissance. En revanche, une hausse de la PGF, induite par le progrès technique, se

traduira par une hausse durable du niveau de la production globale et du taux de croissance de l'économie. *A contrario*, un ralentissement de la croissance de la PGF se traduira par une moindre hausse, voire une baisse, du niveau de la production globale et un ralentissement du taux de croissance de l'économie.

Une accélération de la croissance de la PGF doit entraîner une accélération des gains de productivité du travail et donc de la croissance potentielle.

Une décélération de la croissance de la PGF doit entraîner une décélération des gains de productivité du travail et donc de la croissance potentielle.

8.5 LES ORIGINES DU PROGRÈS TECHNIQUE

8.5.1 *Le progrès technique*

Selon Jean Fourastié (économiste français du xx^e siècle), le progrès technique désigne une capacité d'action de plus en plus efficace sur les éléments matériels.

Pour les économistes Olivier Blanchard et Daniel Cohen, « *dans un sens restreint, on peut considérer que l'état de la technologie est la liste de tous les plans de fabrication de tous les produits et des techniques qui permettent de les produire.* » Pour autant, dans un sens plus large, « *la quantité de production possible dans une économie dépend également de l'organisation des entreprises, du fonctionnement du marché, de tout le système de loi, de leur application, de l'environnement politique* ».

Au sens strict, le progrès technique désigne l'ensemble des modifications qui affectent les procédés ou la nature des biens réalisés, ou encore l'**ensemble des innovations** qui améliorent l'efficacité de la combinaison productive et la qualité des produits. Cette notion se réfère à l'idée de progrès. Autrement dit, le progrès technique suppose une amélioration. Dès que l'on parle de progrès, on fait

l'hypothèse d'une évolution positive, d'un changement en bien, en mieux.

8.5.2 *Les différentes formes d'innovations*

En ce qui concerne les innovations, elles peuvent prendre plusieurs formes. Selon l'économiste autrichien Joseph Schumpeter (1883-1950), l'innovation est une combinaison nouvelle des facteurs de production (travail, capital) qui peut prendre cinq formes : nouveau produit, nouveau procédé, nouveau marché (ou nouveau débouché), nouvelles formes d'organisation des entreprises, nouvelles sources de matières premières. Nous allons ici en retenir trois.

Les innovations de produit consistent à mettre sur le marché de nouveaux produits. Si l'innovation est réussie, les économies d'échelle qui en résultent se traduisent par des gains de productivité et de la croissance. **Les innovations de procédé** correspondent à de nouveaux procédés (machines) dans les processus de production. Il en est ainsi de l'automatisation, qui permet de rationaliser le travail, de gagner du temps et donc de réaliser des gains de productivité. On peut retenir comme exemple le convoyeur de Ford. Les innovations de procédé s'accompagnent souvent d'**innovations organisationnelles** en matière de travail. Ainsi, le convoyeur de Ford s'est accompagné du développement du travail à la chaîne. Ceci étant, d'autres formes d'organisation du travail post-tayloriennes, comme le toyotisme ou ohnisme (en référence à Taïchi Ohno, patron de Toyota), favorisent le travail en équipe, le travail par projet, le temps de travail flexible, l'implication personnelle dans le travail, donc une moindre division du travail, et peuvent permettre d'améliorer la productivité. Les innovations organisationnelles peuvent aussi favoriser l'émergence de nouveaux produits, de nouveaux services et de procédés plus performants.

Il y a donc de vraies **interactions entre les différentes formes d'innovation**. On voit, encore une fois, que la

classification n'implique pas la séparation, mais juste la distinction, dans un but pédagogique.

Quel que soit le type d'innovation présenté ci-dessus, quand on s'intéresse aux effets des innovations sur la croissance économique, on ne peut échapper à la distinction entre les **innovations mineures**, progressives ou incrémentales (par incrémentation) et les **innovations majeures**, radicales, de rupture. Les premières induisent des améliorations des techniques utilisées ou des produits commercialisés, sans introduire de changement radical. On va observer une association entre différentes technologies connues, mais dans une fonctionnalité nouvelle. Les secondes provoquent un saut qualitatif par le biais d'améliorations techniques d'un niveau supérieur. Elles accompagnent ou entraînent des modifications des systèmes productifs, commerciaux, financiers. Toutes choses égales par ailleurs, les innovations majeures se traduiront par une accélération plus forte de la PGF et, *in fine*, de la croissance économique.

8.5.3 Les origines des innovations

Qui est à l'origine de l'innovation ? D'où viennent les innovations ? Pour répondre à ces questions, il faut distinguer l'analyse de Schumpeter de celle des théories, plus récentes, de la croissance endogène, fondées sur l'accumulation du capital sous toutes ses formes.

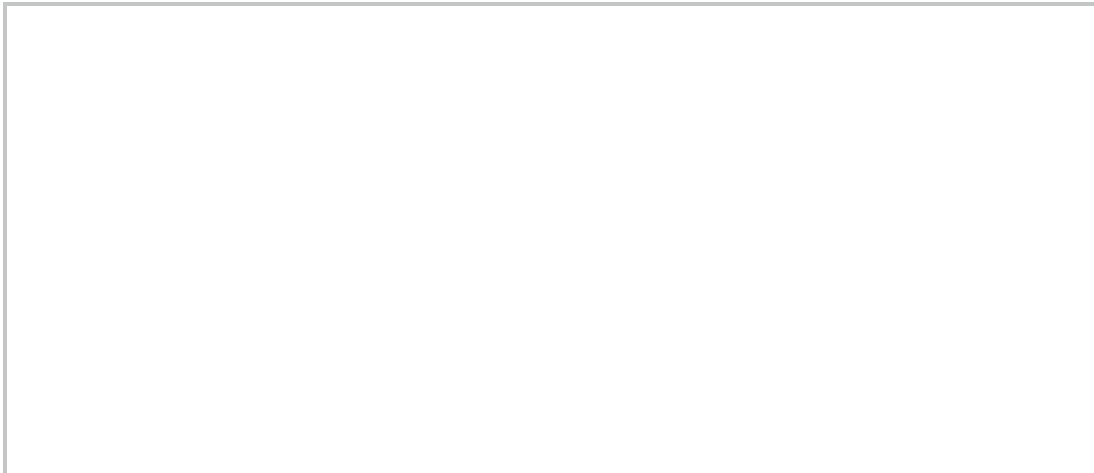
8.5.3.1 Le rôle central des entrepreneurs et le processus de destruction créatrice (PDC)

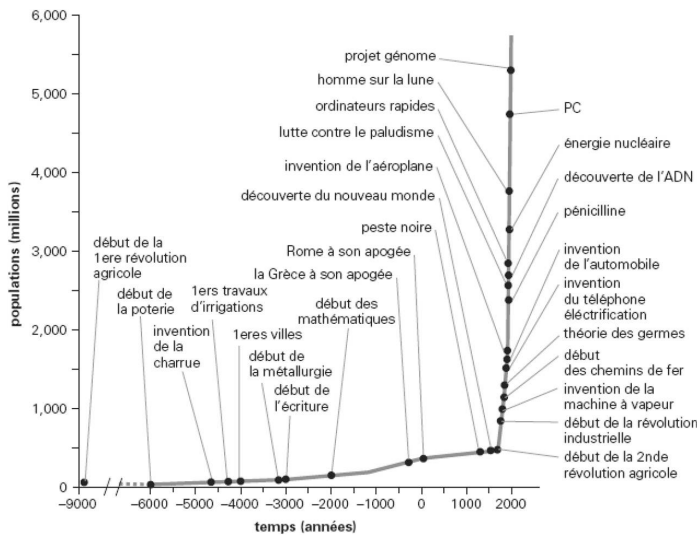
Selon les intuitions de Joseph Schumpeter, le processus d'innovation est le résultat de l'activité d'une catégorie d'individus exceptionnels, et contingent au capitalisme de marché : **les entrepreneurs**. Nous verrons que cette approche aura des implications majeures sur les politiques économiques à mener. Pour Joseph Schumpeter, « [...] *le rôle de l'entrepreneur consiste à réformer ou à révolutionner la routine de production en exploitant une invention ou, plus généralement, une possibilité technique inédite (production*

d'une marchandise nouvelle, ou nouvelle méthode de production d'une marchandise ancienne, ou exploitation d'une nouvelle source de matières premières ou d'un nouveau débouché, ou réorganisation d'une branche industrielle, et ainsi de suite). [...] C'est à ce genre d'activités que l'on doit primordiallement attribuer la responsabilité des "prospérités" récurrentes qui révolutionnent l'organisme économique, ainsi que des "récessions" non moins récurrentes qui tiennent au déséquilibre causé par le choc des méthodes ou produits nouveaux ».

Le processus de destruction créatrice qui en découle, « constitue la donnée fondamentale du capitalisme : c'est en elle que consiste, en dernière analyse, le capitalisme et toute entreprise capitaliste doit, bon gré mal gré, s'y adapter [...] En d'autres termes, le problème généralement pris en considération est celui d'établir comment le capitalisme gère les structures existantes, alors que le problème qui importe est celui de découvrir comment il crée, puis détruit ces structures » (J. Schumpeter, *Capitalisme, socialisme et démocratie*, 1942).

On peut observer que la croissance des revenus s'est accompagnée d'accroissements sans précédent de la population et logiquement, d'une accélération et d'une multiplication exponentielles du rythme des découvertes scientifiques. Les innovations portent bien la croissance économique sur le très long terme.





Robert Fogel, *Catching up with the Economy*, American Economic Review, 89 (1), 1999. Reproduit in Michael Spence, *Rapport sur la croissance*, Commission sur la croissance et le développement, 2008 (ESKA, 2010)

Note de lecture : Il y a généralement un écart entre le moment de l'invention d'un procédé ou d'une machine et la généralisation de son application à la production. « Début » signifie : stade initial du processus de diffusion.

GRAPHIQUE 8.6.

Évolution de la population mondiale et progrès technique depuis le néolithique

8.5.3.2 La RD et le processus d'innovation

Selon les théories de la croissance endogène, les innovations sont davantage le résultat des activités de recherche et développement (R&D). Si l'on se réfère au rapport du Conseil d'analyse économique (CAE) intitulé « Innovation et croissance » (1998), on peut distinguer un modèle linéaire et un modèle interactif pour expliquer l'émergence et la diffusion des innovations. Le **modèle linéaire** est le modèle traditionnel d'explication de l'innovation. L'innovation y est conçue comme un processus linéaire au cours duquel se succèdent, de manière linéaire, des événements (phases), sans qu'il existe de possibilité de rétroaction d'une phase sur la (les) phase(s) antérieure(s). Ainsi, selon ce modèle, tout effort d'innovation passe par un effort antérieur de recherche relevant, avant tout, du domaine de la « Science ». L'innovation relevant du domaine de l'économie, en particulier des entreprises. Or, des études récentes montrent clairement

que l'innovation ne coïncide pas avec l'intensité de la recherche technologique. Autrement dit, il n'existe pas de lien direct et mécanique entre l'intensité de la recherche et la capacité à innover. En outre, cette thèse peut être abondée par des comparaisons internationales. Dans ces conditions, la mesure de la capacité d'innovation par les dépenses en recherche-développement n'est pas fiable). Dans le cadre du **modèle interactif**, l'innovation n'est plus une succession de phases isolées, mais un aller-retour permanent entre des possibilités offertes par la technologie et le marché, des moyens financiers (privés et publics), des stratégies d'acteurs plus ou moins facilitées par l'environnement économique et social. De cette interaction (aller-retour permanent) va dépendre la performance collective. On pourra parler de système d'innovation. L'effort d'innovation devient alors le moyen de la recherche constante d'avantages concurrentiels par les entreprises les plus dynamiques. L'innovation est donc étroitement liée à : la structure de la concurrence entre entreprises (conditions de création, conditions de concurrence...), l'efficacité des facteurs de production et de leurs combinaisons (potentiel d'efficacité), la demande du marché (en qualité et en quantité), la compétitivité des industries proches (en amont et en aval), l'action des pouvoirs publics (qui peuvent influencer les quatre premiers paramètres).

Pour mieux comprendre, on peut mettre en évidence le processus d'innovation et le rôle de la recherche en amont en partant d'un exemple qui décrit la recherche et développement à partir de la mise au point du lecteur CD (apparu sur le marché en 1983 ou en 1979) par Philips et Sony en 1979.

8.5.3.3 De la recherche fondamentale à l'innovation

La recherche fondamentale ou théorique n'a pas d'objectif immédiat ni d'application concrète : il s'agit d'étudier des phénomènes (par exemple, les rayonnements) sans produire d'éléments concrets (matériels). La découverte est suscitée

par la recherche fondamentale et appliquée. La recherche appliquée implique l'acquisition de nouvelles connaissances dirigées vers un but ou un objet déterminé. Elle a pour objectif une application concrète, qu'on appelle une invention. L'invention correspond à l'application technique d'une découverte fondamentale ou plus simplement à la découverte technique. Le développement, le perfectionnement des prototypes initiaux permet le développement expérimental (ou industriel) du produit ou du procédé (à distinguer bien sûr du développement économique et social d'un pays). Cette étape correspond à la mise au point d'un prototype, c'est-à-dire du premier exemplaire d'un produit (prototype). La commercialisation se traduit par la fabrication en série et le lancement du produit sur le marché (il s'agit alors d'un produit grand public).

TABLEAU 8.1.
De la recherche à l'innovation

Étapes	Recherche fondamentale	Recherche appliquée	Développement industriel	Commercialisation
Objectifs	Découverte Théorique Lois générales	Invention Nouvelles conceptions techniques	Innovation (1 ^{re} étape) : mise au point pratique de ces nouvelles conceptions : prototype	Innovation (2 ^e étape) : mise en œuvre de ces nouvelles conceptions grâce à des investissements : Nouveau produit (ou un procédé) destiné au marché
Exemples	Lois physiques des rayonnements	Rayon laser	Prototype du lecteur CD	Lecteur CD

Un grand nombre d'inventions ne dépassent jamais le stade du prototype, faute de rentabilité (aucune entreprise ne veut financer la production en grande série), d'intérêt, etc. L'innovation correspond donc à l'application économique d'une invention. Les étapes 3 et 4 correspondent à la mise au

point d'une innovation. La recherche et développement correspond donc à un processus qui, de la recherche fondamentale à la recherche appliquée et au développement expérimental, permet l'apparition d'une innovation. Le processus n'aboutit pas nécessairement. Nombre d'inventions ne deviennent pas des innovations et toute innovation n'est pas un succès commercial : c'est le risque inhérent à l'innovation. À chaque étape, des obstacles techniques, mais surtout financiers, peuvent ralentir le processus, voire y mettre fin.

8.6 LES THÉORIES DE LA CROISSANCE ENDOGÈNE ET LE RÔLE DES INSTITUTIONS

Même si le modèle de croissance proposé par Solow a été largement critiqué, il propose des enseignements robustes sur les origines de la croissance. Ses enseignements sont précisés et complétés par les théories de la croissance endogène, lesquelles permettent de mieux comprendre les interdépendances entre les facteurs primaires (de production et de croissance) et le progrès technique, au sens le plus large. En effet, dans le cadre des théories de la croissance endogène, nous verrons que l'accumulation du capital va incorporer du progrès technique, lequel va moderniser le stock de capital fixe et donc entraîner des améliorations qualitatives du capital productif.

8.6.1 Les différentes formes de capital et les théories de la croissance endogène

Les théories de la croissance endogène permettent d'insister sur l'importance de l'accumulation du capital sous toutes ses formes. Les modèles de la croissance endogène se sont développés progressivement à partir du début des années 1980, afin de combler les vides explicatifs de la théorie de la croissance de Robert Solow. Le point crucial est que le

processus de croissance trouve ses origines dans l'économie elle-même (endogène), via les décisions des agents économiques et les opérations qui en découlent. De surcroît, les effets sont cumulatifs, dans la mesure où les décisions d'investissement et d'innovation présentes dépendent de celles du passé et que, donc, la croissance économique présente est conditionnée par la croissance antérieure.

– C'est **Paul Romer** qui, en 1986, a proposé un premier modèle fondé sur la logique de l'**apprentissage par la pratique** (*learning by doing*), mis en évidence par Kenneth Arrow, qui permet d'expliquer les rendements croissants qui accompagnent l'accumulation de connaissances et de compétences. Il observe des rendements marginaux non décroissants sur les facteurs accumulables (capital physique, capital humain) et l'explique par le savoir-faire accumulé, des externalités pécuniaires et les facteurs de production publics.

– En 1990, dans un second modèle, **Paul Romer** insiste sur le rôle de la **recherche-développement** (R&D) dans le processus de progrès technologique. La recherche permet d'**accumuler des connaissances**, lesquelles vont se diffuser librement, car la connaissance est un bien collectif accessible à tous ; et, du fait des externalités positives qui en découlent, les pouvoirs publics ont intérêt à inciter le secteur privé à développer ces activités en les protégeant (brevets) ou en les subventionnant pour soutenir durablement une croissance soutenue.

– En 1990 toujours, **Robert Barro** présente un modèle fondé sur l'impact majeur de l'**accumulation de capital public** (infrastructures publiques telles que les routes, chemins de fer, ports, hôpitaux...) sur l'accumulation de capital privé, et donc sur la croissance.

– En 1998, **Robert Lucas** développe un modèle fondé sur l'**accumulation de capital humain** et ses effets externes positifs (externalités positives). Le capital humain désigne un stock de connaissances, de savoir-faire et de

compétences que les individus peuvent intérioriser et valoriser dans l'économie. Chaque individu, entreprise ou institution peut investir en capital humain et en tirer un avantage durable ex post, car le capital humain a une productivité marginale constante, il facilite une accumulation ultérieure et il est à l'origine d'externalités positives, de sorte que l'accumulation par les uns favorise et facilite l'accumulation par les autres. Les pouvoirs publics ont donc intérêt à améliorer les politiques de formation initiale, professionnelle et continue pour promouvoir une croissance durablement soutenue.

– En 1992 et 1998, **Philippe Aghion et Peter Howitt** ont tenté de formaliser un **modèle néo-schumpétérien de l'innovation**. Ce modèle, né de la théorie de l'organisation industrielle moderne, présente les effets de **l'accumulation de capital technologique**. Il se fonde sur la dimension destruction-créatrice du processus d'innovation. L'innovation dépend des dépenses privées et du stock d'innovation réalisé précédemment. Le modèle permet de distinguer les innovations qui permettent de rattraper la **frontière technologique** (imitation), de celles qui permettent de la repousser. Les auteurs insistent sur le rôle des innovations en matière de différenciation verticale. Ainsi, les politiques économiques seront différentes selon que le pays se trouve proche de la frontière technologique, ou qu'il est éloigné de celle-ci. Dans le premier cas, les politiques de recherche et développement, d'innovation, de concurrence et de mobilité (des biens et des facteurs) seront indispensables et, dans le second, pour une économie en phase de rattrapage, l'accumulation de capital et le transfert de technologie seront à privilégier dans un premier temps.

Ces approches endogènes de la croissance économique permettent de montrer que l'investissement et l'accumulation de capital matériel, immatériel, technologique, humain et public améliorent la capacité productive et permettent d'élever la productivité du travail.

Par exemple, le progrès des connaissances va avoir une incidence sur la formation du capital humain, lequel permet de faire progresser les compétences et la capacité à utiliser de manière optimale le capital matériel. Il en découle une augmentation de la productivité du travail et de la productivité globale des facteurs. De manière générale, cette approche, en termes de capital humain, permet d'insister sur le rôle majeur de l'éducation dans la croissance à long terme. On peut encore ajouter les dépenses en recherche et développement qui conduisent à accumuler du capital immatériel dont les effets vont se diffuser via différents canaux complémentaires. De manière générale, les dépenses de recherche et développement correspondent à de l'investissement immatériel. Cette accumulation de capital immatériel va contribuer à améliorer l'efficacité de la combinaison productive et à accroître durablement la production et la croissance potentielle. Enfin, les investissements publics se traduisent par une hausse quantitative et qualitative (modernisation) du stock de capital public composé, pour l'essentiel, des infrastructures de transport, de communication, des investissements dans la recherche, l'éducation ou la santé... En fait, on prend la mesure ici des **interdépendances entre les investissements publics et privés**, dans la mesure où les premiers vont conditionner et favoriser les seconds, et réciproquement. L'émergence des PPP, **partenariats publics privés**, par exemple, en est une illustration limpide.

On peut donc retenir que l'accumulation de ces différentes formes complémentaires de capital va conduire à l'amélioration de l'efficacité de la combinaison productive et va permettre d'entretenir une **dynamique endogène de la croissance économique en favorisant une accélération de la croissance de la PGF et de la productivité du travail**. Autrement dit, les investissements d'aujourd'hui vont favoriser la croissance de demain et la capacité à investir et accumuler du capital après-demain...

8.6.2 *Le rôle central des institutions*

Les **institutions économiques** encadrent le fonctionnement des marchés de biens et services, le marché du travail, les marchés de capitaux (financiers)... via les droits de propriété, les conditions de concurrence, la protection par les brevets, la stabilité politique (démocratie et bonne gouvernance).

Le concept d'**institution** a d'abord été développé par **Douglas North** et **William Fogel** (prix Nobel d'économie 1993). Pour eux, les institutions désignent « **les règles du jeu dans la société ou, plus formellement, les contraintes créées par les hommes qui régissent les interactions entre les hommes** ». Ils considèrent que posséder de « bonnes » institutions est essentiel pour la croissance et pour le développement, mais les institutions adéquates dépendent du niveau de développement. Le lien entre croissance (développement) et institutions est difficile à étudier empiriquement, car le lien de causalité est circulaire.

Il s'agit bien plus d'une interaction que d'une causalité à sens unique. Par exemple, des institutions adaptées sont nécessaires pour que les marchés financiers permettent la rencontre de l'épargne et des investissements, et donc le financement d'une croissance soutenue et durable. Ce n'est pas le cas en Chine depuis les années 2000, où l'inefficacité des intermédiaires financiers et les imperfections des marchés réduisent l'efficacité de l'appariement entre capacité et besoins de financement et, *in fine*, la PGF et la croissance potentielle.

Le rôle des institutions, en général, et des **droits de propriété**, en particulier, apparaît crucial dans la mesure où une économie capitaliste de marché, dans laquelle on souhaite promouvoir le libre-échange, l'initiative individuelle, a besoin de règles multiples pour créer un environnement favorable à l'activité économique.

Selon Dani Rodrik et Arvind Subramanian, **les institutions** contribuent à la croissance économique quand elles sont « **créatrices de marché** », car en leur absence les marchés

n'existent pas ou fonctionnent très mal. Ce faisant, par les incitations qu'elles créent, elles stimulent l'esprit d'entreprise, l'investissement, l'innovation : les institutions et mécanismes de marché favorisent alors la croissance économique à long terme. De surcroît, selon Dani Rodrik, « *un cadre réglementaire et un système judiciaire qui permettent le respect des droits de propriété et offrent un avantage aux inventeurs – par exemple via le brevet – assurent aux entrepreneurs efficaces qu'ils conserveront leur profit et les incitent à innover. Cependant, cette protection est également susceptible de freiner la diffusion des innovations* », Rodrik, D. et Subramanian, A., « La primauté des institutions » (Finance & Développement, juin 2003).

Dani Rodrik et Arvind Subramanian ont proposé **une typologie** simple et claire **des institutions** favorables à la croissance économique. Leur absence est donc préjudiciable à la croissance. Au-delà des droits de propriété, pour qu'une économie capitaliste de marché fonctionne efficacement, il est indispensable de mettre en place « **trois autres types d'institutions** pour tout à la fois soutenir la dynamique de croissance, renforcer la capacité de résistance aux chocs, faciliter une répartition des charges socialement acceptable en cas de chocs. »

Les réglementations des marchés permettent de gérer et de contrôler les externalités (positives et négatives), les économies d'échelle, l'imperfection de l'information. Des institutions solides permettent la stabilisation de ces marchés, afin de maintenir la stabilité des prix, la confiance et les échanges. Nous le développerons dans d'autres chapitres, mais l'instabilité macroéconomique et financière nuit aux anticipations et à l'avenir des économies.

Enfin, les institutions doivent permettre de légitimer les marchés et le cortège d'inégalités et d'injustices qui découle inexorablement de leurs fonctionnements. L'existence de systèmes de protection sociale, de mécanismes de redistribution des revenus et d'instances de régulation des conflits sociaux est indispensable.

Sur l'importance des institutions dans le fonctionnement des économies, les développements du chapitre 7 (sur les mécanismes de marché) et du chapitre 6 (sur la monnaie) sont incontournables.

L'INSTABILITÉ DE LA CROISSANCE ÉCONOMIQUE À COURT TERME

« Les crises ne sont jamais que des solutions violentes et momentanées des contradictions existantes, de violentes éruptions qui rétablissent pour un instant l'équilibre rompu. »

- Karl Marx, 1867

« Le vrai remède au cycle économique ne consiste pas à supprimer les booms et à maintenir en permanence une semi-dépression, mais à supprimer les dépressions et à maintenir en permanence une situation voisine du boom. »

- John Maynard Keynes, 1936

« Une fois la crise déclarée, nous en pouvons rien faire pour en sortir avant son terme naturel. »

- Friedrich Von Hayek, 1966

« Une expédition de martiens arrivant sur terre après avoir lu cette littérature se serait attendue à ne voir que les épaves d'un capitalisme qui aurait de lui-même éclaté en morceaux depuis longtemps. À vrai dire, l'histoire économique était marquée par des fluctuations aussi bien que par la croissance, mais la plupart des cycles semblaient être auto-correcteurs. La croissance soutenue, quoique perturbée, n'était pas une rareté. »

- Robert Solow, 1987

SOMMAIRE

[9.1 L'instabilité de la croissance](#)

[9.2 Fluctuations et cycles](#)

[9.3 Observation des fluctuations économiques et croissance tendancielle](#)

[9.4 Principes de base pour l'analyse des fluctuations économiques](#)

[9.5 Les principales explications des fluctuations de la croissance économique](#)

[9.6 Chocs de demande et fluctuations](#)

[9.7 Chocs d'offre et fluctuations](#)

[9.8 Le cycle du crédit et ses effets cumulatifs sur les fluctuations](#)

[9.9 L'investissement : entre chocs d'offre et chocs de demande](#)

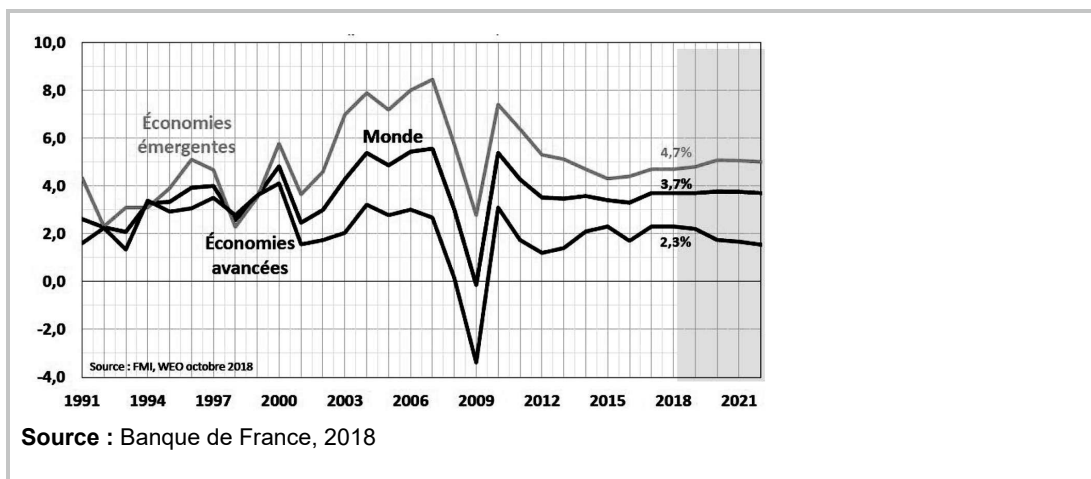
[9.10 Le rôle des politiques conjoncturelles et des politiques structurelles](#)

9.1 L'INSTABILITÉ DE LA CROISSANCE

Une observation rapide des statistiques, dans un tableau ou un graphique, nous confirme que la croissance économique est intrinsèquement instable. Pour le dire simplement, le taux de croissance du PIB varie d'un trimestre à l'autre, d'une année à l'autre. C'est d'ailleurs la raison pour laquelle nous pouvons et nous devons calculer le taux de croissance annuel moyen afin d'avoir une idée plus précise du rythme de l'activité économique globale : la croissance tendancielle ou *trend*.

Les inévitables fluctuations de la croissance

Sur ce graphique on peut observer le caractère instable de la croissance mondiale, de la croissance respectivement dans les économies avancées et dans les économies émergentes entre 1991 et 2018. La corrélation entre les taux de croissance des économies ouvertes est le plus souvent positive (les taux de croissance augmentent et baissent à peu près en même temps), car les économies sont interdépendantes du fait des nombreux échanges qu'elles réalisent.



GRAPHIQUE 9.1.

Croissance du PIB mondial (prix constants, en %)

Logiquement, les taux de croissance sont tendanciellement plus élevés dans les économies émergentes, en situation de rattrapage, que dans les économies avancées. La « grande récession » de la fin des années 2000 se traduit par un taux de croissance quasi nul pour le Monde (la moyenne) en 2009, mais ce taux de croissance reste positif pour les économies émergentes (le PIB croît toujours, mais moins rapidement) alors qu'il devient négatif (le PIB baisse) pour les économies avancées.

Intuitivement, on pourrait considérer que cette instabilité est inhérente à l'activité économique et donc tout à fait logique,

dans la mesure où le PIB n'est que le résultat de la combinaison de millions de décisions, prises par des millions d'agents économiques partiellement irrationnels qui échangent (au niveau microéconomique), soumises à la puissance des courants globaux (représentés par les variables macroéconomiques) qui orientent ces décisions et dépendent simultanément de celles-ci.

De surcroît, les forces macroéconomiques vont modifier, en continu, l'environnement dans lequel ces décisions sont prises. Pour le dire autrement, les grands courants de l'océan macroéconomique vont contribuer à donner une orientation aux gouttes d'eau que sont les décisions microéconomiques, lesquelles, en retour et de manière agrégée, vont se traduire par des inflexions, voire des retournements, des sens du courant global. Il serait fort étonnant, fortuit, et ce serait même un pur hasard d'observer une économie dont le taux de croissance serait parfaitement stable ou presque, c'est-à-dire un taux de croissance annuel égal au taux de croissance annuel moyen. Donc, par nature, l'activité économique est instable et son indicateur le plus global, le produit intérieur brut (PIB), est très variable.

L'instabilité est donc la règle et la stabilité est l'exception. Cette instabilité de la croissance se traduit par ce que les économistes nomment des fluctuations économiques, celles-ci pouvant présenter un caractère plus ou moins cyclique. Afin d'y voir plus clair, nous allons définir les concepts de base relatifs à cette instabilité de la croissance. Puis nous présenterons les principaux indicateurs et signes de ces fluctuations, avant d'en observer les évolutions. Ensuite, nous ferons le tour des principales explications avancées et acceptées. Enfin, nous montrerons pourquoi, comment, et dans quelles mesures certaines politiques économiques peuvent avoir pour objectif de lisser ces fluctuations, de limiter cette instabilité.

9.2 FLUCTUATIONS ET CYCLES

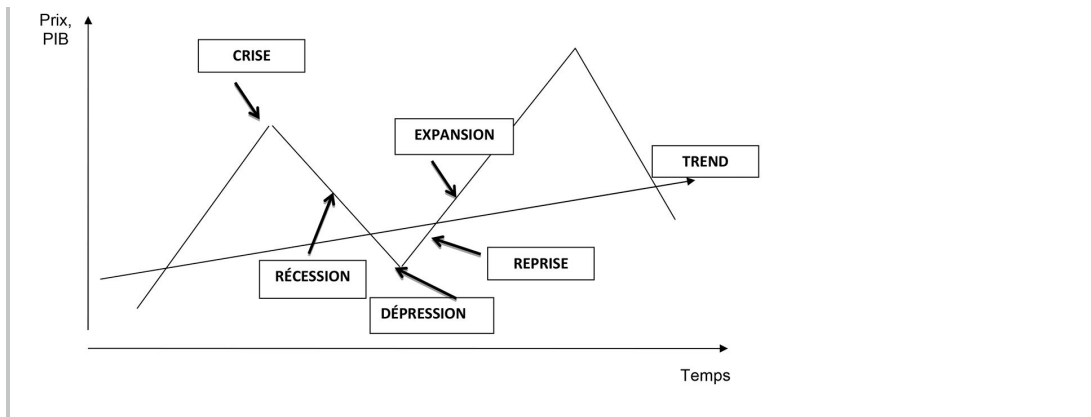
Les **fluctuations économiques** désignent une succession de phases au cours desquelles le taux de croissance du PIB est variable. Les fluctuations rythment de manière plus ou moins régulière l'activité économique. On utilisera le concept de **cycle économique** pour qualifier des fluctuations d'amplitude et de régularité constantes de l'activité économique. Concrètement, lors d'un cycle, on peut observer la succession de phases d'accélération (hausse du taux de croissance), de ralentissement (baisse du taux de croissance), de baisse du PIB (quand le taux de croissance devient négatif).

9.2.1 *Les différentes phases des cycles*

Dans le cadre des théories des fluctuations et des cycles, quand la période de hausse du taux de croissance est de courte durée, on parle d'**expansion**. La période de ralentissement correspond à une **récession**. Quand le PIB baisse, on préfère le terme de **dépression**. Ceci étant, au sens dit « technique » des conjoncturistes qui étudient l'activité économique dans le très court terme (quelques mois à deux ans), la récession désigne une baisse du PIB pendant au moins deux trimestres consécutifs. Et la dépression désigne une période durable de forte baisse du PIB.

Ici, la **crise** et la **reprise** ne sont que des moments de retournement de la tendance de la croissance et non des périodes longues. Ainsi, la crise correspond au moment à partir duquel le taux de croissance se met à baisser. Il reste positif, mais devient plus faible. Autrement dit, le PIB augmente, mais moins rapidement. Un peu plus tard, si cette tendance se poursuit, le PIB lui-même se met à baisser. La reprise intervient quand le taux de croissance du PIB redevient positif et que le PIB se met à augmenter à nouveau. Par exemple, à partir de 2007, on observe bien le ralentissement de la croissance, avant la baisse du PIB en 2008 et 2009, puis une reprise légère en 2010.





GRAPHIQUE 9.2.

Représentation du cycle économique

Les économistes distinguent plusieurs types de cycles. Il est d'usage de différencier les cycles sectoriels des cycles généralisés à l'économie dans son ensemble. En ce qui concerne les cycles sectoriels, on peut retenir les cycles du café, du coton, du bétail, du porc, du bâtiment... Ces cycles concernent surtout les secteurs de production de biens durables, de biens d'équipement... Les secteurs des biens non durables, des biens de consommation ou encore des services sont peu cycliques.

9.2.2 Typologies des cycles

Pour ce qui est des cycles généraux, ils correspondent aux cycles des affaires (business cycles) tels que définis par Burns et Mitchell : « [...] *Les cycles des affaires sont un type de fluctuation que l'on trouve dans l'activité économique agrégée des pays qui organisent leur production au moyen d'entreprises d'affaires (comprenez : privés).* » Le cycle du bâtiment et le cycle majeur sont étroitement liés, selon A. Hansen ("Fiscal policy and business cycles", 1941).

Depuis Joseph Schumpeter (dans son ouvrage *Business Cycles*, 1939), les économistes classent aussi les cycles en trois catégories, selon la durée des phases des fluctuations de l'activité économique et selon l'ampleur de celles-ci.

- Les cycles courts (40 mois environ) ou cycles de **Kitchin**. Ils s'observent à partir des variations des stocks et constituent, surtout aux États-Unis, le mouvement conjoncturel fondamental.
- Les cycles « majeurs » (six à douze ans) ou cycles de **Juglar**. Il s'agit du cycle des affaires ou cycle majeur, selon Alvin Hansen. Ces cycles sont construits essentiellement sur la base de séries statistiques portant sur les prix.
- Les cycles longs (40 à 60 ans) ou cycles de **Kondratieff (ou Kondratiev)**. Ils sont liés à des vagues d'innovations. Ces cycles sont construits essentiellement sur la base de séries statistiques portant sur les productions et les prix.

On peut encore distinguer les cycles selon leur amplitude. Ils peuvent être amortis quand ils convergent vers le trend, explosifs s'ils s'écartent du trend, ou réguliers quand leur amplitude est constante. Notons que, selon la théorie des cycles économiques réels (Real Business Cycles), la distinction entre cycle et tendance n'a pas beaucoup de sens, car, en faisant l'hypothèse d'une tendance stochastique, les chocs économiques sont intégrés à la tendance.

9.3 OBSERVATION DES FLUCTUATIONS ÉCONOMIQUES ET CROISSANCE TENDANCIELLE

Pour observer les fluctuations économiques et éventuellement les cycles, il convient de les **repérer**. Or, comme nous l'avons déjà indiqué, **les faits** ou observations économiques ne sont que des **statistiques construites** et non des données de la nature. Nous avons donc besoin de disposer de nombreuses séries statistiques régulières et fiables et d'en connaître le processus de construction, pour mieux en connaître les limites.

Afin de repérer des fluctuations des variables économiques, nous utilisons **des séries statistiques** (annuelles,

trimestrielles, mensuelles) sur des données qui ont été collectées, ordonnées, et qui portent sur la production en valeur et en volume, sur les prix, sur le niveau du chômage, sur les niveaux des stocks, sur les carnets de commandes des entreprises, etc. En France, les enquêtes de conjoncture sont réalisées par l'Insee (aux États-Unis par le NBER). Nous devons également avoir une **valeur de référence** à l'aune de laquelle on pourra mesurer, ou tout au moins estimer, les **écarts** (accélération, ralentissement, hausse, baisse).

Cette **tendance** ou trend de la croissance du PIB nous est donnée par le **taux de croissance tendanciel**, dont on a vu qu'il nous était donné par le calcul du TCAM. Le TCAM permet de déterminer le trend de la croissance sur une période donnée et de mesurer l'ampleur de la hausse du PIB en moyenne chaque année sur la période. En tant que moyenne, le TCAM est donc constant. Cette vitesse moyenne de la croissance du PIB est très utile d'une part pour comprendre le sens des évolutions longues, faire apparaître des ruptures de tendance et donner du sens à la notion de cycle (lequel ne peut être perçu que par rapport à son trend), d'autre part pour envisager les mesures de politique économique adaptées. Il est crucial de savoir distinguer l'évolution cyclique du trend de la croissance.

9.4 PRINCIPES DE BASE POUR L'ANALYSE DES FLUCTUATIONS ÉCONOMIQUES

Le point de départ de l'analyse des fluctuations sera la distinction entre le PIB effectif (observé, car mesuré) et le PIB potentiel (estimé) d'une part, et la croissance effective et la croissance potentielle d'autre part.

9.4.1 *PIB effectif, PIB potentiel et output gap*

Le **PIB effectif** correspond à la quantité de biens et services produits par les facteurs de production utilisés, quantitativement et qualitativement, à un moment donné. Toute sous-utilisation de ceux-ci constitue un gaspillage de

ressources productives et une réduction de la capacité à satisfaire les besoins des agents économiques. Ce **gaspillage** de ressources productives va se traduire par un écart de production (**output gap**) avec le PIB potentiel estimé au même moment. Le **PIB potentiel** correspond à la production maximum possible, quantitativement et qualitativement. Il désigne encore le niveau maximal de production durablement soutenable sans accélération de l'inflation.

L'écart entre **PIB effectif** et **PIB potentiel** correspond à l'« **output gap** » ou **écart de production**. Cet écart permet de connaître l'intensité de la récession, ou dépression, à court terme. L'objectif est donc de rapprocher le PIB de son niveau potentiel. Pour réduire cet output gap, il convient de faire croître le PIB effectif en accélérant la croissance effective.

Attention : Il ne faut pas confondre la **croissance tendancielle**, qui est une moyenne, et la **croissance potentielle**. Le trend peut être durablement inférieur au potentiel, mais, à long terme, il peut avoir un impact sur le potentiel et donc en devenir une approximation.

9.4.2 *Croissance effective, croissance potentielle*

La **croissance effective** est mesurée à partir du taux de croissance du PIB effectif. La **croissance potentielle** est estimée à partir du taux de croissance du PIB potentiel. Lorsque la croissance effective est supérieure à la croissance potentielle, il y a « surchauffe » et on risque d'observer des tensions inflationnistes. Lorsque la croissance effective est inférieure à la croissance potentielle, il y a gaspillages et on peut observer une hausse du chômage.

Quand la croissance effective est en dessous de son potentiel, le niveau de PIB effectif s'éloigne du potentiel et les gaspillages de ressources productives augmentent. À la longue, le PIB potentiel peut diminuer en raison des pertes de capacité productives induites. Quand la croissance effective devient plus élevée que la croissance potentielle, les tensions inflationnistes peuvent survenir si la situation devient durable,

et si la demande globale augmente durablement plus vite que l'offre et la capacité de production.

9.4.3 *Output-gap et politiques économiques*

9.4.3.1 Output-gap, pertes de production et de croissance

Dans ce cas de figure, il conviendra d'établir un diagnostic juste afin d'établir une ordonnance adaptée. **L'output gap** résulte-t-il d'une croissance effective inférieure à la croissance potentielle ou d'une croissance potentielle trop faible ?

Dans le premier cas, on utilisera les **politiques économiques conjoncturelles**, car ce sont des facteurs de demande qui déterminent le niveau et la croissance du PIB effectif et, dans le second, on mettra en œuvre des **politiques économiques structurelles**, car ce sont des facteurs d'offre qui déterminent le niveau et la croissance du PIB potentiel.

9.4.3.2 Output gap : politiques de l'offre ou politiques de la demande

On retrouvera ici la **distinction entre politiques de la demande et politiques de l'offre**, même si elle n'est pas identique en tout point à la précédente, dans la mesure où, par exemple, certaines politiques conjoncturelles de gestion de la demande globale, quand elles sont menées durablement, ont des effets structurels. Pensons simplement aux politiques budgétaires d'austérité (et exagérément restrictives en période de récession généralisée) qui, en dégradant durablement la demande globale, peuvent nuire à l'investissement productif, à l'accumulation de capital humain, à l'investissement en recherche et développement et à l'innovation et donc, *in fine*, à la croissance potentielle.

9.5 LES PRINCIPALES EXPLICATIONS DES FLUCTUATIONS DE LA CROISSANCE ÉCONOMIQUE

Typologie des chocs économiques : fluctuations et cycles

Comment expliquer l'instabilité de la croissance économique ? Pour expliquer les fluctuations de l'activité économique, et donc l'instabilité de la croissance, on distingue en général les chocs d'offre des chocs de demande, mais également les chocs **endogènes** des chocs **exogènes**. On peut aussi différencier les **chocs réels** des **chocs monétaires et financiers**. Ces différents chocs peuvent être positifs ou négatifs. On accordera une attention particulière au **cycle du crédit**, compte tenu de ses impacts immédiats sur la croissance effective et de ses effets différés sur la croissance potentielle.

9.5.1 Chocs endogènes et exogènes

9.5.1.1 Les chocs exogènes

Les chocs exogènes désignent des chocs qui ont ou auraient pour origine des phénomènes qui se situent en dehors de l'économie, *stricto sensu*. Il faut remonter aux économistes classiques pour retrouver la trace de cette conception du fonctionnement de l'économie.

Les catastrophes naturelles comme les tremblements de terre, les tsunamis et les dérèglements climatiques (sécheresses, pluies diluviennes) peuvent avoir des effets économiques majeurs. De manière générale, on admet que les chocs suivants sont exogènes, car ils ne sont pas le résultat de dysfonctionnements internes à l'économie.

On peut les classer par catégories et distinguer les chocs liés à :

- La nature, comme les catastrophes naturelles ou climatiques et qui se traduisent par des épidémies et famines ;
- L'instabilité politique avec les guerres, révolutions, troubles de la société, insécurité extrême ;

- Les mouvements démographiques liés au solde migratoire et à l'accroissement naturel ;
- L'évolution technique (dans une certaine mesure) ;
- Les changements culturels avec l'acculturation (liée à l'immigration), les religions, les traditions, les modes de vie, les représentations sociales.

9.5.1.2 Les chocs endogènes

Les **chocs endogènes** correspondent à des chocs qui trouvent leur origine dans le fonctionnement de l'économie elle-même, en raison des modifications des décisions des agents économiques et de leurs effets sur leurs opérations économiques, les échanges et les interactions. Les vagues d'innovation et les gains de productivité induits, les mouvements de pessimisme et d'optimisme qui entraînent des mouvements collectifs d'achat ou de vente d'actifs, de biens et services, et qui peuvent provoquer des crises financières et économiques ou un retour de la croissance économique.

La distinction entre chocs exogènes et chocs endogènes est cependant très délicate, voire discutable. Ainsi, les dérèglements climatiques se traduisent depuis des années par des catastrophes naturelles diverses, brutales ou graduelles, qui affectent l'économie. A priori, on se trouve là en présence de chocs exogènes. Mais si on admet que dans le très long terme (un à deux siècles), c'est l'activité économique qui en est en grande partie responsable, il est alors fallacieux de qualifier ces chocs d'exogènes.

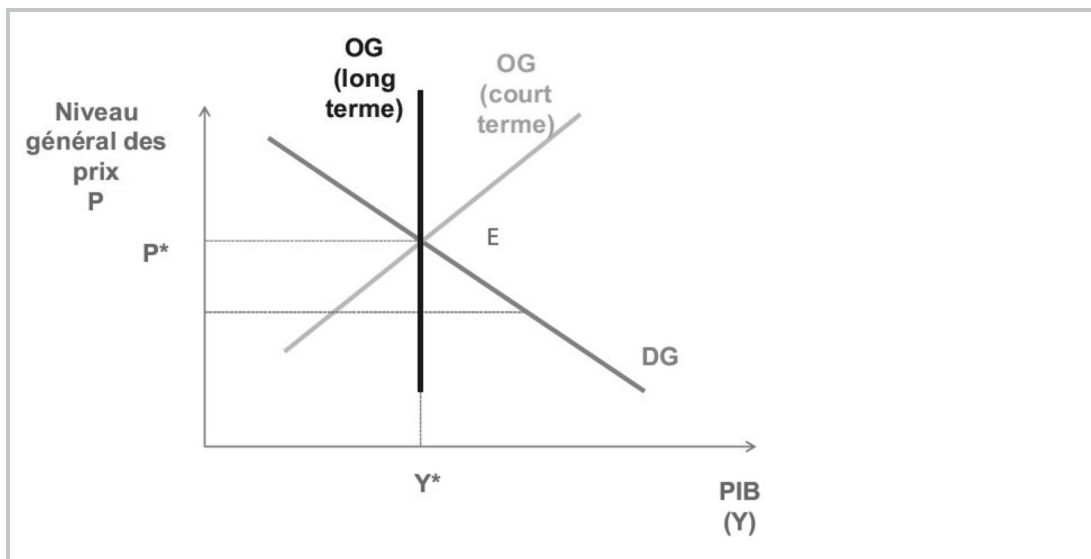
9.5.2 Chocs d'offre et chocs de demande

On opère également une distinction entre chocs d'offre et chocs de demande. Étant donné que l'économie se traduit toujours par des échanges multiples, il semble logique de distinguer les chocs d'offre et les chocs de demande. De surcroît, comme nous observons les évolutions de variables globales, nous considérons essentiellement des chocs sur l'offre globale (OG) et la demande globale (DG).

Pour visualiser les principaux mécanismes, nous nous appuyons sur le **modèle offre globale/demande globale** (ou offre agrégée/demande agrégée). Ce modèle est donc le modèle économique de base approprié pour l'analyse des fluctuations. Nous retrouverons ici avec intérêt l'équilibre entre ressources en biens et services et emplois (utilisations) de ces biens et services pour une économie, fermée ou ouverte.

$$\begin{array}{ccc} \text{Ressources en biens et services} & = & \text{Emplois de ces biens et services} \\ \mathbf{Y + M} & & \mathbf{C + I + G + X} \\ \mathbf{OG} & & \mathbf{DG} \end{array}$$

Dans le cadre de ce modèle, les variations du PIB effectif s'expliquent par des variations des composantes de la demande globale. Les variations du PIB potentiel s'expliquent par des variations des déterminants de l'offre globale.



GRAPHIQUE 9.3.
L'offre globale et la demande globale

9.5.2.1 Chocs de demande

- Les **chocs de demande** correspondent aux variations des composantes de la demande globale, lesquelles vont se traduire par une hausse ou une baisse de la demande

globale. Un choc économique va donc affecter la consommation finale des ménages, les décisions d'investissement des entreprises, la gestion des stocks, les exportations nettes... On peut penser aux facteurs de modification des anticipations, des revenus, de la richesse ou à l'utilisation de la politique budgétaire ou monétaire... *In fine*, les anticipations des agents économiques sont modifiées et leurs comportements changent. Il en découlera un déplacement de la courbe de demande globale. Graphiquement, les chocs de demande se traduisent par un déplacement de la courbe de demande globale. S'il s'agit d'un choc positif, la demande globale croît et la courbe de demande globale se déplace vers la droite (sud-est). Dans le cas d'un choc négatif, la demande globale décroît et la courbe de demande globale se déplace vers la gauche (nord-ouest).

Toutes choses égales par ailleurs :

- Lorsque la demande globale est affectée par des chocs positifs, sa hausse peut impulser une phase d'expansion, via une accélération de la croissance effective (hausse du taux de croissance du PIB effectif). L'output gap se réduit et l'économie se rapproche de son potentiel du moment, de sa frontière des possibilités de production. Les gaspillages de ressources productives diminuent.
- Lorsque la demande globale est affectée par des chocs négatifs, sa baisse peut entraîner une phase de récession, via un ralentissement de la croissance effective (baisse du taux de croissance du PIB effectif). L'output gap augmente et l'économie s'éloigne de son potentiel du moment, de sa frontière des possibilités de production. Les gaspillages de ressources productives augmentent.

9.5.2.2 Chocs d'offre

Les **chocs d'offre** désignent des modifications des déterminants (facteurs, inputs) de l'offre globale qui vont avoir un effet sur la capacité de production, les coûts de production, les quantités produites, la productivité des

facteurs de production et de leur combinaison, les gains de productivité... Graphiquement, les chocs d'offre se traduisent par un déplacement de la courbe d'offre globale. On peut penser aux variations du prix des produits de base, des salaires nominaux, de la productivité, des coûts de production de manière générale. S'il s'agit d'un choc positif, l'offre globale croît et la courbe d'offre globale se déplace vers la droite (nord-est). Dans le cas d'un choc négatif, l'offre globale décroît et la courbe d'offre globale se déplace vers la gauche (sud-ouest).

Toutes choses égales par ailleurs :

- Lorsque l'offre globale est affectée par des chocs positifs, sa hausse peut se traduire aussi par une hausse de la croissance potentielle. Autrement dit, le taux de croissance du PIB potentiel augmente.
- Lorsque l'offre globale est affectée par des chocs négatifs, sa baisse peut entraîner une baisse de la croissance potentielle. Autrement dit, le taux de croissance du PIB potentiel baisse.

Ici encore, la distinction n'est pas exclusive, dans la mesure où les interdépendances, entre les offres et les demandes sur les différents marchés et entre les marchés eux-mêmes, peuvent conduire un choc initial sur l'offre à avoir un impact sur la demande et un choc initial de demande à affecter l'offre. Les typologies et classifications permettent de distinguer pour mieux comprendre, mais en aucun cas il ne s'agit de séparer excessivement, au mépris des interdépendances inhérentes au fonctionnement de l'économie.

9.5.3 Chocs réels, chocs monétaires et financiers

On peut distinguer les chocs liés à l'économie réelle des chocs monétaires et financiers, soit les causes monétaires et financières des cycles. Mais ces distinctions n'impliquent pas que ces explications soient exclusives les unes des autres.

Selon les théories de référence, les cycles ont des causes réelles. Ils sont la conséquence de variations des composantes de la demande globale (consommation, investissement, demande externe, dépenses publiques) et des facteurs fondamentaux de l'offre globale (productivité, rentabilité, compétitivité coût et prix, innovation...).

Selon la théorie du cycle économique réel (TCER), ce sont des chocs d'offre correspondant essentiellement à des modifications aléatoires de la technologie qui induisent des variations de la productivité du travail et provoquent des variations cycliques de l'activité. Mais la grande nouveauté, avec la TCER, est qu'il n'y pas de variation autour de la tendance, car c'est la tendance elle-même qui devient variable. Les fluctuations réelles peuvent correspondre à une évolution aléatoire de la tendance elle-même.

Selon les théories monétaires et financières des cycles, ceux-ci sont la conséquence des variations de la quantité de monnaie en circulation, des variations du crédit qui entraînent des comportements excessifs en matière d'investissement et de placements divers et qui favorisent le « boom » économique. La structure productive et celle de la consommation peuvent être modifiées également. In fine, les phases d'expansion dopées par la monnaie et les effets financiers se retournent brutalement pour entraîner un « krach », une récession, voire une dépression.

9.6 CHOCS DE DEMANDE ET FLUCTUATIONS

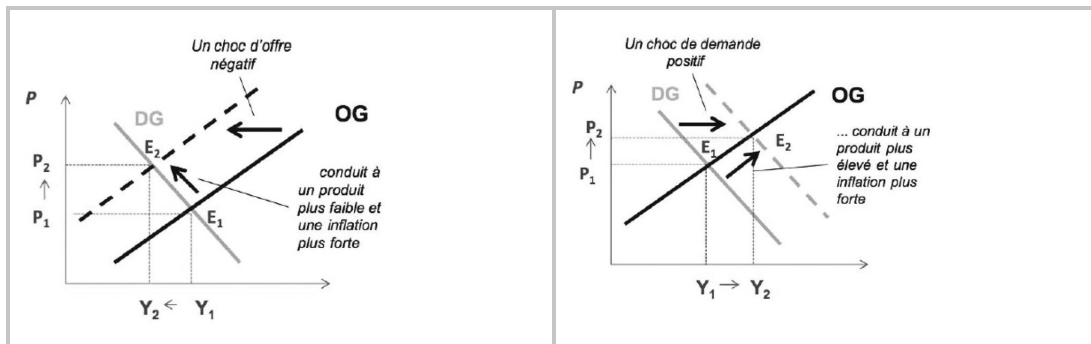
9.6.1 *Les chocs de demande négatifs*

Un choc négatif de demande entraîne une baisse de la croissance, voire de la production globale. À partir de 2008, les économies des États-Unis et de la zone euro ont subi des chocs de demande négatifs. Aux États-Unis, la chute des prix des actifs immobiliers s'est traduite par une baisse de la richesse des ménages, une hausse de leur taux

d'endettement, une hausse des défauts et une restriction du crédit, tout cela entraînant une baisse de la consommation et de l'investissement, donc de la croissance économique. La crise de 1929 fut un **choc de demande négatif** sur l'économie américaine. Une baisse des exportations, toutes choses égales par ailleurs, est un choc de demande négatif pour une économie.

9.6.2 Les chocs de demande positifs

Un choc positif de demande provoque une hausse de la croissance et de la production globale. La réunification de l'Allemagne a pris d'abord la forme d'un grand choc de demande positif, dans la mesure où l'alignement des salaires des Allemands de l'Est sur ceux des Allemands de l'Ouest s'est traduit par une forte hausse de la demande globale. La très forte augmentation des dépenses de l'État au cours de la Seconde Guerre mondiale fut un puissant choc de demande positif pour les économies belligérantes.



GRAPHIQUE 9.4.
Choc d'offre négatif

GRAPHIQUE 9.5.
Choc de demande positif

9.7 CHOCS D'OFFRE ET FLUCTUATIONS

9.7.1 Les chocs d'offre négatifs

Un choc d'offre négatif entraîne une réduction des capacités de production, et donc une baisse de la production elle-même. Ainsi, l'enchaînement tremblement de terre puis tsunami au Japon, début 2011, s'est traduit par

une forte destruction des capacités de production avec la chute de la production d'électricité due à l'arrêt des réacteurs nucléaires, une forte dégradation du commerce extérieur du Japon, avec la baisse de la capacité d'offre, et une forte hausse du prix de l'énergie, avec l'augmentation importante de la demande de gaz naturel par substitution au nucléaire. Les transports ont été également gravement perturbés. *Ipsa facto*, la production a baissé. Une hausse importante du coût des matières premières (chocs pétroliers de 1973 et 1979, par exemple) est un choc d'offre négatif, car elle implique de fortes augmentations de salaires (supérieures aux gains de productivité, comme au cours des années 1970). Une hausse de la fiscalité pour les entreprises peut se traduire par un choc d'offre négatif en réduisant la rentabilité des entreprises et leur capacité d'investissement, et donc en diminuant le stock de capital. Cela rend les entreprises moins productives et moins compétitives. Certaines d'entre elles peuvent être acculées à la faillite. La réunification de l'Allemagne a pris, dans un second temps, la forme d'un fort choc d'offre négatif, dans la mesure où l'alignement des salaires des Allemands de l'Est sur ceux des Allemands de l'Ouest s'est traduit par une forte baisse de compétitivité de l'économie allemande.

9.7.2 Les chocs d'offre positifs

Un choc d'offre positif entraîne une hausse des capacités de production et donc une hausse de la production elle-même. Les États-Unis ont connu un choc d'offre positif entre 1995 et 2000, avec la diffusion à grande échelle des technologies de l'information et de la communication (TIC), l'utilisation d'Internet et d'autres technologies de l'information (téléphonie mobile, multimédia). Ces innovations en grappe ont favorisé des gains de productivité et une baisse des coûts unitaires de production. Ce choc d'offre positif s'est traduit par une baisse des prix de vente des produits, qui facilite leur diffusion en masse chez les consommateurs et donc l'augmentation de la production. Depuis 2009, les États-Unis

profitent d'un nouveau choc d'offre positif, en raison de la baisse des coûts salariaux et de la baisse du prix de l'énergie liée à l'exploitation du gaz de schiste. Cela a entraîné une reprise de l'investissement productif et donc une hausse de la capacité de production de l'industrie, l'accroissement des parts de marché à l'exportation, l'amélioration progressive de la balance commerciale pour les produits manufacturés, la baisse des importations de pétrole et donc de la facture énergétique et une augmentation de l'emploi manufacturier.

9.8 LE CYCLE DU CRÉDIT ET SES EFFETS CUMULATIFS SUR LES FLUCTUATIONS

Le cycle du crédit désigne les effets des mouvements du montant global des crédits accordés aux agents économiques sur l'économie. L'histoire nous montre que le crédit va souvent jouer un rôle d'amplificateur des cycles économiques. Il y aura accélération à la hausse (boom) ou à la baisse (krach) de nombreuses variables macroéconomiques, sous l'effet du crédit. Autrement dit, la hausse du montant global du crédit, et donc de la création de monnaie (voir sixième partie), peut stimuler la demande de biens et de services et la demande d'actifs immobiliers et financiers, dont les cours montent, ce qui est alors source d'effets de richesse positifs, laquelle en retour peut renforcer la hausse du crédit, doper la croissance, mais accroître le risque inflationniste sur les prix des biens ou sur les prix des actifs : c'est le « boom » monétaire, financier et économique.

Mais il vient un moment où les agents économiques sont trop endettés relativement à leurs revenus et à la valeur des actifs qu'ils ont pu acheter à crédit. Pour diverses raisons, les anticipations se retournent et les mouvements des variables monétaires, financières et économiques s'orientent à la baisse. Les demandeurs de crédit, les débiteurs, préfèrent se désendetter.

La demande de crédit baisse. Simultanément, l'offre de crédit baisse avec les défauts auxquels doivent faire face les

créanciers. Tout ce qui s'achetait à crédit jusque-là disparaît. La demande globale baisse et, logiquement, l'offre globale aussi. La croissance ralentit. Dans certains cas, le PIB peut baisser. La récession devient dépression et la déflation menace, car le cercle vicieux déflation/baisse de la demande peut s'enclencher.

Notons qu'en général, la solvabilité du crédit est fondée sur un revenu, plutôt stable. Cela limite la hausse du montant du crédit et les risques d'excès sont faibles. En revanche, lorsque la solvabilité est fondée sur la hausse des prix des actifs qui sont achetés à crédit, alors la situation est potentiellement très dangereuse (cf. crise de l'immobilier et des subprimes aux États-Unis), car, au boom artificiel, aux tensions inflationnistes et à la bulle sur les prix des actifs peuvent se substituer brutalement le krach, la récession, la dépression et les pressions déflationnistes. En effet, dans la première phase, les effets de richesse positifs facilitent et accélèrent l'endettement et, dans la seconde phase, les effets de richesse négatifs provoquent le surendettement, obligent les agents à un désendettement rapide, ce qui provoque un effondrement de la demande globale et de la croissance. Cette baisse de la croissance effective et du PIB effectif peut s'accompagner d'une baisse du PIB potentiel et de la croissance potentielle, en raison des destructions de capacités de production par la baisse du stock de capital productif et la dépréciation du capital humain (voir le point 8.2.5).

H. Minsky explique la phase de boom du crédit, période où les agents s'endettent facilement, car l'optimisme partagé incite à voir l'avenir en rose, par une sorte de « paradoxe de la tranquillité ». En effet, derrière cette tranquillité paradoxale, ce calme apparent, se prépare inévitablement la tempête économique qui viendra renverser le sens des anticipations, transformer l'optimisme béat en pessimisme contagieux et pathologique et conduire les agents à voir l'avenir en noir. Pour I. Fisher, seule la « déflation par la dette » peut permettre au système de se purger complètement des excès

du passé afin de construire un nouvel avenir, moins rose, mais plus réaliste et soutenable. Avec le cycle du crédit, l'histoire grève notre présent et hypothèque notre avenir !

9.9 L'INVESTISSEMENT : ENTRE CHOCS D'OFFRE ET CHOCS DE DEMANDE

L'investissement occupe une place particulière dans le processus de croissance ; facteur d'offre ou moteur de la demande. En effet, comme nous l'avons vu dans le chapitre 8, à long terme, l'investissement se traduit par une accumulation de capital et donc une hausse possible du stock de capital productif, donc de la capacité de production. Il est alors un facteur de la croissance potentielle.

À court terme, il est davantage un moteur de la croissance, une composante de la demande globale. Il agit sur la croissance effective. Avec l'investissement, on oscille entre chocs sur l'offre et choc sur la demande. On peut résumer les effets de l'investissement et de l'accumulation sur la croissance de la manière suivante.

9.9.1 *L'investissement : un choc d'offre*

L'investissement en capital fixe matériel principalement (achat de machines) permet d'augmenter la production en s'appuyant sur les mécanismes suivants :

- **L'investissement de capacité** augmente les capacités de production de l'entreprise. Il permet de produire davantage, il permet une hausse de la production de l'entreprise à condition que les nouvelles capacités soient utilisées.

Si le taux et la durée d'utilisation des capacités de production sont proches du maximum, l'entreprise ne pourra pas répondre à un surcroît durable de demande sans réaliser un investissement permettant d'augmenter sa capacité de production (son stock de capital).

Le taux d'utilisation des capacités de production (machines et équipements) est égal au rapport

entre les capacités de production utilisées et les capacités de production disponibles.
L'existence de goulots de production (ou **goulot d'étranglement**) traduit les difficultés rencontrées par l'appareil de production (ou par une partie de celui-ci) pour répondre à l'accroissement de la demande en raison d'une insuffisance des équipements (nombre de machines insuffisant) et/ou d'une pénurie de main-d'œuvre. Lorsque la demande adressée aux entreprises augmente de manière durable, celles-ci cherchent à ajuster leurs capacités de production à leurs débouchés. Toutefois, le niveau d'utilisation des capacités de production existantes, mesuré par le taux d'utilisation des capacités de production, intervient dans la décision d'investir ; en effet, une entreprise qui souhaite produire plus, mais qui sous-utilise ses capacités productives (certaines machines ne sont pas en service) n'investira pas, mais mettra en service le capital inutilisé. À l'inverse, une entreprise dont les capacités de production sont proches de la saturation (toutes les machines sont en service) devra investir si l'augmentation de la consommation est durable.
Source : Insee

– **L'investissement de productivité** permet d'augmenter la productivité du stock de capital, et plus largement de la combinaison productive, de l'entreprise. Les gains de productivité sont issus du progrès technique qui reste le principal facteur de croissance. En effet, le progrès technique permet de produire plus avec la même quantité de main-d'œuvre ou de produire autant avec moins de main-d'œuvre. Le choix de combinaison productive, de l'intensité capitaliste (capital par travailleur) est un choix qui appartient à l'entreprise. Parmi les entreprises qui réalisent un investissement de productivité, certaines vont augmenter leur production, d'autres non.

– **L'investissement de remplacement** n'augmente ni la capacité de production ni la production de l'entreprise, mais ne les diminue pas non plus. Il est neutre vis-à-vis de la capacité de production et du stock de capital de l'entreprise.

Quand l'investissement s'accompagne d'une hausse du stock de capital productif, il augmente la capacité de production et contribue à la hausse de la production (offre) globale. D'un point de vue macroéconomique, il favorise la croissance économique potentielle.

L'effet accélérateur a été mis en évidence par A. Aftalion en 1909 et J. M. Clark en 1917. Il permet de montrer que les variations de l'investissement sont corrélées positivement avec celles de la demande, mais de manière amplifiée et décalée dans le temps.

L'effet d'accélération désigne l'effet de la variation de la demande sur l'investissement : une variation de la demande va entraîner une variation amplifiée (plus que proportionnelle) de l'investissement. L'hypothèse à la base du modèle d'accélération est qu'il existe un rapport constant entre production et équipement. De surcroît, il n'y a pas de stock : la quantité produite est égale à la quantité demandée. Il y a donc un rapport constant entre la production et le capital nécessaire :

$$v = \frac{\text{K nécessaire}}{\text{Production}} = \text{coefficient de K} \rightarrow v = K / Y$$

Autrement dit, $1/v = Y/K \rightarrow$ la productivité du capital est constante.

Selon ce modèle, une variation de la demande entraîne une variation accélérée de l'investissement. Autrement dit, quand la demande augmente, celle de l'investissement est beaucoup plus rapide. Symétriquement, quand la demande ralentit, l'investissement diminue. Concrètement, l'achat de capital fixe varie de manière irrégulière. Certaines années, compte tenu de la conjoncture, il faut acheter une grande quantité de capital fixe quand d'autres années, un déstockage et une hausse du taux et de la durée d'utilisation des capacités de production suffisent.

9.9.2 *L'investissement : un choc de demande*

L'investissement est, au niveau macroéconomique, une composante de la demande globale. De plus, l'investissement est une opération économique consistant à acheter des biens de production durables (biens d'équipement). De fait, une variation de l'investissement global constitue un choc de demande. L'effet de l'investissement sur la demande globale et sur la croissance transite par deux canaux.

L'investissement est **une composante de la demande globale**. Pour $DG = C + I + G + (X - M)$, toutes choses égales par ailleurs, une hausse de l'investissement se traduit par une hausse de DG et donc de la croissance du PIB à court terme.

Via **l'effet multiplicateur**, une hausse initiale de l'investissement entraîne une hausse plus importante de la production globale.

En effet, un investissement produit des vagues successives de revenus qui provoquent une hausse de la production globale qui est un multiple du montant de l'investissement initial.

Autrement dit, le multiplicateur d'investissement montre comment un accroissement de l'investissement provoque une expansion du PIB.

Par exemple, pour simplifier, dans le cas d'une économie fermée, on a $R = C + S$. Autrement dit, le revenu (production)

est la somme de la consommation et de l'épargne ou épargne = S

La propension moyenne à épargner est $s = S/R$ et la propension moyenne à consommer est $c = C/R$.

Étant donné que $c + s = 1$

on obtient donc $c = 1 - s$ ou $s = 1 - c$

Si dans cette économie fictive, la propension moyenne à épargner est de 20% alors la propension moyenne à consommer est de 80% : On a $c = 0,8$ et $s = 0,2$

Ainsi, sur 100 euros de revenus supplémentaires, 80 sont consommés et 20 sont épargnés.

Dans ce cadre, une augmentation initiale de l'investissement va provoquer des vagues successives de revenus qui augmentent le revenu (production) global.

Le supplément de revenu global est égal à l'investissement initial multiplié par un coefficient k qu'on appelle multiplicateur d'investissement.

L'augmentation du revenu global est proportionnelle à l'investissement initial : $\Delta R = k \Delta I$

k est appelé multiplicateur d'investissement

On peut montrer que $k = 1 / 1 - c = 1 + c^1 + c^2 + c^3 + \dots$

On obtient donc : $\Delta R = (1 / 1 - c) \Delta I$

$$\Delta R = (1 + c^1 + c^2 + c^3 + \dots) \Delta I$$

$$\Delta R = (1 + 0,8 + 0,64 + 0,512 \dots) \Delta I$$

$$\Delta R = 5 \Delta I, \text{ donc } k = 5$$

En effet, $k = 1 / 1 - 0,8 = 1 / 0,2 = 5$

Finalement, le supplément de revenu global est égal à l'investissement initial multiplié par un coefficient $k = 5$ qu'on appelle multiplicateur d'investissement. Dans notre exemple, le supplément de revenu global est égal à l'investissement initial multiplié par 5. Si $\Delta I = 100$, alors $\Delta R = 5 \times 100 = 500$.

L'investissement prend bien la forme d'un choc de demande qui permet dans cet exemple d'accélérer la croissance effective. On pourrait montrer qu'une hausse de la

dépense publique (G dans $DG = C + I + G$) peut présenter les mêmes types d'effets multiplicateurs.

Ces deux types d'effets de l'investissement nous conduisent logiquement à deux catégories de politiques économiques de nature très différentes, mais complémentaires.

9.10 LE RÔLE DES POLITIQUES CONJONCTURELLES ET DES POLITIQUES STRUCTURELLES

Pour faire face aux différents types de chocs économiques qui rendent l'économie très instable, les pouvoirs publics peuvent mettre en œuvre différents types de politiques économiques. Les économistes distinguent usuellement les politiques conjoncturelles, les politiques structurelles, les politiques macroéconomiques, la politique budgétaire, la politique monétaire et la politique de change.

Les politiques macroéconomiques conjoncturelles visent à agir à court et moyen terme sur la situation économique, afin de la stabiliser. Concrètement, les politiques budgétaires des États et les politiques monétaires des banques centrales ont pour objectif la gestion des fluctuations plus ou moins cycliques des économies.

Les **politiques économiques conjoncturelles** désignent l'ensemble des décisions prises et des mesures appliquées par les pouvoirs publics qui visent à un résultat à court terme. Les objectifs sont ceux du carré magique de Nicholas Kaldor (1971). Les principaux instruments de la politique conjoncturelle sont la politique monétaire et la politique budgétaire. Les politiques macroéconomiques conjoncturelles agissent sur les composantes de la demande globale (consommation, investissement), afin de stabiliser la croissance effective.

Les **politiques économiques structurelles** correspondent à l'ensemble des interventions publiques destinées à agir sur les structures de l'économie qui encadrent, organisent et

conditionnent le fonctionnement des économies à long terme. Les politiques structurelles sont des politiques microéconomiques qui agissent sur les déterminants de l'offre globale afin de renforcer la croissance potentielle.

LA CROISSANCE ÉCONOMIQUE ET LA DYNAMIQUE DÉMOGRAPHIQUE

« Il ne faut jamais craindre qu'il y ait trop de sujets, trop de citoyens vu qu'il n'y a de richesse, ni force que d'hommes. » - Jean Bodin, 1577

« 1. Le niveau de la population est nécessairement limité par les moyens de subsistance. 2. La population s'accroît partout où croissent les moyens de subsistance, à moins que des obstacles puissants ne l'arrêtent. » - Thomas R. Malthus, 1803 [1798]

« Il n'y a aucun autre moyen de favoriser la population que d'augmenter la production. » - Jean-Baptiste Say, 1815

« On a parfois la double tentation de considérer les problèmes démographiques comme fondamentaux pour l'avenir de nos sociétés, mais aussi comme très éloignés dans le temps. En fait, ces problèmes sont à nos portes. » - Jacques de Larosière, 2017

SOMMAIRE

[10.1 De la démographie en France à la dynamique démographique mondiale](#)

[10.2 La démographie et l'économie : des questionnements associés](#)

[10.3 Économie et démographie : une histoire de gâteau à partager](#)

[10.4 De la transition démographique aux enjeux économiques du vieillissement démographique](#)

10.5 Évolution des structures démographiques, épargne, accumulation du capital et croissance

10.6 Démographie, vieillissement et migrations

10.1 DE LA DÉMOGRAPHIE EN FRANCE À LA DYNAMIQUE DÉMOGRAPHIQUE MONDIALE

10.1.1 *La population dans le monde en 2016*

Selon l'Insee, « la population mondiale en 2016 est de 7,4 milliards et devrait atteindre 8,5 milliards d'ici 2030, 9,7 milliards en 2050 et 11,2 milliards en 2100 d'après le scénario moyen des projections des Nations unies. La majorité de l'augmentation prévue incombe à un petit nombre de pays africains à la fécondité élevée ou à des pays ayant déjà une grande population. Sur la période 2015-2050, neuf pays devraient concentrer la moitié de la croissance démographique mondiale : l'Inde, le Nigéria, le Pakistan, la République démocratique du Congo, l'Éthiopie, la Tanzanie, les USA, l'Indonésie et l'Ouganda. La Chine et l'Inde sont les pays les plus peuplés du monde avec chacun plus d'un milliard de personnes, représentant respectivement 19% et 18% de la population mondiale. La population de l'Inde devrait dépasser celle de la Chine en 2022. En 2015, parmi les dix pays les plus peuplés du monde, un est en Afrique (Nigéria), cinq en Asie (Bangladesh, Chine, Inde, Indonésie et Pakistan), deux en Amérique latine (Brésil et Mexique), un en Amérique du Nord (États-Unis) et un en Europe (Russie). Parmi eux, le Nigéria, classé au 7^e rang dans le monde, voit sa population augmenter le plus rapidement au point qu'elle pourrait dépasser celle des États-Unis vers 2050, le pays devenant alors le troisième le plus peuplé au monde. En 2050, six pays devraient dépasser les 300 millions d'habitants : la Chine, l'Inde, l'Indonésie, le Nigéria, le Pakistan et les États-Unis ».

10.1.2 *Un aperçu de la dynamique de la population mondiale à très long terme*

Pendant très longtemps, la population humaine ne comptait tout au plus que quelques centaines de milliers d'individus, et ce n'est qu'il y a trente à quarante mille ans, c'est-à-dire très récemment dans l'histoire humaine, que sa population a franchi le seuil d'un million d'habitants. La croissance s'est alors poursuivie jusqu'à atteindre un milliard d'hommes vers 1800 et six milliards en 1999.

La population mondiale s'élève à environ 7,2 milliards d'habitants à la mi-2013 et devrait grimper à 8,1 milliards d'habitants en 2025 et 9,6 milliards en 2050. D'ici 2050, plus de la moitié de la croissance de la population mondiale devrait s'observer en Afrique. Enfin, si la population du reste du monde devrait augmenter d'un peu plus de 10% entre 2013 et 2100, la population européenne devrait diminuer de 14%, avec une fécondité en dessous du seuil de renouvellement des générations dans presque tous les pays européens. Le vieillissement démographique serait la caractéristique majeure des vieux pays développés au cours des décennies à venir.

10.2 LA DÉMOGRAPHIE ET L'ÉCONOMIE : DES QUESTIONNEMENTS ASSOCIÉS

10.2.1 *Des disciplines scientifiques aux problématiques liées*

C'est un médecin hygiéniste français, A. Guillard (1799-1876), qui est à l'origine du terme démographie en 1855, par la réunion de deux termes grecs : *demos* (population) et *graphein*, (décrire). La démographie comme science ne s'autonomise sur le plan institutionnel (à l'Université) qu'à partir de la seconde moitié du xx^e siècle, notamment lorsque l'ONU se dote d'une division population. Le dictionnaire démographique des Nations Unies définit la démographie

comme « *une science ayant pour objet l'étude des populations humaines, et traitant de leur dimension, de leur structure, de leur évolution et de leurs caractères généraux envisagés principalement d'un point de vue quantitatif* ». En France, l'Institut National d'Études Démographiques (INED) est créé en 1945 sous l'impulsion d'A. Sauvy (1898-1990).

Par ailleurs, on peut rappeler que pour P.A. Samuelson, la science économique étudie « *la manière dont les sociétés utilisent les ressources rares pour produire des marchandises ayant une valeur et pour les répartir entre une multitude d'individus.* » En résumé, la science économique étudie les sources de la production et de sa croissance, mais aussi la répartition de celle-ci entre des individus de plus en plus nombreux : les problématiques démographiques et économiques sont donc indiscutablement liées dans leurs analyses.

Dans un rapport du Conseil d'analyse économique (CAE) publié en 2002, J. Pisani-Ferry souligne ainsi que : « *Démographie et économie forment un vieux couple, mais qui prête aux fantasmes* ».

Selon J.M Charpin, ancien directeur de l'Insee, « *la France occupe une position très singulière en Europe en matière démographique. Il s'agit d'une position dynamique, à l'inverse de la plupart des autres domaines, où la tendance est davantage de singulariser la situation française, non par des qualifications positives, mais par des qualifications négatives. Or, ce n'est pas du tout le cas en matière démographique* ».

10.2.2 Un débat classique sur les interactions entre les dynamiques économiques et démographiques : Marx vs Malthus

Dès la fin du XVIII^e siècle, à la suite du théoricien politique Jean Bodin, pour lequel « *il n'y a de richesse ni de force que d'hommes* » (1575), la plupart des observateurs considèrent que la taille de la population détermine la puissance des nations. En publiant son pamphlet en 1798, le pasteur

protestant Thomas Robert Malthus (1766-1834) va rompre avec l'optimisme ambiant en affirmant qu'il n'y a pas de place pour tout le monde « *au grand banquet de la nature* ». Cette rareté vient de la disjonction entre la « *loi d'accroissement* » des hommes (qui suit une progression géométrique) et celle de la nature (qui suit une progression arithmétique). Malthus insiste sur la loi des rendements décroissants qui exerce une contrainte forte sur la production alimentaire. Cette lecture correspond d'ailleurs à une période spécifique à l'Angleterre de la fin du XVIII^e : le pays connaît une croissance de la fécondité alors que les effets de la révolution agricole du début du XVIII^e tendent à s'épuiser. Seuls des freins préventifs (contrainte morale, célibat, etc.) ou des freins plus « destructifs » (maladie, famine, guerre) peuvent alors, selon Malthus, permettre d'atténuer la pression qu'exerce « *la tendance perpétuelle de la race humaine à s'accroître au-delà de ses moyens de subsistance* ». De même, les lois protectrices des pauvres (*poor laws*) doivent être abandonnées, car, en incitant les pauvres à avoir des enfants, elles participent à l'entretien de cette pauvreté. Ce faisant, Malthus va aussi fixer les termes du débat entre démographes et économistes pour les deux siècles suivants en explicitant l'interaction de la population et des ressources tout en lui donnant un rôle déterminant dans la dynamique du système social.

Malthus met en effet en évidence une interdépendance :

- 1) la croissance de la population dépend du niveau de vie puisqu'elle est bornée par la croissance des subsistances ;
- 2) le niveau de vie dépend de l'effectif de la population.

Cette causalité circulaire condamne le niveau de vie des populations à fluctuer autour d'une tendance stationnaire sur le long terme. À partir du XVIII^e siècle, il est vrai que la Grande-Bretagne, puis l'ensemble des pays connaissent une série de bouleversements démographiques et économiques qui conduisent à une croissance démographique jamais

démentie depuis sur le plan mondial, avec ce que les experts démographiques appellent la transition démographique.

Les liens entre la dynamique démographique et la croissance économique sont ainsi l'objet de controverses récurrentes très importantes dans l'histoire de l'analyse économique.

Contestant la thèse de la surpopulation absolue, Karl Marx (1818-1883) souligne que la surpopulation n'a rien de naturel : la surpopulation est relative au mode de production capitaliste. Dans les faits, la croissance démographique qu'observe Malthus est avant tout liée au progrès que constitue le recul de la mortalité et non à la fécondité exubérante des plus modestes. La révolution agricole anglaise atteste de la capacité des populations à contrecarrer les rendements décroissants agricoles par le progrès technique. Par ailleurs, le courant populationniste s'oppose au pessimisme malthusien, et considère que la croissance démographique peut être favorable à la croissance économique. L'économiste et démographe A. Sauvy a par exemple insisté sur les bienfaits de la croissance démographique. Une forte densité de population permet de rentabiliser des investissements lourds que seules les nations très peuplées peuvent entreprendre. La *railroadization* américaine ne débute que lorsque le coût par habitant des routes et des chemins de fer le permet. En outre, la croissance démographique induit une population jeune, moins réfractaire au changement et aux idées nouvelles, ce qui est source d'innovations et donc facteur de croissance. Enfin, la croissance démographique est pour Ester Boserup (1910-1999) à l'origine de la pression créatrice. La forte densité démographique conduit les populations à trouver des solutions innovantes pour accroître les rendements agricoles, comme l'attestent les rendements de l'agriculture hollandaise ou japonaise.

10.2.3 *La permanence du débat dans les sociétés contemporaines*

Démographes et économistes s'attachent ainsi à analyser les causes et les conséquences des évolutions des populations sur le plan économique et plus globalement sur le fonctionnement des sociétés. Le « *malthusianisme* » (politique qui vise, même par la contrainte, à empêcher la population de croître) a connu une fortune diverse depuis deux siècles. Le terme lui-même a été galvaudé puisqu'il désigne aujourd'hui toute politique restrictive, qu'elle concerne la population ou non (on dénoncera par exemple le malthusianisme des stratégies de réduction du temps de travail visant à partager le volume des emplois existants). Malthus lui-même avait déjà largement amendé sa thèse dès 1803 et l'avait progressivement laissée de côté pour se concentrer sur d'autres questions économiques, dont le rôle de la demande dans la croissance. L'histoire démographique a effectivement infirmé certains postulats malthusiens : la croissance démographique est venue d'une baisse de la mortalité et, à l'inverse du monde animal, les populations humaines ont limité leur croissance grâce à une modification des comportements de fécondité. De fait, la population mondiale suit une courbe logistique qui conduit à l'arrêt probable de sa croissance dans la deuxième moitié du XXI^e siècle. De même, on reproche souvent à Malthus d'avoir sous-estimé le rôle des progrès techniques dans l'accroissement de la productivité agricole (à l'instar des « *révolutions vertes* » dans les pays en développement). Pourtant, les analyses malthusiennes retrouvent aujourd'hui un certain écho : les pressions qui s'exercent sur l'environnement et le capital naturel ont redonné une actualité à ces thèses en les décalant vers la problématique du développement durable. La hausse du prix des matières premières agricoles, masquées un temps par la crise financière, montre aussi que la capacité de la terre à nourrir une population même stabilisée est loin d'être assurée, tandis que l'urbanisation rapide dans les pays émergents présente le risque d'une « *bidonvilisation* » du monde et d'une ségrégation urbaine aux conséquences, ajoutées à celle du

changement climatique, potentiellement incontrôlables et dramatiques au XXI^e siècle pour les populations démunies, dans un contexte de tension croissante sur le partage des ressources.

Comme l'écrit l'économiste Daniel Cohen dans son essai *Le monde est clos et le désir infini* (Albin Michel, 2015), « *il se pourrait bien que la croissance économique soit indispensable aux sociétés humaines pour éviter qu'elles ne chutent dans la violence et la colère, même si les contraintes écologiques conduisent à se montrer raisonnable aujourd'hui et accepter son essoufflement inéluctable* ». La croissance démographique pourrait alors conduire les sociétés à changer leurs modes de production et de consommation selon D. Cohen (viser une croissance nettement plus économe en ressources naturelles), en particulier dans les pays avancés, car « *telle est la contradiction fondamentale de la société occidentale, l'équation qu'elle doit résoudre. L'une des voies possibles serait de passer, grâce à un changement dans les mentalités, de la quantité à la qualité, et de hâter l'avènement d'une société où la pacification des relations sociales devrait prendre le pas sur la culture de la concurrence et de l'envie* ». En définitive, même s'il demeure de fortes inégalités entre les pays en matière d'impact sur l'environnement, la croissance démographique pourrait bien constituer au XXI^e siècle un défi pour l'économie mondiale, tout autant qu'un facteur d'aggravation de la crise écologique.

UNE TYPOLOGIE DES INTERACTIONS ENTRE ÉCONOMIE ET DÉMOGRAPHIE

Courant populationniste, courant malthusien et courant neutraliste

Pour Didier Blanchet (Insee et Crest/IPP), nous pouvons distinguer deux courants opposés chez les économistes, et un courant plutôt neutraliste en ce qui concerne ces interactions.

- Le courant populationniste insiste sur les effets positifs de la croissance démographique sur l'économie. De fait, l'inquiétude s'accroît quand la croissance démographique ralentit.
- *A contrario*, le courant malthusien, sur la base de la thèse des rendements décroissants, considère que la croissance démographique est plutôt défavorable à la croissance du niveau de vie.
- Enfin, beaucoup d'économistes ont une approche largement neutraliste de la question. Ils estiment que les changements démographiques, au regard des autres déterminants de l'évolution du niveau de vie, n'ont qu'un rôle secondaire.

10.3 ÉCONOMIE ET DÉMOGRAPHIE : UNE HISTOIRE DE GÂTEAU À PARTAGER

Pour bien comprendre les relations fondamentales qui unissent économie et démographie en tant que disciplines d'une part et entre activité économique et dynamique des populations d'autre part, il convient de se souvenir qu'en situation de rareté (quantité limitée, car abondance = quantité illimitée), les humains doivent résoudre, comme l'a montré le prix Nobel d'économie Paul Anthony Samuelson, trois problèmes économiques fondamentaux : que produire ? Comment produire ? Pour qui produire ?

Les humains doivent donc produire un « gâteau » (nommé PIB) en utilisant les ressources rares à leur disposition. Le nombre d'humains est en quantité limitée (rare). Ils doivent donc faire des choix. Ils sont confrontés au coût d'opportunité et à la loi des rendements marginaux décroissants. Pour produire (fabriquer) ce gâteau, il faut combiner des facteurs de production (le facteur travail, le facteur capital et le progrès technique). La quantité disponible de facteur travail dépend d'abord de la dynamique de la population. De sa combinaison avec le facteur capital et le progrès technique vont dépendre la productivité des facteurs de production considérés isolément et la productivité globale des facteurs (PGF, considérée essentiellement comme une estimation des effets du progrès technique). La combinaison de l'ensemble de ces moyens de production va déterminer le niveau du PIB potentiel et sa croissance.

Ce gâteau-PIB, composé d'une très grande quantité de biens et services, est destiné à satisfaire les besoins humains de tous les agents économiques (individus, organisations...) dans toute leur diversité. Il sera ensuite partagé entre tous les membres de la population. Autrement dit, des plus jeunes (nouveau-nés) aux plus vieux membres d'une population,

chacun recevra en moyenne une part égale à (PIB / PT) . Or, dans une population, tout le monde n'est pas en âge ou en situation de participer à la production du gâteau-PIB qui sera partagé entre tous les membres. Mais chaque membre de la population doit pouvoir en recevoir une part pour vivre. Sauf à considérer que ceux qui ne participent pas à la production doivent être ignorés, éliminés lentement ou brutalement.

10.3.1 *De la population totale... à la population active occupée et à la production à partager*

Signification des initiales pour les catégories de la population :

PT = population totale, PA = population active et PI = population inactive E = emploi = population active occupée (PAO) U = unemployment = chômage PâT = population en âge de travailler PI = PIâT + PInâT avec PIâT (population inactive en âge de travailler) + PInâT (population inactive non en âge de travailler)	On obtient les égalités simples suivantes : $\text{PT} = \text{PA} + \text{PI}$ $\text{PT} = \text{PAO} + \text{U} + \text{PI}$ $\text{PA} = \text{PAO} + \text{U}$ $\text{PAO} = \text{E}$ $\text{PA} = \text{E} + \text{U}$ $\text{PâT} = \text{PA} + \text{PIâT}$ $\text{PI} = \text{PInâT} + \text{PIâT}$

On peut tirer quelques enseignements de ce modèle simple sur les interactions entre démographie et économie.

La production de biens et services destinés à satisfaire les besoins de la PT est le résultat de l'activité exclusive de la PAO, la population qui occupe un emploi. Autrement dit, toutes choses égales par ailleurs, plus le rapport (PAO/PT) sera élevé et plus la part moyenne de gâteau-PIB de chaque habitant sera élevée. *A contrario*, plus ce rapport sera faible et plus le gâteau-PIB produit par une minorité devra être partagé avec une majorité de personnes.

Ici commence à se poser la question de la répartition de la production et donc des revenus (car nous répartissons la production de manière indirecte, via la répartition des revenus) entre différentes catégories de la population : les « jeunes », les « adultes en âge de travailler » qui travaillent - produisent et les « vieux ». Si on admet qu'une partie de la PT n'est pas en âge de travailler - produire, mais doit recevoir une partie du gâteau-PIB, alors nous devons considérer que le rapport PAO/PT n'est pas complètement pertinent et nous devons surtout observer le rapport PAO/PâT. Ce rapport correspond au taux d'emploi de la population en âge de travailler. $TE = PAO/PâT$. On peut aussi calculer le taux d'activité avec $TA = PA/PâT$. Mais ceci ne change finalement que peu de choses au raisonnement suivant. Quand TE baisse, cela veut dire que le pourcentage de personnes non en âge de travailler - produire augmente et donc que ceux qui sont (relativement) de moins en moins nombreux devront être de plus en plus productifs pour compenser le manque de ressources en main-d'œuvre nécessaire pour faire croître la taille du gâteau à partager. Dans le cas contraire, il faudra réduire (absolument ou relativement) la taille moyenne des morceaux de gâteau-PIB que chacun pourra recevoir. C'est toute la question de la dépendance évoquée plus haut.

10.3.2 Capacité de production, production et répartition du « gâteau-PIB »

Ce raisonnement basique nous montre qu'il existe un lien direct et simple entre capacité de production, production et répartition de cette production. Il convient donc de faire des choix sur le partage du gâteau-PIB. Si ceux qui travaillent - produisent reçoivent plus, alors, *ipso facto*, ceux qui ne travaillent pas, pas encore, ou plus du tout, reçoivent moins. On peut penser aux situations des jeunes, des chômeurs et des retraités relativement à celle des actifs occupés qui sont les seuls à produire. Ce qui nous permet de mettre en perspective les questions relatives aux réformes de la redistribution : réformes des retraites, de l'assurance

chômage, des allocations familiales, de l'assurance maladie, de l'éducation nationale... Ces systèmes de protection sociale doivent permettre de réduire les écarts (les inégalités) dans le partage primaire du gâteau. Pour cela, il existe des prélèvements obligatoires, qui permettent de financer des prestations sociales et des biens collectifs afin de corriger les inégalités et de les rendre plus justes (si possible), sans pour autant nuire au potentiel de croissance de l'économie et de la population.

Notons que, compte tenu de la loi des rendements marginaux décroissants des facteurs de production, le progrès technique et le commerce international peuvent transitoirement permettre de repousser le mur des rendements décroissants et de la rareté.

10.3.3 De quelques interactions simples entre dynamiques économiques et démographiques

Si chaque membre d'une population reçoit en moyenne (PIB/PT) et que le PIB augmente moins vite que la PT, on peut soit chercher à accélérer la croissance du PIB, soit réduire la croissance de la PT. De plus, sachant que la taille du PIB dépend du TE (avec $TE = PAO/P\hat{a}T$) alors, toutes choses égales par ailleurs, pour augmenter PIB/PT, on peut chercher à augmenter le PIB en augmentant le TE avec une hausse de TA et baisse de TC, le taux de chômage.

On peut distinguer différents facteurs d'évolution de la population active. La population active est en effet un sous-ensemble de la population en âge de travailler (15-64 ans), qui dépend des naissances et des mouvements migratoires. Le « babyboom » a ainsi eu une incidence forte sur l'évolution de la population active de la période 1975-1982.

Les principales explications des évolutions de la PA et du TA sont les comportements d'activité et les dispositions légales en matière de durée de la vie active. De manière analogue, les principales explications des évolutions de la PAO, du TE

et du TC sont liées aux comportements d'activité et aux dispositions légales sur le marché du travail.

Certains facteurs jouent en faveur de la hausse de la population active (c'est le cas notamment de l'activité des femmes) ; d'autres jouent au contraire en faveur de la baisse du nombre des actifs (âge de la scolarité obligatoire, poursuite d'études supérieures au-delà de cette limite, âge de la retraite).

10.4 DE LA TRANSITION DÉMOGRAPHIQUE AUX ENJEUX ÉCONOMIQUES DU VIEILLISSEMENT DÉMOGRAPHIQUE

Dans ce cadre général, on peut déjà commencer à comprendre pourquoi le vieillissement démographique (VD) pourrait poser de nombreux problèmes économiques, car, à règles inchangées, il va se traduire par une hausse de la part de ceux qui vont recevoir une part du gâteau-PIB sans y contribuer directement par leur participation à la production. Ce VD va non seulement interagir avec la production, mais également en dynamique avec la croissance effective et la croissance potentielle. Autrement dit, l'évolution de la population et de sa composition par âge va modifier le fonctionnement de l'économie, dans la mesure où la population est composée de millions d'agents économiques interdépendants qui font des choix de production, de consommation, d'investissement, de placements financiers différents selon leurs catégories d'âge. Ces choix et leur évolution vont modifier les conditions du présent et de l'avenir.

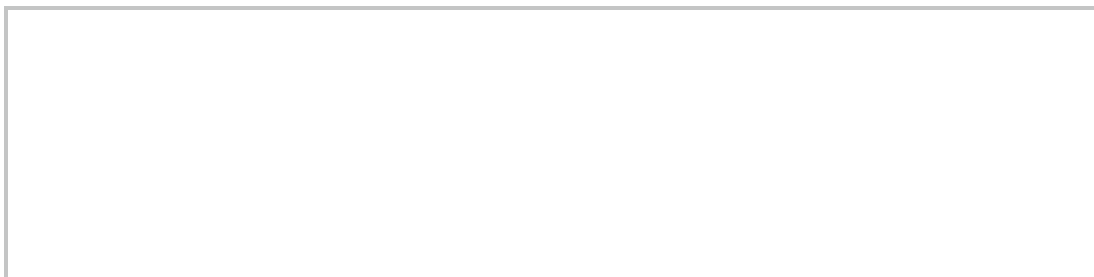
Ils vont aussi modifier les équilibres-déséquilibres sur des marchés interdépendants en économie ouverte. Ces choix vont modifier les règles de partage-répartition du gâteau-PIB à un moment donné (statique) et la dynamique de croissance de ce gâteau à court terme et à long terme. Ils vont donc avoir un impact sur la croissance effective, la croissance

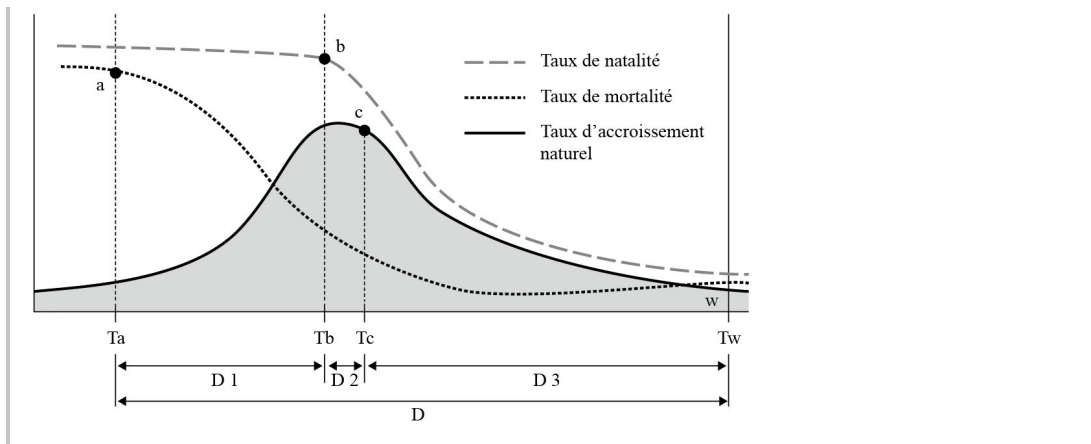
potentielle, mais aussi sur les politiques économiques conjoncturelles et structurelles qui doivent en découler.

10.4.1 *Dynamique démographique et transition démographique (TD)*

Afin de mieux comprendre la dynamique démographique, l'approche par la transition démographique (TD) est utile et éclairante. La TD exprime le passage d'un régime démographique traditionnel (fort taux de natalité et fort taux de mortalité, accroissement démographique faible) à un régime démographique moderne (faible taux de natalité et de mortalité = accroissement démographique faible). Il y a transition démographique dans la mesure où le passage d'un régime à un autre se fait progressivement, sur une longue période. Au cours de la période de TD, il y a d'abord une baisse du taux de mortalité avec un taux de natalité élevé, donc un fort accroissement naturel, la population croît très vite. Puis il y a une baisse du taux de natalité et le retour à une croissance démographique moins explosive.

La transition démographique désigne ainsi pour Jean-Claude Chesnais le passage d'un régime démographique ancien de « quasi-équilibre » haut, à forte natalité et forte mortalité, à un régime démographique moderne de « quasi-équilibre » bas, à faible natalité et faible mortalité. Selon ce modèle explicatif, dans une première phase, la mortalité commence par baisser, provoquant un excédent naturel croissant. Puis, la prise de conscience de la baisse de la mortalité conduit les couples à ajuster la natalité et l'accroissement naturel se poursuit, mais à taux décroissant. À la fin de la transition, natalité et mortalité sont très proches, avec un léger excédent naturel.





GRAPHIQUE 10.1.

Le modèle de la transition démographique

Le nouveau régime démographique se traduit par un vieillissement démographique (VD), lequel peut s'accroître et perdurer si la durée de vie continue de s'allonger et si la fécondité descend en dessous de deux enfants par femme. En France, la transition a été très rapide. Surtout, après la Seconde Guerre mondiale, l'accroissement naturel a été très important, sous l'effet du baby-boom. En Allemagne, ce baby-boom a été moins marqué. Depuis le début des années 2000, en France, l'accroissement naturel reste assez rapide et la population croît encore suffisamment pour contrebalancer les effets du VD lié au « papy-boom ». En Allemagne, le VD se pose avec beaucoup plus de force. Depuis les années 1970, on y observe un déficit naturel en raison d'un faible taux de natalité. Le solde migratoire très positif est censé compenser cette faiblesse pour réduire le VD et ses effets.

Sur ce schéma représentant la transition démographique d'une société on peut distinguer quatre phases. Au cours de la première phase, période qui précède l'industrialisation, les taux de natalité et de mortalité sont élevés et l'écart est faible. Lors de la seconde phase, une amélioration de l'apport nutritionnel, de l'hygiène et des progrès technologiques et médicaux, entraîne une baisse du taux de mortalité et une croissance rapide de la population. Pendant la troisième phase, la fécondité, le nombre d'enfants par femme, diminue,

et logiquement, le taux de natalité baisse. Cela s'explique par la massification et la démocratisation de l'éducation, la croissance des revenus individuels et l'amélioration des moyens de contraception (les femmes ont moins d'enfants et les ont plus tardivement). La croissance démographique ralentit et la société entre dans une quatrième phase au cours de laquelle les taux de natalité et de mortalité sont faibles et proches. Avec une fécondité et une natalité qui baisse et une hausse de l'espérance de vie à la naissance, une dynamique de vieillissement de la population se met inexorablement en place. La dynamique démographique d'un pays développé devient de plus en plus dépendante de l'immigration quand l'accroissement naturel reste durablement faible.

10.4.2 *Le vieillissement démographique (VD)*

Selon les projections de population de l'Insee, publiées dans les *Tableaux de l'économie française 2018*, si les tendances démographiques observées jusqu'ici se maintiennent, la France compterait 76,5 millions d'habitants au 1^{er} janvier 2070. D'après ce scénario dit « central », la quasi-totalité de la hausse de la population d'ici 2070 concernerait les personnes âgées de 65 ans ou plus. L'augmentation serait très forte pour les personnes de 75 ans ou plus. Jusqu'en 2040, la proportion des personnes de 65 ans ou plus progresserait fortement : à cette date, environ un habitant sur quatre aurait 65 ans ou plus. Cette forte augmentation est inéluctable et correspond à l'arrivée dans cette classe d'âge de toutes les générations issues du baby-boom.

En première approche, le vieillissement d'une population désigne la combinaison de l'augmentation de la part des personnes âgées dans une population relativement à celle des personnes plus jeunes et l'allongement de la durée de vie des plus âgés. Autrement dit, il y a de plus en plus de « vieux », relativement (c'est-à-dire en pourcentages), qui vivent de plus en plus longtemps. Les facteurs qui expliquent le vieillissement d'une population sont la diminution de la fécondité, de la natalité, la réduction de la mortalité et

l'allongement de la durée de vie. Ces explications sont liées aux progrès de la science et de la médecine, aux comportements des couples en matière de natalité, aux politiques familiales.

10.4.3 *Deux faits stylisés de la dynamique démographique mondiale*

Au niveau mondial, deux faits stylisés caractérisent la dynamique et les perspectives démographiques mondiales :

- D'une part, après s'être accélérée entre la fin de la Seconde Guerre mondiale et le début des années 2000, la croissance démographique connaît un ralentissement et la population devrait diminuer dans certains pays développés.
- D'autre part, le VD va se traduire dans presque tous les pays par une élévation du ratio de dépendance des personnes âgées. L'un des indicateurs de dépendance utilisés est généralement le ratio des « 65 ans et plus » rapporté aux « 15-65 ans ». Ce ratio va s'élever partout, mais l'ampleur ne sera pas la même dans les différentes régions du monde, car les points de départ, les intensités et les calendriers sont très différents. Dans ces conditions, la structure par âges idéale dans une société serait une structure avec un taux de dépendance minimal et beaucoup d'actifs relativement aux inactifs. Cela permettrait d'avoir une population active importante, donc une capacité de production plus importante et un niveau de production et de moyens financiers plus élevé pour financer la collectivité. Pour la France, on observe un vieillissement démographique, mais le taux de natalité reste assez important. Les effets sur la croissance peuvent *in fine* être positifs avec une hausse de la population active (donc de la quantité de travail disponible), et favoriser la production de richesses, d'autant plus que le taux de chômage pourrait finir par diminuer à long terme...

À partir des données empiriques, on peut mettre en évidence l'effet mécanique du ratio ou taux de dépendance (au sens large, rapport inactifs/actifs, et au sens strict, rapport entre inactifs vieux/actifs). Toutes choses égales par ailleurs, quand le ratio de dépendance augmente, la croissance du produit par tête ralentit et les transferts intergénérationnels sont en hausse pour maintenir les niveaux de vie relatifs des différents groupes d'âge. L'évolution de ce ratio de dépendance nous enseigne que, d'une part, une croissance démographique trop rapide n'est pas favorable, car le nombre de dépendants jeunes augmente et que, d'autre part, la décroissance démographique trop rapide fait croître le nombre de dépendants âgés. Il est donc difficile de trancher sur les effets de la dynamique démographique sur la croissance économique, sauf à verser dans le simplisme.

Mais en tout état de cause selon une majorité d'observateurs, l'essor des systèmes complémentaires d'épargne retraite et de nouvelles réformes avec un recul de l'âge de la retraite semblent inévitables dans les pays avancés. Il faudra également dégager de nouvelles ressources pour financer le risque dépendance, en créant de nouvelles places d'hébergement pour faire face à la demande en croissance, et aider les familles à y faire face.

10.4.4 *Vieillissement démographique, production et capacité de production*

Le vieillissement démographique se traduit certes par l'augmentation de la part de « vieux » dans la population totale, mais, de surcroît, le pourcentage d'adultes d'âge intermédiaire (qui travaillent et produisent) dans la population totale peut diminuer. Il peut y avoir baisse de la part de la population qui produit pour tout le monde. Toutes choses égales par ailleurs, cette population risque de devoir travailler plus longtemps, ou de s'organiser autrement afin d'être plus productive.

Ce vieillissement démographique est, dans une perspective longue, le résultat du baby-boom ayant eu lieu en France après la Seconde Guerre mondiale, entre 1945 et 1973. De plus, la génération issue du baby-boom a eu très peu d'enfants. Dès lors, toutes choses égales par ailleurs, un baby boom engendre un « papy-mamy boom » environ 40 ans plus tard (surtout si les baby-boomers ont peu d'enfants). Dès lors, toutes choses égales par ailleurs, un baby boom engendre un « papy-mamy boom » environ 40 ans plus tard (surtout si les baby-boomers ont peu d'enfants).

Le vieillissement démographique va, c'est une certitude, impacter particulièrement l'Europe (même si le phénomène touche d'autres zones), et provoquer une baisse de la part de sa population en âge de travailler d'ici à 2050. Ce vieillissement générera un poids croissant des retraites, un risque d'affaiblissement de l'innovation et de l'entrepreneuriat, une aversion au risque des épargnants, et des tensions de plus en plus fortes sur les systèmes de santé.

10.5 ÉVOLUTION DES STRUCTURES DÉMOGRAPHIQUES, ÉPARGNE, ACCUMULATION DU CAPITAL ET CROISSANCE

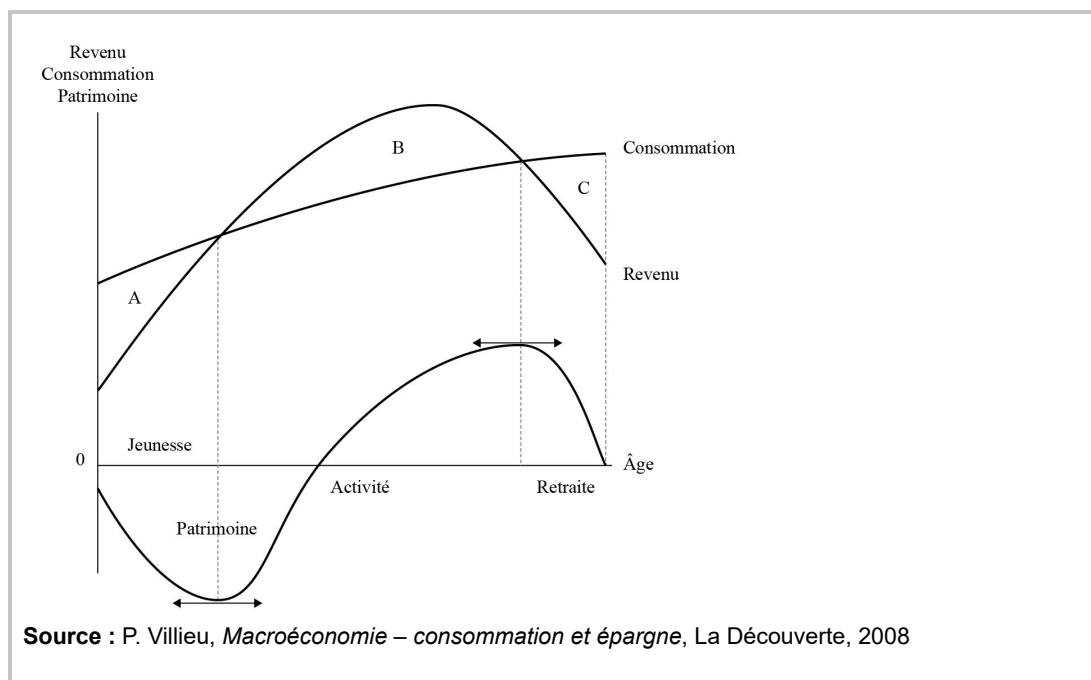
Au-delà de ces effets sur la capacité de production, le niveau de production et la capacité à financer la collectivité, le VD va aussi avoir des effets sur le niveau global de l'épargne et de l'accumulation du capital et donc sur la croissance économique potentielle.

S'il est certain que de nombreux économistes considèrent que le vieillissement démographique constituera un frein à la croissance économique (dans le cadre des hypothèses de la « *stagnation séculaire* » de la croissance), il pourra aussi générer de nouvelles opportunités d'investissement et de création d'emplois (essor des technologies médicales,

domotiques, services et loisirs pour les retraités, *silver economy*).

10.5.1 La théorie du cycle de vie

Pour comprendre les mécanismes sous-jacents, nous disposons de la théorie du cycle de vie de l'épargne (donc de la consommation et de l'investissement), proposée par l'économiste Franco Modigliani. Cette théorie est une représentation simple des choix de consommation et d'épargne en fonction de la structure par âges de la population. Selon la théorie du cycle de vie, le vieillissement démographique va provoquer une baisse du montant global d'épargne dans l'économie. En effet, sous certaines conditions, le taux d'épargne change selon la catégorie d'âge. Autrement dit, le vieillissement démographique doit entraîner une variation à la baisse du montant d'épargne global, du taux global de l'épargne et donc une variation à la baisse de l'accumulation de capital.



GRAPHIQUE 10.2.

Répartition du revenu et du patrimoine durant le cycle de vie

Ainsi, un agent économique rationnel cherche à lisser sa consommation tout au long de sa vie, car il a conscience des évolutions de son niveau de revenu et de patrimoine. Son revenu étant faible quand il est jeune, il s'endette (épargne nette négative et patrimoine net négatif). La hausse de son revenu, au cours de sa vie active, le conduit à se constituer une épargne (épargne nette positive et croissante) et un patrimoine (patrimoine net positif et croissant). Il commence aussi à rembourser ses dettes, surtout en fin de période. Enfin, pendant sa retraite, période d'inactivité de fin de vie, il désépargne (épargne nette positive mais décroissante), finit de rembourser ses dettes et utilise son patrimoine (patrimoine net positif et décroissant) pour satisfaire une partie de sa consommation courante.

Logiquement, le montant global de l'épargne dans l'économie va diminuer, d'autant plus que le VD et ses effets sont puissants. Cette théorie met en évidence un effet d'âge sur les comportements d'épargne, car une population vieillissante épargnera moins.

Selon le modèle du cycle de vie, les consommateurs font varier leur patrimoine net de manière à compenser les variations du revenu. Ils ont un patrimoine net négatif en début de carrière, l'emprunt permettant de consommer davantage que le revenu courant ne le permet (phase A du graphique), ils augmentent ensuite progressivement leur patrimoine pour préparer leur retraite (revenu > consommation dans la phase B), et enregistrent enfin une désépargne à partir de la retraite, c'est-à-dire que le patrimoine net diminue de manière à consommer davantage que ne le permet le revenu courant (phase C).

Le vieillissement de la population, dans les pays développés, devrait donc s'accompagner selon la théorie du cycle de vie d'une évolution à la baisse du taux d'épargne.

Pourtant, les observations empiriques viennent contrarier partiellement les enseignements de cette théorie.

Ainsi, les pays qui ont les taux d'épargne les plus élevés devraient être ceux dont la part des actifs âgés est la plus forte (Allemagne, par exemple). Mais l'Irlande par exemple a un taux d'épargne plus élevé que le Royaume-Uni, alors qu'elle est plus jeune que lui. Les pays les plus vieux devraient avoir des taux d'épargne faibles (baisse du stock de

patrimoine par sa consommation). Or, c'est plutôt l'inverse (à nouveau, il s'agit plus vraisemblablement d'altruisme générationnel).

De plus, cette causalité peut être nuancée parce que l'épargne ne dépend pas seulement des variables démographiques, mais aussi de la confiance des personnes dans les produits d'épargne proposés, des taux d'intérêt, du risque estimé, de la situation économique. De nombreuses incitations peuvent modifier les comportements d'épargne des agents économiques.

Plus largement, il faut souligner que l'impact de la démographie sur le taux d'épargne dépend pour une large part de facteurs politiques et institutionnels : l'évolution de l'offre d'actifs financiers ou les choix de financement de la protection sociale ont une influence déterminante (poids de l'épargne salariale, évolution de la fiscalité, etc.).

Les retraités peuvent épargner afin de léguer un héritage à leurs descendants, mais également pour aider leurs enfants et leurs petits-enfants, en cas de nécessité (solidarité familiale). Une population vieillissante pourrait signifier une élévation au moins temporaire du montant de l'épargne si les personnes âgées souhaitent transmettre un héritage (patrimoine), surtout en période de chômage massif pour les jeunes.

Mais cet effet de structure par âges de la population ne s'arrête pas à une baisse de l'épargne globale. Le vieillissement démographique risque d'impacter le processus d'accumulation mondiale du capital. En effet, pour accumuler du capital (stock de capital), il faut investir, et c'est l'épargne (domestique ou externe) qui permet cet investissement. Cependant, l'accumulation du capital ne dépend pas que de la dynamique démographique : l'ampleur des effets du VD sur l'économie dépendra donc aussi des autres facteurs déterminant le taux d'épargne. Mais si on s'en tient au modèle du cycle de vie, le VD, via ses effets sur l'épargne et

l'accumulation du capital, devrait bien être au xxi^{e} siècle un facteur de ralentissement de la croissance potentielle.

10.5.2 *La relation entre vieillissement démographique (VD), épargne et investissement en économie ouverte*

Nous avons vu qu'en économie ouverte, il existe des échanges de biens et services, mais également de capitaux. Dans ces conditions, quels enseignements pouvons-nous tirer de la théorie du cycle de vie ?

Lorsqu'une économie est dite fermée, il n'y a aucun échange, pas d'importations ni d'exportations de biens et services, de capitaux.

Nous avons la production globale $PG = Y = C + I$ ou le revenu global $RG = Y = C + S$.

L'ensemble du revenu est utilisé soit pour consommer, soit pour épargner.

Il en découle l'identité suivante : $I = Y - C$ et $S = Y - C$, il vient $I = S$.

$I = S$ est l'identité comptable de l'épargne et l'investissement dans l'économie fermée.

Dans une économie fermée, l'investissement (I) dépend strictement de l'épargne (S). Si S augmente, I augmente. Si S diminue, I diminue. Or, selon la théorie du cycle de vie de l'épargne, dans un pays où on observe un vieillissement démographique, on sait que S va diminuer progressivement et les personnes âgées vont utiliser leur épargne pour consommer. Donc, si VD entraîne une baisse de S , alors cela provoque une baisse de I , donc une baisse de l'accumulation du capital et, *in fine*, une baisse de la croissance effective et de la croissance potentielle.

Mais, dans la réalité, la grande majorité des économies sont ouvertes, elles exportent et importent des biens et services. Les résidents et non-résidents échangent des capitaux.

L'identité $I = S$ va changer de nature dans la mesure où une partie de l'épargne domestique pourra financer des investissements à l'étranger et une partie de notre investissement pourra être financée par l'épargne d'un autre pays. Autrement dit, en économie ouverte, la causalité qui va du vieillissement démographique à la baisse du taux d'épargne, et donc à la baisse de l'accumulation du capital, est remise en cause par l'existence de la possibilité de financer l'accumulation du capital par des agents non-résidents. En économie ouverte, le vieillissement démographique peut provoquer une baisse du taux et du montant d'épargne global, mais celle-ci peut être compensée par l'afflux de capitaux extérieurs.

De plus, selon la théorie du cycle de vie, une économie ouverte en excédent extérieur courant et en excédent d'épargne ne devrait pas traverser un vieillissement démographique et disposerait donc d'une population jeune et adulte active qui travaille, produit et épargne. *A contrario*, une économie ouverte en déficit d'épargne disposerait donc d'une population vieillissante qui désépargne. *A contrario*, une économie ouverte en déficit extérieur courant et déficit d'épargne devrait traverser un vieillissement démographique et serait caractérisée par une épargne en baisse.

Quand la Chine est en excédent extérieur, elle est en excédent d'épargne et dispose donc d'une capacité de financement. Si la Chine avait une population globalement jeune, le VD ne serait pas d'actualité. Symétriquement, si les États-Unis sont en déficit extérieur, en déficit d'épargne, il existe un besoin de financement. Ce serait un pays en plein VD. Les agents économiques chinois (publics ou privés) peuvent donc financer, grâce à leur épargne, les agents économiques américains.

Ce n'est pas tout à fait ce que nous observons, dans la mesure où l'excédent d'épargne de la Chine est concomitant à un processus de VD bien engagé. Nous pouvons nuancer les enseignements de ce modèle, car les pays vieillissants n'auront pas forcément un taux d'épargne qui devient faible,

car il baisse. De même, les pays jeunes n'auront pas forcément un taux d'épargne qui augmente et devient élevé.

Dans une certaine mesure, on observe dans la zone euro la même chose, avec une Allemagne vieillissante qui dispose cependant d'un excédent extérieur et d'un excédent d'épargne durablement élevés.

En conclusion, dans le cadre d'une économie ouverte, les conséquences du vieillissement démographique sur l'épargne, l'investissement, l'accumulation du capital et la croissance économique sont moins puissantes, et ce parce que l'investissement et l'accumulation du capital peuvent être financés par d'autres pays. On observe que les échanges de capitaux, comme les échanges de biens et services, peuvent avoir des effets positifs sur la croissance. Mais un problème demeure : faire financer la croissance de la production nationale par des non-résidents revient à s'endetter, à être redevable et donc dépendant de ses créanciers.

10.6 DÉMOGRAPHIE, VIEILLISSEMENT ET MIGRATIONS

10.6.1 *Le vieillissement démographique selon les régions du monde*

Selon la Commission sur la croissance et le développement dirigée par Michael Spence et Robert Solow, la population mondiale vieillit. Il y a deux causes principales à ce vieillissement : une baisse de la fertilité et une importante augmentation de la longévité, excepté dans les pays les plus pauvres où le taux de fertilité demeure élevé, mais où des maladies comme le SIDA ont considérablement réduit la longévité.

10.6.2 *Le vieillissement démographique et la croissance économique*

Le vieillissement démographique aura un impact majeur sur la croissance mondiale et les variables comme l'épargne, l'investissement... Les tranches de population vieillissantes représentent plus de 70% du PIB mondial. Dans ces conditions, la stagnation ou la baisse du nombre d'adultes en âge de travailler combinée à la hausse du nombre de retraités va provoquer une diminution du revenu par habitant. Il y aura moins de personnes pour travailler et produire, mais il y en aura autant pour partager le gâteau production-revenu.

Cependant, il est peu probable que la population « en âge de travailler » reste la même, car, dans de nombreux pays et régions (la majeure partie de l'Europe, l'Amérique du Nord, le Japon et la Chine), le vieillissement démographique menace la solvabilité des régimes de retraite.

Des réformes indispensables vont modifier l'équilibre entre période de vie professionnelle et période de retraite. Cela se traduira nécessairement par un allongement de la période d'activité. Ces changements au niveau du taux de participation et du taux d'emploi pourront permettre d'éviter un ralentissement important de la croissance mondiale.

Si le vieillissement démographique est surtout le problème des vieux pays les plus riches, il affecte de plus en plus des pays comme la Chine.

Symétriquement, de nombreux pays moins développés du monde connaissent le problème inverse avec une dynamique démographique portée par un accroissement naturel important. Les populations y sont jeunes et, dans ces pays ravagés par des maladies comme le SIDA, l'effet anti-vieillessement peut être spectaculaire.

Cependant, dans ces pays, des millions de jeunes vont prochainement et massivement quitter l'école pour entrer sur des marchés du travail qui ne pourront pas leur proposer d'emploi. Sachant que les jeunes sont désavantagés sur le marché du travail par rapport à des travailleurs qui ont de l'expérience, le problème du taux d'emploi de jeunes va s'aggraver.

10.6.3 *Le vieillissement démographique et les migrations internationales*

Au niveau mondial, une augmentation de 3 milliards d'habitants est prévue d'ici 2050. Mais cette croissance démographique va être très inégalement répartie : 100 millions dans les pays riches contre un milliard dans les pays comme l'Inde et la Chine.

Le reste, soit les deux tiers de l'augmentation (environ 2 milliards) de la population mondiale, concernera les pays les moins développés à faible croissance et à faible capacité d'absorption des nouveaux actifs sur le marché du travail.

Il y aura donc un décalage croissant entre la croissance de l'offre de travail et celle de l'offre d'emploi.

Les personnes en âge de travailler vont et devront émigrer des pays où la main-d'œuvre est abondante et les créations d'emplois insuffisantes pour se rendre dans des pays où la main-d'œuvre peut se révéler insuffisante et les marchés du travail plus dynamiques.

Ce type de migration massive devra être supervisé au niveau international pour éviter les abus dans la manière dont est traitée cette main-d'œuvre mobile.

LA SOUTENABILITÉ DE LA CROISSANCE ÉCONOMIQUE

« L'essor industriel entraîne une consommation croissante de charbon, que ni les économies possibles ni les substituts envisageables ne pourront réduire. » - William Stanley Jevons, 1865

« La thermodynamique et la biologie sont les flambeaux indispensables pour éclairer le processus économique [...] la thermodynamique parce qu'elle nous démontre que les ressources naturelles s'épuisent irrévocablement, la biologie parce qu'elle nous révèle la vraie nature du processus économique. » - Nicholas Georgescu-Roegen, 1971

« L'important est de comprendre que le processus économique, à l'impossible autonomie, produit du fait de ses multiples interactions avec la nature des conséquences irréversibles. Nous puisons dans des stocks de ressources naturelles non renouvelables (pétrole, matières premières, etc.) et dégradons ou modifions qualitativement les fonds environnementaux en leur imposant un rythme d'exploitation supérieur à leur capacité de régénérescence (terres agricoles, eau, ressources maritimes, etc.). La loi d'entropie nous rappelle qu'il existe une flèche de temps et que nous laisserons ainsi aux générations futures un patrimoine naturel moindre et sans doute moins adapté à leurs besoins que celui dont nous avons hérité. Plus encore, le fait que l'exploitation des stocks de ressources épuisables libère la "vitesse" économique (la croissance) du rythme écologique contribue à la dégradation des fonds, notamment la biosphère, et peut

susciter des changements irréversibles dans l'évolution des climats. » - Jean-Paul Fitoussi, 2006

SOMMAIRE

[11.1 Les questions écologiques : dégradations de l'environnement et dérèglements climatiques](#)

[11.2 La « tragédie des biens communs »](#)

[11.3 Au-delà de la croissance économique : le développement et le bien-être](#)

[11.4 L'analyse économique de la croissance soutenable et du développement durable](#)

[11.5 Analyse des différentes formes de soutenabilité de la croissance](#)

[11.6 Les instruments des politiques climatiques](#)

11.1 LES QUESTIONS ÉCOLOGIQUES : DÉGRADATIONS DE L'ENVIRONNEMENT ET DÉRÈGLEMENTS CLIMATIQUES

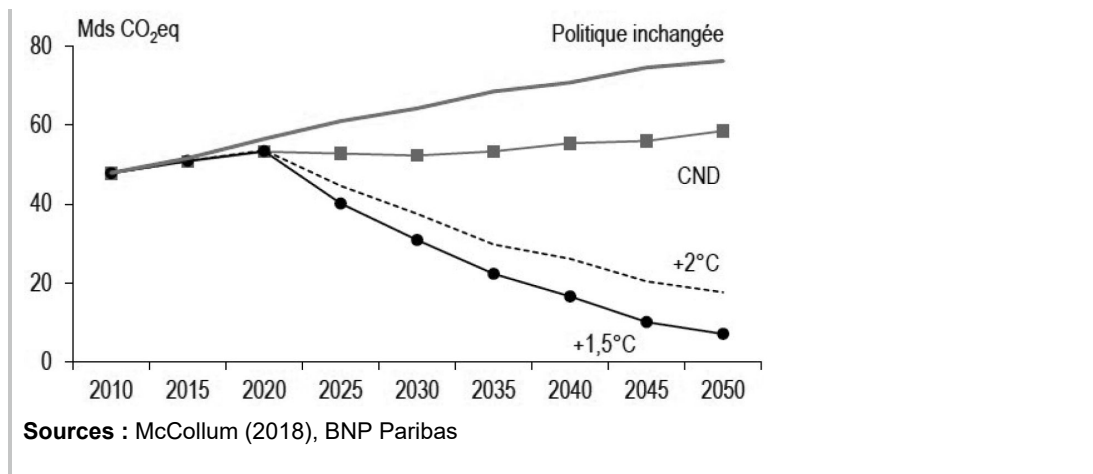
La montée des inquiétudes quant aux dérèglements climatiques, les catastrophes environnementales, la pollution de l'air, des eaux, des terres, les débats récurrents autour de réglementations et taxes diverses sur les activités économiques qui sont sources d'externalités négatives sur notre environnement naturel, la question de l'énergie, etc. sont des signes clairs de la dépendance du système économique par rapport à son environnement. L'économie ne fonctionne pas dans un univers clos, autonome, indépendant de la nature.

11.1.1 *Les conséquences des changements climatiques*

Selon la Convention cadre des Nations unies sur le changement climatique, « *la tendance actuelle du réchauffement prévoit des extinctions d'espèces. Les êtres humains, bien que n'étant pas menacés de cette manière, vont probablement faire face à des difficultés de plus en plus grandes. Les récentes tempêtes, inondations et sécheresses, par exemple, ont tendance à démontrer ce que les modèles d'ordinateurs estiment de plus en plus fréquemment comme des événements météorologiques extrêmes* ».

Le changement climatique peut même devenir une menace pour la sécurité alimentaire mondiale, dans la mesure où les événements météorologiques extrêmes peuvent affecter les rendements agricoles et la stabilité des prix au cours des prochaines décennies.





GRAPHIQUE 11.1.
Émissions de CO₂ selon les différents scénarios

11.1.2 Les interdépendances entre économie, environnement et climat

Les interdépendances irréductibles entre le fonctionnement de l'économie et l'évolution de notre environnement naturel d'une part, et entre analyse économique et analyse écologique d'autre part font de plus en plus consensus. La multiplication des sommets internationaux sur l'environnement, le climat et l'avenir de la planète et de l'être humain en sont une autre preuve. Le processus de croissance économique aboutit, de fait, à une raréfaction croissante, voire à un épuisement des ressources non renouvelables et renouvelables. **La croissance économique sous contrainte de rareté produit une rareté accrue.** Les ressources énergétiques fossiles s'épuisent (pétrole, charbon, gaz, minerais métalliques), les réserves halieutiques aussi, les espèces naturelles en voie de disparition se multiplient, la déforestation s'accélère, l'augmentation de la concentration des gaz à effet de serre GES (dont le dioxyde de carbone, le CO₂) accroît la rareté de l'air pur.

D'un point de vue économique, les défaillances des marchés telles que la présence d'externalités, l'existence de biens collectifs, de biens communs permettent de comprendre les

limites écologiques de la croissance économique. L'absence ou les mauvaises incitations ne conduisent pas les agents économiques à internaliser les conséquences environnementales et climatiques dans leurs calculs (souvent imprécis) préalables à toute décision. De fait, le coût d'opportunité de choix nuisibles à l'environnement et à autrui est très faible.

11.1.3 Croissance économique et destruction des ressources naturelles

Une croissance soutenable serait une croissance qui serait compatible avec la préservation de l'environnement. Si la croissance n'est pas soutenable, c'est que le processus de croissance de la production globale (hausse du PIB) se traduit par de nombreux problèmes écologiques.

En effet, pour produire davantage et accélérer la croissance, il faut utiliser de plus en plus de ressources naturelles et donc puiser dans le stock de ressources naturelles : l'air, l'eau, la flore et la faune... Mais les ressources naturelles sont rares, elles ne sont pas disponibles en quantité illimitée. De nombreuses ressources naturelles ne sont pas renouvelables ou le sont très difficilement.

Concrètement, le processus de croissance se traduit par un prélèvement direct sur le stock de ressources naturelles et par prélèvement indirect via les dégradations qui en découlent.

Tous ces prélèvements se traduisent, *in fine*, par une réduction de notre capacité à satisfaire de nos besoins et une réduction du niveau de satisfaction de ces besoins.

Pour les économistes, on peut commencer à expliquer ces limites écologiques par l'existence d'externalités négatives et des comportements irresponsables. Dans nos économies, rien n'incite suffisamment les agents économiques à ne pas être à l'origine d'externalités négatives nuisibles à l'environnement naturel. Il y a donc un réel manque

d'incitations pour les agents économiques à adopter des comportements écologiquement responsables.

11.1.4 « Accords » pour réduire le réchauffement climatique

L'Accord de Paris plaide en faveur de la limitation du réchauffement climatique sous le seuil de 1,5°C par rapport à l'ère préindustrielle. Toutefois, selon le GIEC, les mesures de limitation actuelles ne permettront pas d'atteindre cet objectif.

« Lors de la Conférence des Parties (COP 21), qui s'est tenue en décembre 2015 à Paris, les 196 parties (195 États, plus l'Union européenne) sont parvenues à la conclusion qu'il fallait limiter le réchauffement climatique à 2°C et poursuivre les efforts visant à le contenir à 1,5°C. La COP 24, qui s'est déroulée en décembre 2018 dans la ville polonaise de Katowice, a confirmé ces objectifs mais sans adopter les mesures nécessaires pour y parvenir » (op. cit.).

La question écologique permet de renouveler la réflexion sur le processus de croissance économique, ses causes et ses conséquences. Pour essayer de mieux comprendre les problèmes écologiques posés par la croissance, on peut s'appuyer sur les différentes catégories de biens que distinguent les économistes (cf. chapitre 7, partie sur les défaillances de marché).

11.2 LA « TRAGÉDIE DES BIENS COMMUNS »

11.2.1 Biens collectifs et biens communs

Les économistes distinguent différentes catégories de biens dans nos économies de marché (cf. chapitre 7). Ici, on se contentera de distinguer des biens collectifs et des biens communs. Les biens collectifs ne peuvent donner lieu à un échange marchand. Ils sont « non rivaux » et « non excluables ». Un bien est qualifié de rival dans la mesure où la consommation par l'un empêche complètement celle de

l'autre. Celui qui paye le prix peut satisfaire son besoin, celui qui ne paye pas, ne le peut pas. Un bien collectif respecte la non-rivalité (la consommation de ce bien ne peut empêcher la consommation de ce même bien par un autre individu) et la non-exclusion (il est impossible d'exclure par les prix l'usage de ce bien) Pour ces biens, les mécanismes de marché sont complètement défaillants. On ne peut exclure par le prix, car tout le monde a intérêt à ce que tout le monde y ait accès !

Les **biens collectifs** les plus purs sont le climat, la qualité de l'air, la pureté de l'eau.

Les **biens communs** sont « rivaux », mais « non excluables ». Les biens communs ne sont pas des biens collectifs. Ainsi, les pâturages communs ne font pas l'objet d'une exclusion par les prix, mais l'herbe mangée par un troupeau de vaches n'est plus disponible pour d'autres. Les poissons pêchés par un pêcheur ne sont plus disponibles pour les autres.

11.2.2 *La tragédie des biens communs selon Garret Hardin*

La nature des biens communs conduira à leur tragédie, selon le biologiste malthusien Garret Hardin, qui voit dans l'accroissement de la population un grand danger et estime que la liberté de procréer conduira inévitablement à la ruine. Dans *The Tragedy of commons*, 1968 (La tragédie des biens communs), il montre comment l'usage collectif de terres communales aboutit, en l'absence de régulation, à la ruine des paysans. On peut dire que, contrairement à la main invisible du marché théorisée par Adam Smith, **la somme des intérêts individuels ne conduit pas à l'intérêt général, mais à la ruine générale**. De surcroît, en l'absence de droits de propriété sur les biens environnementaux, on aboutit à une très mauvaise utilisation et affectation des ressources. De manière générale, selon l'approche de Hardin, l'inexistence de droits de propriété sur les biens environnementaux conduit à la destruction inévitable et complète de l'environnement.

L'existence de ces situations de marché dans une économie capitaliste est favorable à l'émergence et à la multiplication de comportements nuisibles à l'environnement, au climat, au bien-être et à l'aggravation des limites écologiques de notre mode de croissance. Pour lui, la gestion optimale des communs exige soit la privatisation du bien considéré, soit la nationalisation.

11.2.3 *L'approche d'Elinor Ostrom : les biens communs, entre tragédie et coopération*

Jusqu'aux travaux d'Elinor Ostrom, les économistes semblaient avoir oublié la notion de biens communs mise en avant en 1968 par Garrett Hardin dans un article célèbre et incontournable : « La tragédie des communs ».

L'article de Hardin va rester longtemps la référence en matière de gestion des biens communs. Elinor va réexaminer ce qui se passe réellement en matière de gestion des biens communs existants et remettre la question des biens communs au centre des débats et des recherches. Elle va montrer que d'autres formes de gouvernance des biens communs que la privatisation ou la nationalisation sont possibles. Les travaux d'Elinor Ostrom présenteront l'efficacité de la gestion directe des réseaux d'irrigation par les parties prenantes en Californie du Sud, ou les façons dont des copropriétaires peuvent gérer correctement et collectivement les immeubles. Les analyses montreront que la gestion de ressources partagées peut s'appuyer de manière efficiente sur la constitution d'arrangements institutionnels, souvent informels, mais dotés d'une grande capacité opérationnelle en vertu de l'implication des acteurs directement concernés.¹⁴

11.3 AU-DELÀ DE LA CROISSANCE ÉCONOMIQUE : LE DÉVELOPPEMENT ET LE BIEN-ÊTRE¹⁵

11.3.1 Naissance du concept de développement durable

Depuis le début des années 1970, avec le rapport du Club de Rome intitulé « Halte à la croissance », la dimension qualitative progresse relativement à la dimension exclusivement quantitative. En 1972, la Conférence des Nations Unies sur l'environnement, qui se tient à Stockholm, déclare que « *tout homme a droit à un environnement de qualité et il a le devoir de le protéger pour les générations futures* ». C'est en 1987, dans le rapport de la commission des Nations unies présidée par Gro Harlem Brundtland, « Notre avenir à tous », que la notion de développement durable apparaît clairement. La Conférence de Rio de 1992 en donnera une définition devenue célèbre en précisant que le développement durable est un mode développement qui « *répond aux besoins du présent sans compromettre la capacité des générations futures à répondre à leurs propres besoins* ». Depuis « Notre avenir à tous », on définit donc le développement durable comme étant un développement permettant aux générations présentes (GP) de satisfaire leurs besoins sans compromettre la satisfaction des besoins des générations futures (GF). Le développement durable (DD) implique que le développement des uns ne se fasse pas au détriment des autres. Selon Robert Solow (père des théories de la croissance), il y a DD si la capacité des GF à satisfaire leurs besoins est au moins aussi élevée que les GP. Il insiste sur la notion de progrès. Cette exigence posée par Robert Solow remet en question la thèse de la décroissance et de la régression.

11.3.1.1 Le développement durable : un concept développement multidimensionnel

Le développement durable (DD) est donc un développement à trois dimensions :

- La dimension économique est mesurée par la croissance économique, en particulier celle du PIB par tête.

- La dimension sociale dépend du degré d'inégalité, de justice sociale et de cohésion sociale.
- La dimension environnementale est liée aux aspects écologiques du développement.

Il est aujourd'hui admis que le DD ne peut résulter que de l'interaction vertueuse entre ces trois dimensions du développement. Autrement dit, le développement n'est durable qu'à la condition que, dans ces trois dimensions, nous observions des progrès significatifs du bien-être pour chaque membre de la population.

Selon le rapport du Conseil d'analyse économique (CAE) intitulé « Évaluer la performance économique, le bien-être et la soutenabilité » (2011), *« la conceptualisation standard de la soutenabilité englobe trois dimensions principales, à savoir sociale, économique et environnementale :*

- un système socialement durable doit garantir l'équité dans la répartition et l'opportunité, la prestation adéquate des services sociaux comme la santé et l'éducation, l'égalité des sexes et les responsabilités et participations politiques ;*
- un système économiquement durable doit être en mesure de produire des biens et des services de manière continue, afin de maintenir des niveaux soutenables de dette publique et extérieure, et d'éviter des déséquilibres extrêmes entre les différents secteurs ;*
- un système durable sur le plan environnemental doit pouvoir maintenir une base de ressources stables, éviter la surexploitation des ressources renouvelables ou la production de déchets, et ne réduire les ressources non renouvelables que dans la mesure où des investissements sont réalisés dans des substituts adéquats. Cela comprend le maintien de la biodiversité, de la stabilité atmosphérique et d'autres fonctions d'écosystèmes qui ne sont d'ordinaire pas classifiées comme des ressources économiques.*

La satisfaction simultanée de ces trois conditions est fondamentale pour parvenir à la soutenabilité du bien-être. »

11.4 L'ANALYSE ÉCONOMIQUE DE LA CROISSANCE SOUTENABLE ET DU DÉVELOPPEMENT DURABLE

Pour produire, croître, nous développer, améliorer notre bien-être, nous devons utiliser et accumuler quatre formes de capital. Mais, ce faisant, nous réduisons et dégradons le stock de capital naturel. Compte tenu des limites écologiques inhérentes au fonctionnement de nos économies, nous devons nous interroger sur la soutenabilité de notre modèle de croissance, la durabilité de notre modèle de développement.

La croissance économique, le développement et le bien-être des populations résultent de l'interaction de l'utilisation et de l'accumulation de ces quatre formes de capitaux.

11.4.1 *Les formes de capitaux utilisées pour améliorer notre bien-être*

Le bien-être des populations résulte pour simplifier de l'accumulation et de l'interaction de quatre formes typiques de capital. Pour satisfaire les besoins et améliorer le bien-être de chacun et de l'ensemble, nous devons produire des biens et services en utilisant ces différentes formes de capital. Cette analyse porte sur le niveau et l'évolution des stocks des différentes formes de capital afin d'estimer, *in fine*, le degré de développement durable. Les quatre formes de capital, interdépendantes, sont le capital naturel, le capital physique produit, le capital humain et le capital social et institutionnel. Le bien-être dépend de l'accumulation - interaction de quatre formes de capital :

- Le capital **naturel** correspond à l'ensemble des ressources naturelles (air, eau, matières premières...) ;
- Le capital **physique est constitué** des moyens de production produits ;

- Le capital **humain** désigne la quantité et la qualité du travail ;
- Le capital **social et institutionnel** fait référence à l'ensemble des institutions qui encadrent le fonctionnement de l'économie et de la société, à l'ensemble des ressources sociales et culturelles qui favorisent un cercle vertueux.

11.4.2 Évaluer la soutenabilité environnementale

« La soutenabilité environnementale est un des trois aspects, et vraisemblablement le plus apparent, des trois aspects essentiels énoncés dans le débat autour de la notion de développement durable. Ainsi, un système viable sur le plan environnemental ne doit pas gaspiller sa base de ressources. Or, cela implique d'éviter la surexploitation des ressources renouvelables ou des fonctions d'assimilation naturelle de l'environnement. Il convient, par exemple, de vérifier la capacité d'absorption du dioxyde de carbone par les océans ou les forêts et d'empêcher l'épuisement des nappes d'eau lié à une surconsommation d'eau. La soutenabilité environnementale exige également une gestion de l'épuisement des ressources non renouvelables qui soit efficace et équitable entre les générations. Enfin, des systèmes viables sur le plan environnemental requièrent le maintien d'une biodiversité pour vérifier la résilience du système aux chocs. »

Source : « Évaluer la performance économique, le bien-être et la soutenabilité », rapport du Conseil d'analyse économique, 2010, La Documentation française

À l'aune de cette définition de la soutenabilité environnementale, il semble clair que le PIB n'a pas été conçu pour évaluer la soutenabilité de la croissance. Afin de savoir si nous sommes sur le chemin du DD et des progrès du bien-être, on peut essayer d'observer, d'estimer, voire de mesurer les variations des stocks des différentes formes de capital. Cependant, il convient d'abord de voir si le stock

global de capital diminue, stagne ou augmente. En effet, selon cette logique, si ce stock global de capital baisse, alors la capacité d'amélioration des besoins et du bien-être diminue aussi. On sort alors du DD, car les GF auront une plus faible capacité à satisfaire leurs besoins que les GP.

Donc, si le niveau du stock de capital est menacé, des questions majeures se posent : pourquoi le stock global de capital baisse-t-il ? Quel type de capital voit son stock diminuer, dans ces conditions ? La croissance économique, condition nécessaire au développement économique, peut-elle nuire à l'environnement et au stock de capital naturel ? Y a-t-il substituabilité entre formes de capital ?

Notre mode de croissance exige d'accroître notre stock de capital physique et humain pour augmenter la production de biens et services et notre capacité à satisfaire nos besoins. Mais cela peut entraîner une baisse du stock de capital naturel, ce qui pourrait nuire au bien-être et au DD, dans sa dimension sociale.. De plus, notre mode de croissance peut entraîner davantage d'inégalités et moins de justice sociale, ce qui peut aussi nuire au bien-être et au DD. En fait, pour produire plus de biens et services et améliorer leur qualité, l'accroissement des quantités utilisées de certaines formes de capital peut s'accompagner de la baisse du stock d'autres formes de capital.

11.5 ANALYSE DES DIFFÉRENTES FORMES DE SOUTENABILITÉ DE LA CROISSANCE

À ce stade, la question qui se pose est celle de la soutenabilité de notre mode de croissance fondé sur l'exploitation de ces différentes formes de capital. De manière générale, le développement est durable si la croissance est soutenable du point de vue économique, social et environnemental. Pour que cela soit possible, il faut que la quantité globale de capital ne diminue pas. Mais la stabilité du stock global de capital est-elle compatible avec le caractère substituable ou non substituable des différents types de

capital ? En réponse à cette question, qui fait toujours débat, deux thèses sur la soutenabilité opposent les économistes : la thèse de la soutenabilité faible et celle de la soutenabilité forte. Ces deux approches s'opposent sur le degré de substituabilité entre les différentes formes de capital.

11.5.1 De la « soutenabilité faible » à la « soutenabilité forte »

Les partisans de l'approche de la soutenabilité faible considèrent que la contrainte écologique, sur le fonctionnement de l'économie et de la société, est faible.

A contrario, les défenseurs de l'approche de la soutenabilité forte estiment que la contrainte écologique sur le fonctionnement de l'économie et de la société est forte.

L'analyse économique de la soutenabilité de la croissance repose sur le degré de substituabilité entre les différentes formes de capitaux utilisés pour produire.

Les tenants de la « soutenabilité faible » estiment que la nature est un capital productif comme les autres. Il est donc substituable. En revanche, les partisans de la « soutenabilité forte » ne partagent pas cet optimisme. Ils considèrent que les atteintes au stock capital naturel sont, dans une certaine mesure au moins, irréversibles : les dommages causés à l'environnement restent en partie irréparables et certaines ressources épuisables sont irremplaçables. Il n'y a pas de substituabilité.

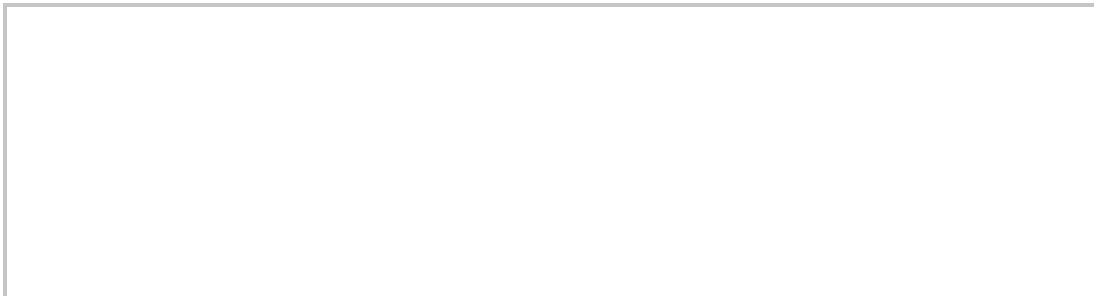
11.5.2 La soutenabilité faible

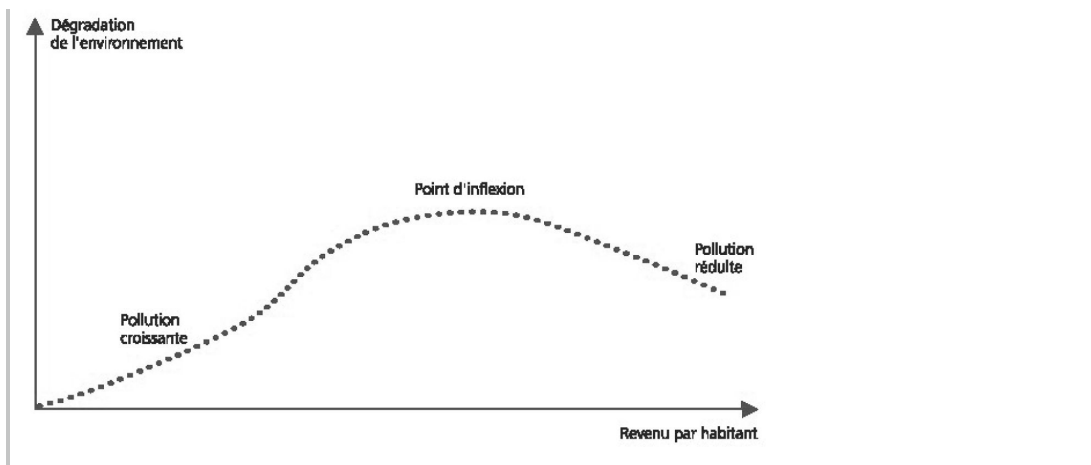
Cette approche, privilégiée dans la théorie néo-classique, considère qu'il existe plusieurs types de capitaux : le capital naturel, le capital social, le capital humain et le capital physique, substituables entre eux. Les mécanismes de marché fondés sur les prix relatifs, les phénomènes de rente liés à l'épuisement des ressources naturelles, ainsi qu'une politique environnementale adaptée doivent permettre d'assurer la soutenabilité de la croissance en maintenant

constant le stock global de capital. Une diminution du stock d'un type de capital peut être compensée par un investissement accru dans d'autres types de capitaux. Selon cette approche, c'est l'évolution du stock global qui prévaut. En effet, le stock de capital global peut croître, même si le stock de certaines formes baisse. Il peut y avoir compensation. Il y a substituabilité, car le capital naturel est considéré comme un capital productif comme les autres. Les coûts relatifs des facteurs de production influencent les choix de combinaison productive. Par exemple, une baisse du capital naturel, social et institutionnel peut être compensée par une hausse du capital physique. De plus, on peut observer que l'homme a pu sauvegarder et même réintroduire des espèces animales ou reconstruire des milieux naturels menacés. Un fleuve pollué peut être dépollué, une forêt détruite replantée, la biodiversité reconstituée.

11.5.3 *La courbe de Kuznets environnementale*

Cette conception repose sur un certain optimisme et sur une confiance dans la capacité du progrès technologique à répondre aux défis écologiques. Elle s'appuie notamment sur « la courbe de Kuznets environnementale » : la croissance économique provoquerait dans un premier temps une dégradation de l'environnement puis permettrait progressivement de réduire ces atteintes (recul de l'industrie, technologies plus efficaces, augmentation des dépenses de protection, délocalisation, adaptation des comportements aux prix, hausse de la réglementation). Ce type de représentation viendrait soutenir la thèse d'une compatibilité entre la croissance économique et la préservation de l'environnement.





GRAPHIQUE 11.2.

La courbe de Kuznets environnementale

L'analyse proposée et illustrée par la courbe de Kuznets environnementale nous conduit fort logiquement à la thèse de la soutenabilité forte.

11.5.4 *La soutenabilité forte*

Selon l'approche dite de la soutenabilité forte, les différentes formes de capital ne sont pas substituables. En particulier, la substituabilité n'est pas possible entre une hausse du capital physique et une baisse du capital naturel. Pour les tenants de cette approche, les destructions du capital naturel sont souvent irréversibles. Les atteintes à l'environnement restent au moins en partie irréparables et il existe des ressources épuisables. De même, une hausse du capital physique et une baisse du capital social ne sont pas substituables, au sens où il y a plus d'inégalités. Les progrès technologiques ne peuvent nous permettre de repousser indéfiniment les limites écologiques de la croissance économique.

Certains courants de l'opinion contestent l'idée même de développement durable et se prononcent en faveur d'une « décroissance soutenable » en se réclamant notamment des travaux de N. Georgescu-Roegen (1906-1994) (*La Décroissance. Entropie, écologie, économie*, 1995). Dans cette perspective, il faut mettre en cause de façon radicale

nos modes de production et de consommation, en finir avec le productivisme et mettre l'accent sur la qualité de la vie, plutôt que sur la quantité de biens et services que l'on peut acquérir. Ces positions sont cependant très minoritaires parmi les économistes qui soulignent qu'une part très importante de la population mondiale est privée de l'accès au logement, à la santé, à l'éducation et qu'une croissance négative ne serait sans doute pas le bon moyen d'apporter une réponse à ces exigences. En revanche, une transformation radicale des modes de croissance, des choix technologiques et des politiques publiques, est indispensable pour assurer à la fois une soutenabilité sociale et environnementale de la croissance. On peut même considérer que la conversion écologique des économies (réduction des émissions à effet de serre notamment) serait une source importante de croissance et d'emploi, dans le cadre de ce que l'on appelle la « croissance verte ».

Selon l'approche qui sera retenue (soutenabilité faible ou soutenabilité forte), les politiques de croissance et de développement menées seront différentes. Si le progrès technologique peut compenser la baisse du capital naturel, alors, les politiques environnementales seront très différentes que dans le cas où prévaudrait la soutenabilité forte.

11.6 LES INSTRUMENTS DES POLITIQUES CLIMATIQUES

Le dérèglement climatique désigne l'ensemble des transformations observées et prévues du climat de la planète. Il se traduit par une hausse des températures à l'échelle planétaire, des phénomènes climatiques exceptionnels qui se multiplient et qui sont de plus en plus violents comme les tempêtes, inondations, sécheresses, la réduction de la biodiversité et tout ce qui menace la survie de l'espèce humaine. Le changement climatique s'explique pour une bonne part par « l'effet de serre » produit par l'émission et l'accumulation dans l'atmosphère terrestre de « gaz à effet de

serre » (GES), dont le CO₂. Les principaux GES sont émis en raison de la combustion des énergies fossiles (charbon, pétrole, gaz). Le dérèglement climatique est visiblement le résultat des activités humaines liées à la croissance économique depuis les débuts de la révolution industrielle.

Face à ces situations, que les mécanismes du marché ne peuvent gérer, les pouvoirs publics doivent mener des politiques environnementales et climatiques pour faire face, compenser et réparer les effets des externalités négatives et autres défaillances de marché inhérentes à l'activité économique dans des économies capitalistes de marché. Les **politiques environnementales** sont l'ensemble des objectifs et instruments mis en œuvre par les pouvoirs publics en vue de préserver l'environnement, et les **politiques climatiques** désignent l'ensemble des objectifs et instruments utilisés par les pouvoirs publics pour réduire les émissions de gaz à effet de serre et les changements climatiques induits.

Les pouvoirs publics peuvent tout d'abord **orienter les préférences** des agents, notamment par l'information et l'éducation. On peut ici souligner le rôle du Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC). L'ancien vice-président américain A. Gore, à travers le film *Une vérité qui dérange* (2006), a aussi œuvré dans ce sens. L'information du consommateur passe par un affichage environnemental des produits, facilement identifiable et interprétable. Par ailleurs, de nouvelles méthodes permettent de favoriser des comportements vertueux en matière environnementale. Elles consistent en des « coups de pouce », d'où leur nom de *nudges verts*, en faveur de comportements éco-vertueux sans recourir à la contrainte. Par exemple, proposer par défaut la démarche la plus écologique (c'est le cas des factures numériques dans certaines banques). Cependant, ces mesures, bien qu'utiles, ne sauraient suffire.

En amont des externalités négatives, il s'agit d'inciter les comportements vertueux, responsables. En aval des

externalités négatives, il s'agit de sanctionner les comportements non vertueux, nuisibles, destructeurs et irresponsables. Globalement, l'objectif est de ne pas épuiser les ressources en capital naturel et de lutter contre les dérèglements climatiques et leurs effets.

Pour cela, les politiques environnementales et climatiques disposent de deux grands types d'instruments : les instruments réglementaires et les instruments économiques (la taxation et les marchés de permis ou quotas d'émission).

11.6.1 *Les instruments réglementaires : les normes*

Les pouvoirs publics, instances nationales ou supranationales, vont définir des règles à respecter sous peine de sanctions. Ces normes peuvent prendre la forme de lois, décrets, arrêtés, directives.

Il s'agit d'interdire et d'encadrer les comportements des agents économiques, borner leur choix et de les inciter à « internaliser les externalités » de leurs activités économiques. De cette manière, on espère réduire les activités polluantes ou vecteurs d'externalités négatives pour l'environnement et le climat. Ces normes peuvent avoir pour objectif la réduction de l'empreinte écologique.

Les exemples abondent : gestion des déchets, limitation de la vitesse sur les routes, obligation du pot d'échappement catalytique et fixation d'un seuil maximal d'émissions de CO₂ pour les véhicules automobiles, interdiction des gaz chlorofluorocarbones.

On peut distinguer les normes d'émission ou de rejet (seuil maximal d'émissions de CO₂, principal gaz à effet de serre [GES]), les normes de procédé ou de processus (interdiction de l'utilisation des chlorofluorocarbones dans les circuits de refroidissement des réfrigérateurs), les normes de produit (pots d'échappement catalytiques et interdiction du plomb dans le carburant) et les normes de qualité (pour l'air, l'eau, circulation alternée des véhicules et limitation de vitesse). Les

effets des réglementations sont limités dans leurs effets et donc leur efficacité, car elles sont trop uniformes. Elles ne tiennent pas suffisamment compte des situations et activités relatives des agents économiques, lesquels doivent donc réaliser un même effort. En outre, quand les normes fixent un seuil maximum, les agents ne sont aucunement incités à réduire les externalités négatives de leurs activités, tout simplement parce que le seuil a peut-être été mal défini. Afin de pouvoir différencier les niveaux d'exigence, les pouvoirs publics disposent des instruments économiques des politiques environnementales et climatiques.

La réglementation est un instrument souvent utilisé, car relativement simple. Elle est parfois efficace, comme le démontre l'interdiction du bisphénol A. La réglementation peut faire naître des innovations, plus respectueuses de l'environnement. La question du « trou de la couche d'ozone » a été très largement résolue grâce à l'interdiction des CFC (chlorofluorocarbones) qui étaient notamment utilisés dans les appareils frigorifiques.

Toutefois, sur le plan économique, la réglementation présente l'inconvénient de l'uniformité, toutes les entreprises doivent s'y conformer, quelle que soit leur taille, par exemple. En outre, un effet pervers peut découler d'une norme limitant la pollution par unité produite ou consommée. En effet, elle peut conduire à un effet-rebond, analysé par S. Jevons dès 1865. La baisse du coût d'utilisation par unité conduit à accroître l'utilisation du bien et augmenter la pollution. Par exemple, la réglementation européenne limite les émissions des nouvelles voitures à 130g de CO₂ par kilomètre en 2015. Les producteurs limitent alors la consommation en carburant par kilomètre des véhicules pour s'y conformer. Mais les automobilistes sont alors incités à plus emprunter ce mode de transport puisque son coût relatif diminue. De ce fait, la pollution globale augmente.

11.6.2 *Les instruments économiques : marché de quotas et taxation selon le principe du « pollueur-payeur »*

11.6.2.1 Les taxes environnementales et climatiques

Ces taxes sont des instruments inspirés des travaux de Arthur C. Pigou. On parle donc souvent de « taxe pigouvienne » pour désigner les taxes qui visent à internaliser les effets externes négatifs. Arthur Pigou (1877-1959) est un économiste anglais qui est l'un des premiers auteurs à avoir réfléchi à l'économie environnementale, dans le cadre de l'économie du bien-être. Ce champ de l'économie étudie les conditions dans lesquelles on peut assurer le maximum de satisfaction aux individus qui composent la société. En étudiant un certain nombre de situations non optimales (situations dans lesquelles on peut améliorer le bien-être d'un individu sans détériorer celui d'un autre individu), Pigou met en avant le rôle déterminant des externalités négatives (exemple : la pollution causée par une usine rejetant ses déchets dans une rivière, qui va affecter la situation des pêcheurs). Il s'agit dans ce cas de rétablir l'optimalité en comblant l'écart entre le coût privé et le coût social provoqué par les effets externes environnementaux : c'est le principe du « **pollueur-payeur** ».

Alors qu'il réfléchit à un moyen de réduire la pollution à Londres, Pigou développe un mécanisme permettant d'intégrer les externalités au coût des activités : le principal effet des externalités est que le coût privé diffère du coût pour la société. Par exemple, quand une usine pollue, son coût (dit privé) est plus faible que le coût social, puisqu'elle n'intègre pas la pollution qu'elle génère dans ses coûts. Raisonnant uniquement sur le coût privé (qui est faible), elle va produire plus que si elle prenait en compte le coût social (qui intègre le coût de traitement des déchets). L'externalité négative va donc engendrer une surproduction.

Pigou propose de mettre en place une taxe du montant de l'externalité, afin que le coût social soit le coût effectif pour la firme. La mise en place d'une telle taxe devrait ainsi réduire les effets négatifs. À titre d'exemple, un raisonnement du même type a été appliqué dans le cadre de la taxe carbone : en taxant les pollueurs, on s'attend à ce que ces derniers réduisent leur pollution.

Les taxes peuvent en outre être à l'origine d'un « double dividende » : réduire la pollution par des incitations à modifier ses comportements polluants, et dégager des recettes fiscales pour alléger d'autres prélèvements obligatoires, notamment le coût du travail et/ou l'investissement dans des infrastructures ou des énergies moins polluantes. À ce double dividende statique s'ajoute un double dividende dynamique, car la fiscalité écologique incite à l'innovation. Cependant, la fiscalité écologique souffre de certaines faiblesses : elle est difficile à calibrer, car les pouvoirs publics ne connaissent qu'imparfaitement son niveau optimal. D'autre part, elle nécessite une coopération internationale pour éviter les distorsions de compétitivité, difficile à mettre en œuvre.

On peut étendre le raisonnement aux subventions si certains comportements vertueux des agents ont un rendement social supérieur à leur rendement privé.

11.6.2.2 Les marchés de permis d'émission

Les marchés de permis d'émission (parfois appelés « marchés des droits à polluer » ou « marchés de quotas ») sont des solutions inspirées directement des travaux de Ronald Coase (1910-2012). Il s'agit de modifier les droits de propriété privée pour introduire les sources des émissions polluantes. Les pouvoirs publics créent des quotas d'émissions répartis selon une procédure établie. Ils sont ensuite transférables entre les agents responsables de la pollution que l'on souhaite limiter. Chaque émetteur effectue alors son propre calcul coût-bénéfice pour décider des efforts de dépollution les plus avantageux. Les firmes les moins polluantes peuvent alors revendre les quotas inutilisés aux

firmes les moins capables de comportements vertueux. Par le jeu de la concurrence, ces dernières doivent disparaître à terme, au profit des firmes les moins polluantes.

Selon le rapport du CAE de 2010 (op. cit.), *« indépendamment de la définition du mécanisme d'allocation de permis d'émission dans le cadre d'un système d'échange global, il peut s'avérer judicieux d'informer les responsables politiques et le grand public des émissions nationales de gaz à effet de serre par habitant. Aussi proposons-nous d'inclure les émissions de gaz à effet de serre par habitant actuelles en tant que deuxième indicateur des GES de notre tableau de bord »*.

Ce type de marché a été mis en œuvre aux États-Unis à partir du Clean Air Act en 1990, dont certains amendements créent un mécanisme d'échanges de quotas d'émissions de dioxyde de soufre (SO₂) et d'oxydes d'azotes (NOx), deux facteurs majeurs des « pluies acides ». En décembre 2010, l'Agence américaine de protection de l'environnement (EPA) publiait un bilan, très satisfaisant pour les émissions de SO₂ liées aux centrales électriques utilisant des énergies fossiles, surtout le charbon, avec une diminution de 67% en 2009 par rapport à 1980.

En Europe, ce type de marché a été créé en 2005 pour les émissions de dioxyde de carbone (CO₂), dans le cadre du protocole de Kyoto, nommé le Système Communautaire d'Échange de Quotas d'émission. Il limite les gaz à émettre et crée un marché du carbone, permettant à chaque entreprise d'acheter ou de vendre les quotas obtenus dans un premier temps à titre gratuit. Les entreprises qui font des efforts sont en principe récompensées et les autres, qui ont dépassé leurs plafonds d'émissions et doivent acheter des quotas d'émissions auprès d'entreprises plus vertueuses sur le plan environnemental, sont pénalisées. Mais ce système a connu de profonds dysfonctionnements, avec une multiplication des cas de fraudes et des périodes de spéculations qui éloignent le prix du carbone de sa valeur fondamentale. En outre, le

prix est très sensible aux fluctuations de l'activité et au niveau initial des quotas alloués. Or, les quotas se révèlent trop élevés, notamment pour préserver la compétitivité des entreprises européennes. Le prix du carbone s'est effondré lors de la récession de 2008 et le signal prix envoyé par le marché est alors contre-productif. Un prix dérisoire est une véritable prime à la pollution, ce qui n'est pas l'effet recherché.

LES MARCHÉS DE QUOTAS VUS PAR ROGER GUESNERIE

L'efficacité environnementale de la politique, dès lors qu'elle est effectivement mise en œuvre, **est toute entière définie par la cible de réduction globale et ne dépend aucunement de sa répartition**. Le point est essentiel. La fixation des quotas individuels ne fait que déterminer la répartition du coût global, soit directement si les quotas n'étaient pas échangeables, soit indirectement si un marché de permis conduit à une réallocation en principe mutuellement avantageuse desdits quotas.

L'instauration d'un marché international de permis dans le dispositif Kyoto, à partir du système de quotas nationaux initialement envisagé, qui a été à l'époque vivement critiquée en Europe, **n'affectait donc pas la performance environnementale recherchée**.

À effort global donné, ce marché contribuait en principe à diminuer le coût global de la politique climatique, en transférant l'effort de là où il est le plus coûteux vers là où il est le moins coûteux. Chacun en principe bénéficiait de la possibilité d'échanges et non comme le débat public le laissait croire, les seuls acheteurs potentiels de permis, comme les USA. La vraie question, celle de l'effort maximal qui peut être accepté par un "pollueur", comme l'étaient les USA, renvoie au rapport de forces dans la négociation multilatérale, dont tout laisse à penser qu'il est indépendant de l'existence ou non d'un marché de permis. Contrairement à ce que laissaient suggérer à l'époque certains arguments polémiques, le marché international des permis n'est pas, par nature, un moyen d'exonérer le pollueur. On peut même retourner l'argument : en abaissant les coûts, ce marché permet au contraire, à rapport de forces donné, d'accroître sinon le coût, mais le quota qui peut être imposé au pollueur !

Comme les autres instruments dits économiques, **un marché de permis tente de concilier marché et écologie**. Il n'est pas pour autant exempt de critiques.

Par ailleurs, le protocole élaboré à Kyoto fixe un objectif quantitatif mondial de réduction. Il instaure, même si elle se décline souplement dans l'espace en fonction des marchés de permis d'émission et des quotas nationaux, ce que l'on appelle une politique de quantités. Cette politique peut être jugée trop rigide dès lors que son coût est incertain, ce qui est le cas. Nombre de spécialistes ont fait valoir que le système serait amélioré si les prix de marché de permis étaient encadrés : un prix maximum constituerait ainsi une "soupape de sécurité". »

Source : Roger Guesnerie, *L'économie de marché*, Le Pommier, 2006

14 Contrairement à ce qu'affirme l'argument de la « tragédie des biens communs », les problèmes de ressources communes peuvent être résolus par des organisations volontaires plus efficacement que par un État coercitif. Parmi les cas considérés figure la tenure communale de prairies et de forêts, des communautés d'irrigation, des droits relatifs à l'eau ainsi que des sites de pêche. (Au Moyen-Âge, la tenure désigne le mode de concession d'une terre accordée par un seigneur (qui en conserve la propriété) à celui qui en aura la jouissance.) La gouvernance des biens communs apporte une contribution majeure à la littérature analytique et à notre conception de la coopération humaine. » (*La gouvernance des biens communs*, Elinor Ostrom, De Boeck Supérieur, 2010) Elinor Ostrom (née en 1933 à Los Angeles) est une politologue américaine. Ses travaux portent principalement

sur la théorie de l'action collective et des biens publics (matériels ou immatériels) et s'inscrivent dans le cadre de la « nouvelle économie institutionnelle ». En octobre 2009, elle est la première femme à recevoir le Prix Nobel d'économie, avec Oliver Williamson, « pour son analyse de la gouvernance économique, et en particulier, des biens communs ».

15 Voir la définition du bien-être dans le chapitre 4, partie 4.6.2.

MARCHÉ DU TRAVAIL, EMPLOI, CHÔMAGE

« L'erreur de raisonnement est ici manifeste : la crise ne résulte pas du système capitaliste, puisqu'elle n'est apparue qu'à l'instant et dans les domaines où l'on a empêché de jouer le mécanisme caractéristique du système dont on prétend démontrer l'inefficacité (la flexibilité des prix et des salaires). Ce que prouve le chômage anglais, ce n'est pas l'impuissance du mécanisme des prix, mais bien au contraire, le fait que lorsqu'on en paralyse le fonctionnement, aucun équilibre économique ne saurait subsister. » - Jacques Rueff, 1931

« Les deux vices marquants du monde économique où nous vivons sont, le premier que le plein emploi n'y est pas assuré, le second que la répartition de la fortune et du revenu y est arbitraire et manque d'équité. » - John Maynard Keynes, 1936

« La baisse du chômage est une fin en soi, mais c'est aussi le principal moyen de faire évoluer le pouvoir d'achat, à tous les niveaux de revenu. La vraie bataille pour le pouvoir d'achat passe inévitablement par le retour au plein emploi. Travailler plus pour gagner plus c'est surtout plus de travail pour plus de création de revenus. » - Charles Wyplosz, 2008

SOMMAIRE

[12.1 Population active, emploi et chômage](#)

[12.2 Le fonctionnement du marché du travail : approches théoriques](#)

12.3 Le fonctionnement dynamique du marché du travail et le processus de « destruction créatrice »

12.4 La pluralité des politiques de l'emploi

12.5 La Flexicurité : pour un équilibre entre flexibilité de l'emploi et sécurisation des revenus et des parcours professionnels

12.6 Croissance et chômage : cycle de productivité et vitesse d'ajustement de l'emploi

12.1 POPULATION ACTIVE, EMPLOI ET CHÔMAGE

12.1.1 *Du travail à l'emploi*

Le travail est une activité socialement et institutionnellement organisée par le droit du travail et les conventions collectives, qui peut parfois être confondue avec l'emploi. Pour l'économiste, d'un point de vue théorique, le travail est souvent considéré comme une désutilité, une insatisfaction. Il présente un coût d'opportunité, celui du temps de loisir auquel on renonce lorsqu'on choisit de travailler. La rémunération du travail est censée en être la contrepartie. Le travail permet aux individus de participer à la production, mais il favorise aussi l'intégration sociale des individus. Le travail peut être considéré comme un accomplissement, un épanouissement, mais il peut parfois être perçu comme un facteur d'intégration ou d'aliénation de l'individu.

L'emploi est un cadre pour une activité de production rémunérée, une place dans le processus de production, une situation dans la hiérarchie de l'entreprise. L'emploi est un cadre économique, social, institutionnel et juridique dans lequel le travail va être organisé. Le travail dans le cadre de l'emploi confère à celui qui l'occupe un statut social, c'est-à-dire un ensemble de droits et de devoirs vis-à-vis de la société. L'emploi confère à celui qui l'occupe, à celui qui travaille, une identité professionnelle et sociale. Les conditions de travail, la durée du travail... dépendent du droit du travail, des conventions collectives, des types d'organisation du travail en vigueur. Les différentes formes d'organisation du travail agissent sur le contenu du travail, et donc sur la demande de travail par les employeurs. Les types d'emploi offerts, les types de contrats de travail proposés, les statuts différents proposés dépendent, dans une économie donnée à un moment donné, des formes privilégiées de

l'organisation du travail et sont en partie déterminés par les volumes respectifs de l'emploi et du chômage observés. Quand, sous l'effet combiné des changements d'équilibre du marché du travail, de l'organisation du travail, du droit du travail et des conventions collectives, les formes de l'emploi se multiplient et se diversifient, alors les statuts sociaux changent et ils peuvent se fragiliser.

Les liens entre travail et emploi sont multiformes. Travail et emploi sont donc très étroitement liés, mais ne se confondent que transitoirement. Les changements de la nature du travail et des types d'emploi proposés sont concomitants.

DE LA POPULATION TOTALE À LA POPULATION ACTIVE

La population active regroupe les personnes ayant un emploi et les personnes disponibles à la recherche d'un emploi. La population active est donc la somme emploi + chômage. ($PA = E + U$). Cette définition permet de comprendre qu'il peut y avoir augmentation du chômage, sans dégradation obligatoire de l'emploi, ou augmentation de l'emploi sans diminution sensible du niveau de chômage. Les évolutions des deux phénomènes dépendent en effet de l'évolution de la population active elle-même, ainsi que de l'évolution de la productivité.

PT = population totale PA = population active PI = population inactive

E = emploi = population active occupée U = unemployment = chômage

$PâT$ = population en âge de travailler

$PIâT$ = population inactive en âge de travailler et $PInâT$ + population inactive non en âge de travailler

$PT = PA + PI$ $PT = PAO + U + PI$

$PA = PAO + U$ $PAO = E$

$PA = E + U$ $E = PA - U$ $U = PA - E$

$PâT = PA + PIâT$ $PI = PInâT + PIâT$

Ces notions peuvent être précisées à partir des conventions statistiques, qui considèrent que la population en âge de travailler (« population active disponible » ou « population active potentielle ») est composée des personnes âgées de 15 à 64 ans. La population totale d'un pays comme la France en 2017 peut ainsi être représentée de la façon suivante :

TABLEAU 12.1.

La composition de la population en France

<i>âge</i> 64 ans population active potentielle = 40, 9 M 15 ans 0 an		Population «incapable d'activité » : 13,1 millions (M)		
		Inactifs en âge de travailler : 11,6 M	Chômeurs : 2,8 M	Population active occupée : 26,5 M
		Population «incapable d'activité » : 13,1 M		
		Population totale : 67,1 M		
Population active totale : 29,3 M				

Source : Insee (enquête Emploi 2017, Bilan démographique 2017) in Jean-Paul Brun, David Mourey, *Marché du travail, emploi, chômage*, De Boeck Supérieur, 2019

12.1.2 Le chômage : définitions et mesures

12.1.2.1 La définition du chômage

Le **chômage** désigne une situation dans laquelle se trouvent les personnes, de 15 ans et plus, qui sont sans emploi, mais qui recherchent effectivement un emploi. Comme souvent en économie, définir et mesurer est difficile, car les frontières entre emploi, chômage et inactivité ne sont pas toujours faciles à établir. Dans ces conditions, l'existence de plusieurs indicateurs complémentaires peut se révéler indispensable et utile.

La mesure du chômage exige de distinguer les données en valeur absolue (le nombre de chômeurs) et des statistiques en valeur relative (le taux de chômage). Compte tenu de l'hétérogénéité des tailles respectives des économies dans le temps et dans l'espace, cette différence n'est pas anodine. Ainsi, pour un nombre de personnes en âge de travailler, d'actifs..., on ne peut se contenter de comparer les nombres de chômeurs entre la France, les États-Unis et la Chine, par exemple.

12.1.2.2 La mesure du nombre de chômeurs

En France, nous disposons de deux sources de chiffres pour le chômage : une mesure statistique proposée par l'Insee et une mesure administrative calculée par Pôle emploi. L'Insee mesure le chômage en utilisant une définition et les critères

établis par le Bureau international du travail (BIT). Pôle emploi mesure le chômage en comptabilisant les demandeurs d'emploi en fin de mois (DEFM).

Le chômage selon l'Insee : un chômeur au sens du BIT est une personne en âge de travailler, sans emploi, qui est disponible pour travailler, qui recherche effectivement un emploi. Conformément à la logique du BIT, pour l'Insee, un chômeur est une personne en âge de travailler (conventionnellement 15 ans ou plus) qui n'a pas travaillé (dépourvue d'emploi) ne serait-ce qu'une heure au cours de la semaine de référence (de l'enquête), qui est disponible pour travailler dans les deux semaines, qui a entrepris des démarches actives de recherche d'emploi dans le mois précédent, ou a trouvé un emploi qui commence dans les 3 mois. L'enquête Emploi de l'Insee permet de mesurer le chômage au sens du BIT. Depuis 2003, cette enquête Emploi est réalisée en continu en métropole alors qu'elle était auparavant réalisée sur un seul mois. Elle permet désormais des estimations trimestrielles du chômage au sens du BIT et non plus seulement annuelles.

Pôle emploi mesure le **nombre de demandeurs d'emploi en fin de mois (DEFM)**.

Pôle emploi est un établissement public à caractère administratif qui a une mission de service public de l'emploi. Pôle emploi est le résultat de la fusion entre l'ANPE et les Assedic en décembre 2008. La Dares et Pôle emploi diffusent les statistiques mensuelles sur les demandeurs d'emploi inscrits à Pôle emploi. Celles-ci sont élaborées à partir des fichiers de demandeurs d'emploi enregistrés en fin de mois par Pôle emploi. Le chômage est seulement « estimé », à partir du nombre de demandeurs d'emploi. **Le nombre de DEFM ne mesure pas strictement le nombre de chômeurs** et cette « mesure » peut être très différente d'un pays à l'autre : les règles administratives et juridiques d'inscription, de radiation... sont très hétérogènes.

Cinq catégories de DEFM : Les demandeurs d'emploi inscrits à Pôle emploi sont regroupés dans une nomenclature composée de cinq catégories différentes (**A, B, C, D, E**) **sur la base de deux critères principaux** : – réaliser des actes positifs de recherche d'emploi – avoir un emploi ou avoir exercé une activité réduite (CDD courts, intérim) au cours du mois.

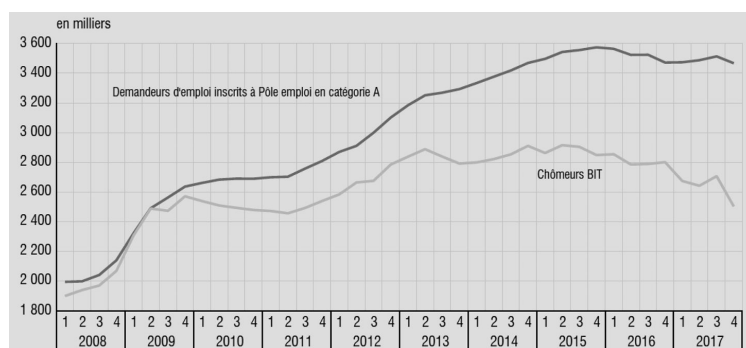
Catégorie A : sans-emploi, demandeurs d'emploi tenus de faire des actes positifs de recherche d'emploi, sans-emploi ;

Catégorie B : demandeurs d'emploi tenus de faire des actes positifs de recherche d'emploi, ayant exercé une activité réduite courte (c'est-à-dire de 78 heures ou moins au cours du mois) ;

Catégorie C : demandeurs d'emploi tenus de faire des actes positifs de recherche d'emploi, ayant exercé une activité réduite longue (c'est-à-dire de plus de 78 heures au cours du mois) ;

Catégorie D : demandeurs d'emploi non tenus de faire des actes positifs de recherche d'emploi (en raison d'un stage, d'une formation, d'une maladie...), sans emploi ;

Catégorie E : demandeurs d'emploi non tenus de faire des actes positifs de recherche d'emploi, en emploi (par exemple : bénéficiaires de contrats aidés).



Champ : France métropolitaine.

Note : données trimestrielles CVS (pour le chômage au sens du BIT), CVS-CJO (pour les demandeurs d'emploi).

Sources : *Emploi, chômage, revenus du travail*, édition 2018 - Insee Références

Dares-Pôle emploi, fichiers STMT (demandeurs d'emploi) ; Insee, enquête Emploi (chômeurs au sens

GRAPHIQUE 12.1.

Évolution du nombre de demandeurs d'emploi inscrits à Pôle emploi en catégorie A et du nombre de chômeurs au sens du BIT

12.1.2.3 Quelques situations particulières de chômage : chômage

de longue durée, chômeurs découragés, chômage partiel

- Un **chômeur de longue durée** est une personne au chômage qui déclare chercher un emploi depuis au moins un an. Le taux de chômage de longue durée est le rapport entre le nombre de chômeurs de longue durée et le nombre de personnes actives.

Au second trimestre 2018, parmi les chômeurs, 1,0 million déclare rechercher un emploi depuis au moins un an (Insee, 14 août 2018). Selon les indicateurs de la Dares de juillet 2018, en France, l'ancienneté moyenne des demandeurs d'emploi en catégories A, B, C est de 607 jours au deuxième trimestre 2018. La durée moyenne d'inscription en catégories A, B, C des demandeurs d'emploi sortis des catégories A, B, C au deuxième trimestre 2018 est de 309 jours.

- Les **chômeurs découragés** sont les personnes qui souhaitent travailler, sont disponibles pour le faire, mais qui déclarent ne plus rechercher d'emploi parce que la perspective d'y parvenir leur paraît trop faible.
- Le **chômage partiel** est une situation liée aux fluctuations de l'activité des entreprises. Lorsqu'une entreprise réduit son activité au-dessous de l'horaire légal ou arrête momentanément tout ou partie de son activité et qu'elle n'entend pas rompre les contrats de travail qui la lient à ses salariés, elle peut avoir recours au chômage partiel.

12.1.2.4 Le calcul du taux de chômage

Le taux de chômage selon l'Insee : au-delà d'une mesure du chômage en valeur absolue, nous disposons d'une

mesure en valeur relative. On calcule le taux de chômage en faisant le rapport entre le nombre de chômeurs et la population active.

TC =	Nombre de chômeurs	=	C	Le taux de chômage est le résultat d'un rapport. Dans ces conditions, la définition de la population qui est au chômage détermine en partie la valeur du taux de chômage : plus le nombre de personnes considérées comme chômeurs est important, plus le taux de chômage est élevé. De même, la définition de la population active peut faire varier le taux de chômage dans le sens inverse.
	Nombre d'actif		PA	

Selon l'Insee, le 14 août 2018, au second trimestre 2018, en moyenne le taux de chômage au sens du BIT est de 9,1% de la population active en France (hors Mayotte), après 9,2% au premier trimestre 2018. Par rapport au deuxième trimestre 2017, le taux de chômage diminue de 0,3 point. En France métropolitaine, le taux de chômage diminue ainsi de 0,2 point sur le trimestre (après + 0,3 point au premier trimestre), à 8,7% de la population active. Le taux de chômage de longue durée s'établit ainsi à 3,6% de la population active au deuxième trimestre 2018, comme au trimestre précédent. Il diminue de 0,4 point sur un an.

Source : Jean-Paul Brun, David Mourey, *Marché du travail, emploi, chômage*, De Boeck Supérieur, 2019

Attention : Le taux de chômage est différent de la part des chômeurs dans la population totale. Il ne faut pas confondre le taux de chômage avec la part des chômeurs au sein de la population totale (C/PT), qui est inférieure au taux de chômage (TC = C/PA). La part des chômeurs est le rapport entre le nombre de chômeurs et le nombre de personnes d'une population considérée ; elle diffère du taux de chômage, qui est le rapport entre le nombre de chômeurs et le nombre de personnes actives (en emploi ou au chômage).

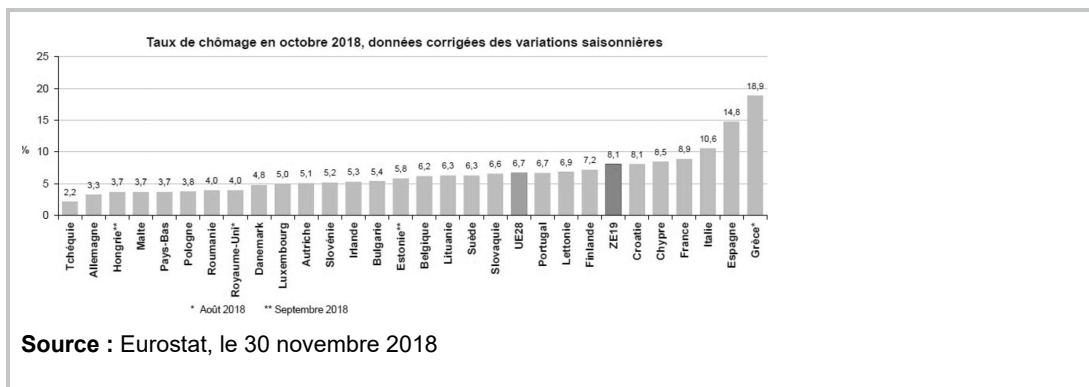
Le calcul du taux de chômage par Eurostat

Pour les pays membres de l'Union européenne (UE), la mesure du chômage par **Eurostat s'appuie sur l'enquête sur les forces de travail** et la définition du BIT.

Les **chômeurs** sont les personnes âgées de 15 à 74 ans qui étaient sans emploi au cours de la semaine de référence, étaient disponibles pour travailler, et étaient activement à la recherche d'un emploi au cours des quatre semaines précédentes ou avaient trouvé un emploi devant débuter dans les trois mois suivants.

Dans la zone euro (ZE19), le taux de chômage corrigé des variations saisonnières s'est établi à 8,1% en octobre 2018, stable par rapport à septembre 2018, et en baisse par rapport au taux de 8,8% d'octobre 2017. Cela demeure le taux le plus faible enregistré dans la zone euro depuis novembre 2008. Dans l'UE28, le taux de chômage s'est établi à 6,7% en octobre 2018, stable par rapport à septembre 2018, et en baisse par rapport au taux de 7,4% d'octobre 2017. Cela demeure le taux le plus faible enregistré dans l'UE28 depuis le début de la série mensuelle sur le chômage en janvier 2000.

Source : Eurostat, le 30 novembre 2018



GRAPHIQUE 12.2.

Taux de chômage des États membres de l'UE et de la ZE

12.1.3 Les inégalités face au chômage

L'observation révèle clairement les inégalités face au chômage selon l'âge, le sexe, le niveau de diplôme et de qualification, la catégorie socioprofessionnelle, la nationalité, le lieu de résidence...

Taux de chômage selon l'âge et le sexe

Jusqu'à la fin des années 2000, le taux de chômage des femmes était plus élevé que celui des hommes mais les inégalités en termes de taux de chômage (TC) devenaient de plus en plus faibles à partir de 2010. Au 2^e trimestre 2018, on peut constater que les TC des hommes et des femmes sont de 8,8 % de la population active. Cependant, même si l'écart s'est réduit, voire inversé, cela ne suffit pas à illustrer une réelle réduction des inégalités entre hommes et femmes face à l'emploi et au chômage. Les jeunes (15-24 ans), avec un TC de 20,6 %, au troisième trimestre 2018, sont nettement plus touchés que les adultes (25-49 ans) avec un TC de 8,2 %. Le TC des seniors (50 ans et plus) est très faible, il est de

6,1 % (la faiblesse du taux d'emploi et du taux d'activité en est la contrepartie).

Taux de chômage selon le niveau de diplôme et la durée depuis la sortie de formation initiale

On peut observer que le TC augmente quand le niveau de diplôme et de qualification baisse. La corrélation est clairement négative. Ainsi, en 2016, le TC global est de 10,1 % ; le TC des personnes sans diplôme s'élève à 17,9 %, celui des individus ayant un BAC, CEP ou BEP est beaucoup plus faible à 10,7 %. Enfin, celui des détenteurs d'un diplôme de niveau BAC+2 ou plus est bien plus faible : 5,7 %. Ainsi, le niveau de diplôme et de qualification n'est pas sans impact sur le niveau de chômage des actifs. Plus tard, fin 2017, selon l'Insee, pour un TC global de 9,4 %, le TC de ceux qui ont un diplôme supérieur à Bac+2 est de 5 %, il est de 5,2 % pour ceux qui ont BAC+2, le TC est de 10,0 % pour ceux qui ont le BAC ou un brevet professionnel et de 10 % également pour ceux qui ont un CAP, BEP ou autre diplôme de ce niveau. Le TC est de 13,9 % pour ceux qui détiennent le Brevet des collèges. Enfin, le TC monte à 18,3 % pour ceux qui n'ont aucun diplôme ou le certificat d'études primaires (CEP).

Taux de chômage selon la nationalité : étrangers, immigrés

Sur le marché du travail, les statuts d'étranger et d'immigré peuvent être des caractéristiques pénalisantes. Pour l'Insee, un étranger est une personne qui réside en France et ne possède pas la nationalité française, soit qu'elle possède une autre nationalité (à titre exclusif), soit qu'elle n'en ait aucune (c'est le cas des personnes apatrides). Un immigré est une personne née étrangère à l'étranger et résidant en France. Un certain nombre d'immigrés ont acquis, depuis leur arrivée, la nationalité française. En revanche, les Français de naissance nés à l'étranger et résidant en France et les étrangers nés en France ne sont pas des immigrés.

En 2016, on observe que le taux de chômage des étrangers (20,5 % pour les femmes et 19,4 % pour les hommes) est deux fois supérieur à celui des personnes de nationalité française (9,2 % pour les femmes et 9,6 % pour les hommes). Cette différence est encore plus marquée pour les étrangers non originaires de l'Union européenne, notamment les femmes (27,3 % pour les femmes et 23,5 % pour les hommes). Environ 6 % des actifs sont étrangers.

En 2014, selon l'Insee, 17,2 % des actifs immigrés étaient au chômage, contre 9,1 % des Français nés en France ; le taux de chômage atteignait 20,7 % pour les actifs non ressortissants de l'Union européenne. Cet écart est en partie lié au diplôme, mais cette explication ne suffit pas : à niveau de diplôme équivalent, le taux de chômage des immigrés demeure supérieur ; par exemple en 2011, 4,7 % des actifs français nés de parents français titulaires d'une licence et plus étaient au chômage, contre 14,8 % des immigrés non ressortissants de l'Union européenne ayant le même niveau de diplôme.

12.1.4 La difficile mesure du chômage et les indicateurs complémentaires du marché du travail

La difficile mesure du chômage en raison de l'effacement des frontières entre emploi et chômage conduit l'Insee à utiliser de nombreux indicateurs complémentaires du taux de chômage : le taux d'emploi, le taux d'activité, le halo autour du chômage, le sous-emploi, le taux de sous-emploi...

- **Le taux d'activité** correspond au rapport (en pourcentage) entre la population active et la population en âge de travailler (personnes âgées de 15 à 64 ans, c'est-à-dire la « population active potentielle » ou « disponible ») ; par exemple selon l'Insee en 2017 le taux d'activité en France est de 71,5%, ce qui signifie que sur 100 personnes âgées de 15 à 64 ans, 71,5 occupent un emploi ou en recherchent un.

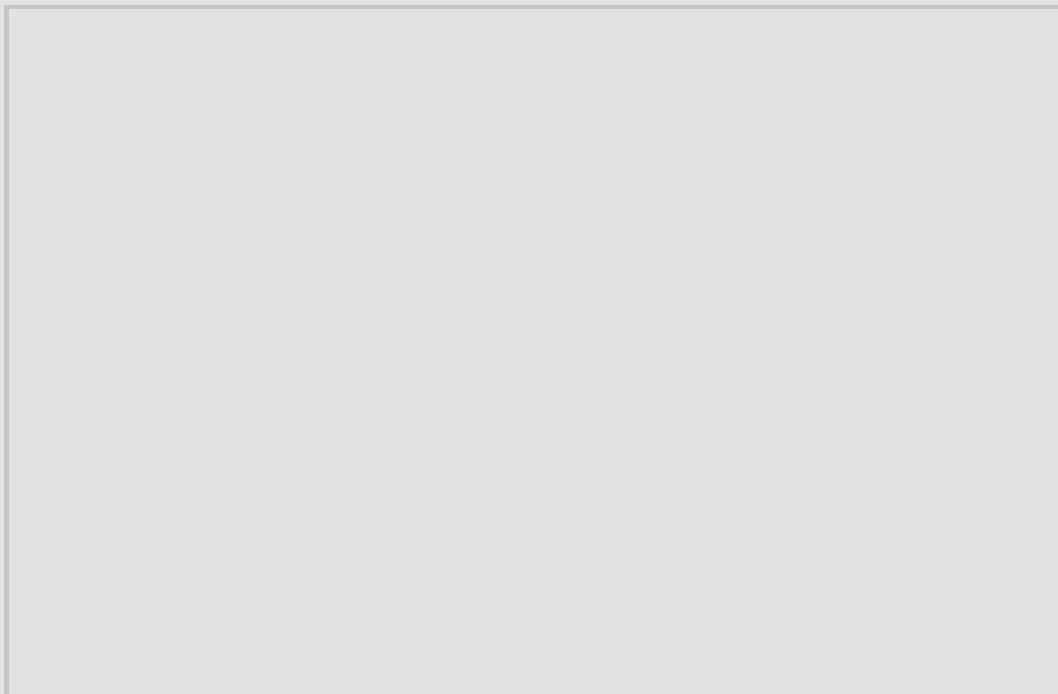
Le taux d'activité (appelé aussi « taux de participation » au marché du travail) mesure la participation, la propension à l'activité d'une population. Il tend à augmenter à long terme en France, du fait de la hausse de l'activité féminine et de la participation croissante des seniors au marché du travail.

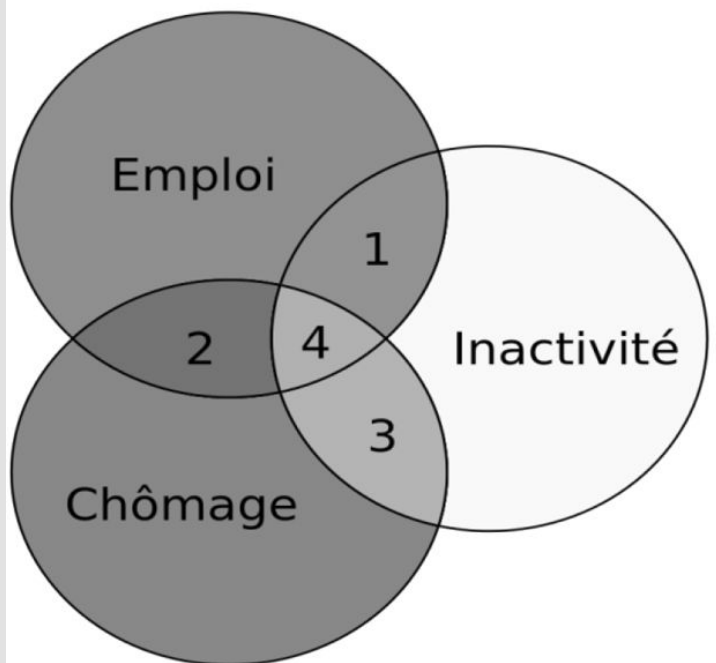
– **Le taux d'emploi** d'une catégorie d'individus est calculé en rapportant le nombre de personnes de la catégorie ayant un emploi au nombre total d'individus dans la catégorie. Il peut être calculé sur l'ensemble de la population d'un pays, mais on se limite le plus souvent à la population en âge de travailler ou à une sous-catégorie de la population en âge de travailler (femmes de 25 à 29 ans...).

Par exemple en 2018 selon l'OCDE, le taux d'emploi aux États-Unis (71% au 4^e trimestre) est plus élevé que dans la zone euro (67,4% au troisième trimestre), en raison notamment d'une plus forte activité des seniors outre-Atlantique.

– Situations intermédiaires et halo du chômage.

LE HALO DU CHÔMAGE





Le chômage, Jacques Freyssinet, 1984

GRAPHIQUE 12.3.

Les cercles de Jacques Freyssinet et le halo du chômage

De manière générale, le « halo du chômage » désigne l'ensemble des situations intermédiaires entre l'emploi, le chômage et l'inactivité (Jacques Freyssinet, *Le chômage*, 1984). Au sens du BIT, un chômeur est une personne en âge de travailler, qui n'a pas travaillé au cours de la semaine de référence, est disponible pour occuper un emploi dans les deux semaines, et effectue des démarches actives de recherche d'emploi (ou bien a trouvé un emploi qui commence dans les trois mois). Cette définition précise laisse de côté des situations complexes dans lesquelles les frontières entre emploi, chômage et inactivité sont difficiles à établir. Par exemple, un étudiant qui travaille quelques heures par semaine et souhaiterait travailler davantage n'est pas considéré comme chômeur au sens du BIT (puisqu'il exerce une activité professionnelle), mais sa situation s'apparente au chômage en termes de revenu et d'insertion sociale. De même, un homme sans emploi qui n'en cherche plus, découragé par le fait de n'avoir rien trouvé d'acceptable, n'est pas classé comme chômeur, mais sa situation est voisine de celle des chômeurs officiellement recensés comme tels.

Source : Jean-Paul Brun, David Mourey, *Marché du travail, emploi, chômage*, De Boeck Supérieur, 2019

12.1.5 L'évolution du niveau du chômage

L'examen du niveau du chômage en France sur longue période révèle une incontestable tendance ascendante. De 1894 (date des premières statistiques sur le taux de chômage) à 1930, la part des chômeurs dans la population

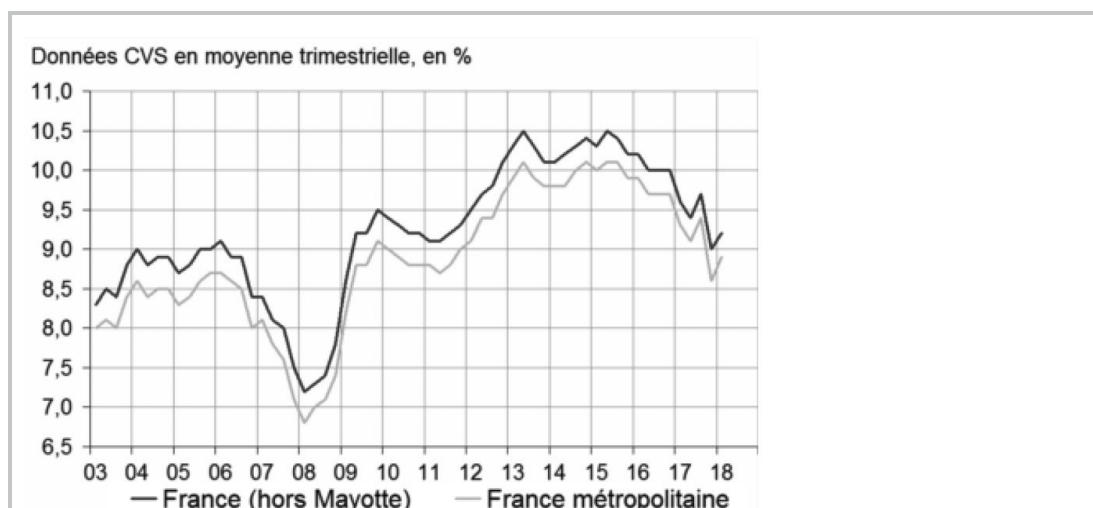
active reste inférieure ou égale à 3,5%, ce qui se rapproche du niveau de plein emploi.

La crise des années 1930 impacte négativement le marché du travail de l'Hexagone : en 1936, le niveau du chômage enregistre un pic, à un peu moins de 5% de la population active (7,5% selon d'autres estimations). Mais cette hausse est nettement plus modérée qu'en Allemagne ou aux États-Unis : dans ce dernier pays, le taux de chômage atteint 25% en 1933...

L'après-guerre voit la France rester durablement au plein emploi, à un taux de chômage inférieur à 3% jusqu'au premier choc pétrolier ; la quasi-totalité des personnes qui souhaitaient travailler trouvaient un emploi, grâce à une forte croissance économique (environ 5% en moyenne par an).

À partir de 1975, la courbe du chômage s'envole jusqu'au milieu des années 1980 où elle se stabilise, et se maintient ensuite à un niveau globalement élevé, en raison notamment du ralentissement de la croissance économique.

Au total, entre 1975 et le premier trimestre 2018, le taux de chômage au sens du BIT a quasiment triplé, passant de 3% à 9,2% en France. Quant au nombre de chômeurs au sens du BIT, estimé à 734 000 personnes en 1975, il atteint 2,6 millions début 2018, soit trois fois plus. Autrement dit, en une quarantaine d'années le marché du travail est passé du plein emploi au chômage de masse.



GRAPHIQUE 12.4.

Taux de chômage en France au sens du BIT (2003-2018), en pourcentage de la population active

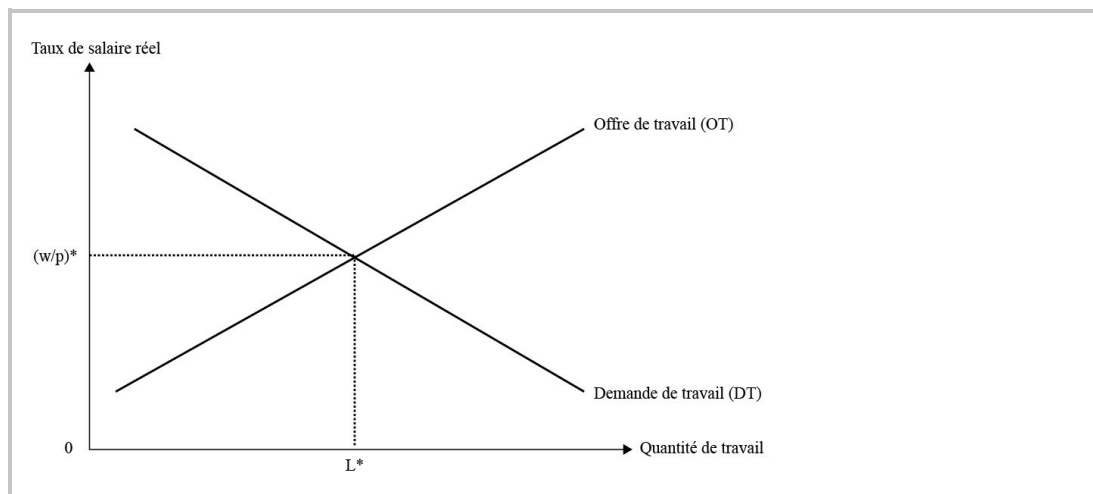
Simultanément, la durée du chômage n'a cessé de croître : un chômeur restait en moyenne 8 mois dans cette situation en 1974, et 15,5 mois en 2017, soit deux fois plus. La crise financière a fortement contribué à cet allongement de la durée du chômage : en France métropolitaine, le nombre de chômeurs de plus d'un an a été multiplié par 2,5 entre octobre 2008 et avril 2016.

12.2 LE FONCTIONNEMENT DU MARCHÉ DU TRAVAIL : APPROCHES THÉORIQUES

12.2.1 *L'équilibre du marché du travail selon les économistes classiques et néoclassiques*

Au sens de John Maynard Keynes, les « classiques » regroupent l'école classique (1776-1871/1874) et l'école néoclassique (1871/1874-1936), c'est-à-dire l'ensemble des économistes qui, à la suite de Smith (1776), privilégient le rôle de la « main invisible » du marché en tant que moyen de régulation efficace de l'activité économique. L'intervention de l'État est inefficace, voire inutile. Pour ces économistes classiques, sur le marché du travail (comme sur tout marché), la rencontre des offreurs de travail (demandeurs d'emploi ; $OT = DE$) et des demandeurs de travail (offreurs d'emploi ; $DT = OE$) aboutit à un équilibre de plein emploi du marché tant que les salaires réels sont flexibles. Il ne peut y avoir de chômage involontaire. Seul le chômage volontaire peut exister, lorsque les offreurs de travail estiment le niveau de salaire réel proposé par le marché trop faible pour compenser leur perte de loisir, quand le coût d'opportunité du loisir devient trop grand. Du côté des employeurs, il convient d'embaucher tant que la productivité marginale du travail

(celle du dernier embauché) est supérieure ou égale au salaire réel d'équilibre. L'offre de travail (OT) désigne la quantité de travail que les agents économiques désirent vendre sur le marché du travail. En résumé, tant que le salaire réel est flexible à la baisse, le mécanisme du marché assure le plein emploi.



GRAPHIQUE 12.5.

L'équilibre du marché de travail

L'emploi dépend du niveau du salaire réel. La flexibilité du salaire réel réalise en permanence l'équilibre entre l'offre et la demande sur le marché du travail. Tout ce qui rend le marché rigide (salaire minimum, syndicats...) est défavorable à l'emploi. Graphiquement, l'offre de travail (OT) est une fonction croissante du taux de salaire (prix du travail). Plus ce dernier est élevé et plus le coût d'opportunité des loisirs est important. La demande de travail (DT) est une fonction décroissante du taux de salaire et de la productivité marginale du travail. Plus le salaire est bas, plus les entreprises sont prêtes à embaucher des travailleurs dont la productivité marginale décroît. L'intersection des deux courbes détermine la quantité de travail d'équilibre (offerte et achetée), ainsi que son prix d'équilibre. À ce niveau d'équilibre E, les offres et les demandes de travail sont satisfaites.

L'hypothèse de survie et la flexibilité à la baisse des salaires sur le marché du travail ont une importance considérable dans le cadre de la théorie néoclassique. En 1995, dans son livre intitulé « Le débat interdit », au chapitre intitulé « Une théorie et son mode d'emploi », Jean-Paul Fitoussi rappelle l'importance incontournable de l'hypothèse de la dotation de survie pour que le marché du travail puisse fonctionner comme un marché. *« Pour les partisans de la théorie libérale des économies de marché, la solution au chômage réside simplement dans la flexibilité à la baisse des salaires, étant donné que les travailleurs doivent être rémunérés selon leur productivité marginale, afin de pouvoir embaucher ceux dont la productivité est faible. Pourtant, selon la théorie pure de l'économie de marché proposée par K. Arrow et G. Debreu (qui sont les artisans modernes de cette théorie) dans un article resté justement célèbre ("Existence of an equilibrium for a competitive economy", *Econometrica*, n° 22, 1954), une hypothèse trop souvent oubliée ou ignorée détermine leur résultat : « Les ressources initiales des agents sont suffisantes pour qu'en toutes circonstances ils disposent du revenu nécessaire pour subsister pendant leur vie entière. Autrement dit, ils disposeraient à tout moment de l'option de ne pas travailler, option qu'ils exerceraient si le prix du travail leur semblait trop bas. Ce qui garantit que le prix du travail ne baissera pas en dessous d'un certain seuil, ce n'est pas l'existence d'un salaire minimum dont les théoriciens supposent l'existence, mais d'une richesse minimum – et on perçoit bien que cela revient au même, sauf quant aux conséquences sur l'emploi. Dans le monde "réel", un travailleur dont l'entreprise jugerait que ses compétences ne valent pas le salaire minimum se retrouverait au chômage. Dans celui décrit par la théorie de l'équilibre général, il s'abstiendrait de travailler. Car, dans l'hypothèse d'une rémunération trop faible du travail, quelle serait l'incitation à l'activité ? » En résumé, pour Arrow et Debreu, si l'on souhaite tirer les enseignements solides d'une théorie, encore faut-il en respecter les conditions de validité. Autrement dit,*

dans une économie de marché, sur le marché du travail, le salaire peut être parfaitement flexible à la baisse si et seulement si la condition de survie est respectée, au même titre que les autres conditions de concurrence pure et parfaite. Aucun travailleur ne doit pouvoir travailler et finalement mourir du fait de la faiblesse de son salaire et de ses revenus. Si le salaire est trop bas, le travailleur doit pouvoir disposer par ailleurs des ressources complémentaires suffisantes. Cette condition est si restrictive qu'elle n'est jamais satisfaite, mais elle nous rappelle dans quelles conditions peut fonctionner un marché du travail conforme à la théorie. »

12.2.2 Le fonctionnement du « marché du travail » selon Keynes

« Je dis que les postulats de la théorie classique ne s'appliquent qu'au cas particulier et non au cas général... En outre, les caractéristiques de ce cas particulier, dont la théorie classique fait l'hypothèse, s'avèrent différentes de celles de la société économique dans laquelle nous vivons effectivement, avec pour résultat que ses enseignements sont trompeurs et désastreux lorsque nous tentons de les appliquer aux faits d'expérience », écrivait John Maynard Keynes durant les années 1930.

Keynes va rejeter l'approche néoclassique pour deux raisons principales :

- D'une part, s'il accepte la courbe de demande de travail (de productivité marginale décroissante), il ne peut accepter le fait qu'il puisse exister des comportements d'offre de travail (et donc de marché du travail *stricto sensu*).
- D'autre part, le niveau de l'emploi ne dépend pas du niveau du salaire réel. En effet, les salariés et les chômeurs sont victimes de l'illusion monétaire. Ils ne s'intéressent qu'au salaire nominal. De surcroît, ils n'ont pas la possibilité d'arbitrer entre travail et loisir, car ils n'ont pas d'autres choix que de travailler pour vivre.

Le niveau de l'emploi dépend, non pas du niveau du salaire nominal, mais de la demande effective ou demande globale anticipée. Il en résulte qu'une baisse des salaires est favorable à l'emploi au niveau microéconomique (toutes choses égales par ailleurs), mais au niveau macroéconomique, une baisse du niveau des salaires (baisse de la masse des salaires) se traduit par une chute de la demande effective. *In fine*, la situation du marché du travail se dégrade avec une destruction nette d'emploi et une hausse du chômage. Un chômage involontaire peut exister. En effet, le salaire n'est pas seulement un coût, il est aussi une demande potentielle. Selon Keynes, la demande effective (DE) détermine le niveau d'emploi. La DE dépend de la demande de consommation (DC) et de la demande d'investissement (DI). La DC est une fonction croissante du revenu. Mais, selon la loi psychologique fondamentale, quand le revenu augmente, la consommation augmente, mais plus lentement. Selon cette loi psychologique fondamentale, « La consommation agrégée dépend essentiellement du revenu agrégé et la propension marginale à consommer ce revenu est positive et inférieure à un. » (John Maynard Keynes, 1936, *Théorie générale de l'emploi de l'intérêt et de la monnaie*, chapitre 8). Autrement dit, l'épargne croît en valeur relative. La DI dépend, pour simplifier, des « esprits animaux » des entrepreneurs, c'est-à-dire de leur confiance en l'avenir de l'économie. S'ils sont optimistes, ils investissent (la DI augmente) et donc la DE croît et l'emploi suit... C'est le cercle vertueux de l'expansion (effet multiplicateur positif). S'ils sont pessimistes, les licenciements et le chômage augmentent. Cette hausse du chômage pèse à la baisse sur les salaires, ce qui réduit d'autant plus le pouvoir d'achat global (effet multiplicateur négatif). *Ipsa facto*, le chômage croît. C'est le cercle vicieux de la récession. D'où le rôle prépondérant de la DE sur le niveau d'emploi.

TABEAU 12.2.

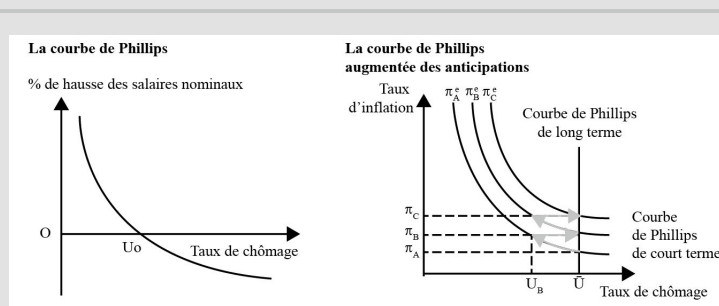
Logiques de détermination de l'emploi selon les classiques et Keynes

Pour les classiques	Pour Keynes
Équilibre sur le marché du travail → Équilibre sur le marché des biens et Services	Équilibre sur le marché des biens et services → Équilibre sur le marché du travail

Pour les classiques, le niveau de l'emploi est déterminé sur le marché du travail et, en vertu de la loi des débouchés (l'offre crée sa propre demande), un équilibre sur le marché des biens et services en résulte. Pour Keynes, *a contrario*, le niveau de l'emploi est déterminé sur le marché des biens et services et, en vertu du rejet de la loi des débouchés et du principe de la demande effective, l'équilibre le marché du travail en résulte.

LES ENJEUX DE LA COURBE DE PHILLIPS ET DU DILEMME INFLATION/CHÔMAGE

Dans un article de 1958, l'économiste d'origine néo-zélandaise A.W. Phillips a mis en évidence une relation empirique inverse entre le taux de chômage et le taux de variation des salaires nominaux (graphique de gauche), à partir d'une estimation statistique portant sur le Royaume-Uni sur longue période (1861-1957). Sur le plan théorique, la relation découverte par Phillips peut s'expliquer par le fonctionnement du marché du travail : les variations du taux de salaire nominal résultent d'un déséquilibre entre l'offre et la demande de travail qui implique un certain taux de chômage (celui qui assure la stabilité des salaires se situerait à $U_0 = 5,5\%$ au Royaume-Uni selon Phillips). Une demande excédentaire de travail des entrepreneurs entraîne une pression à la hausse des salaires nominaux et une diminution du taux de chômage, notamment, car la puissance contractuelle des salariés est renforcée. Inversement, une croissance du taux de chômage implique le ralentissement de la croissance des salaires nominaux, alors que le pouvoir de négociation des travailleurs diminue. Si l'on suppose un lien stable entre la croissance des salaires nominaux et l'inflation, en raison du comportement de marge des firmes (relation stable entre les prix de vente et le coût salarial), et une proportionnalité entre la variation du salaire réel moyen et la productivité du travail, on peut en déduire une relation inverse entre le taux d'inflation et le taux de chômage (graphique de droite, courbe de gauche)..



GRAPHIQUE 12.6.
Courbes de Phillips

La courbe de Phillips devient alors une relation inflation/chômage et fournit une représentation rigoureuse du dilemme des politiques macroéconomiques au sens des économistes de la synthèse

néoclassique (Robert Solow, Paul Samuelson) : une accélération de l'inflation semble alors le prix à payer pour réduire le taux de chômage (politique de relance), tandis que la désinflation entraîne une remontée du taux de chômage (politique de rigueur). Enfin, à long terme, il n'y a pas d'arbitrage inflation-chômage, la courbe de Phillips est verticale en vertu des anticipations adaptatives. Les agents économiques sont transitoirement trompés dans leurs anticipations d'inflation et de salaire réel et confondent, dans un premier temps, la hausse effective de leur salaire nominal avec celle de leur salaire réel. Cependant, ils corrigent leurs anticipations et leurs comportements. **À long terme, la courbe de Phillips est verticale.** C'est ce que montre le graphique de droite avec le déplacement de la courbe de Phillips vers la droite pour obtenir une baisse du taux de chômage, il faut accepter une hausse du taux d'inflation. Après ajustement, adaptation, des anticipations, le taux de chômage revient à son niveau naturel.... Dans ce cas, la politique monétaire (voire budgétaire) de relance peut être efficace à court terme, car elle peut tromper temporairement les agents, victimes de l'illusion monétaire.

12.2.3 Théorie des équilibres à prix fixes : chômage « classique » et chômage « keynésien »

Les fondateurs (R. Clower, 1965 ; A. Leijonhufvud, 1968, et E. Malinvaud, 1977) de cette approche, baptisée « théorie du déséquilibre », rejettent le postulat d'ajustement automatique de l'économie vers l'équilibre de plein emploi. Les déséquilibres économiques résultent de l'imperfection de l'information et de l'absence d'un commissaire-priseur pour centraliser l'information et coordonner les décisions individuelles. Les prix et les salaires sont considérés comme exogènes, c'est-à-dire indépendants des variations relatives de l'offre et de la demande. En courte période, la rigidité des prix rend nécessaires des ajustements par les quantités, selon la règle du côté court. Cela conduit à des situations de rationnement sur chaque marché, car la règle du côté court va déterminer les quantités échangées. Ces rationnements vont se propager d'un marché à l'autre, en raison des interdépendances, et cette théorie cherche à établir des interactions entre les situations de déséquilibre sur les marchés des biens et services et du travail. Ces rigidités et leurs effets au niveau macroéconomique ont donc des fondements microéconomiques. En croisant les équilibres simultanés sur les marchés des biens et services et du travail, les tenants de cette théorie mettent en évidence trois types de déséquilibres, ou équilibres à prix fixes, avec

rationnement. Ainsi, pour chaque agent et sur chaque marché, il existe :

- une offre et une demande effectives (c'est-à-dire achats et ventes réalisés) ;
- une offre et une demande dites « notionnelles » (c'est-à-dire quantités que l'agent aurait souhaité échanger sur un marché donné) ;
- Si les achats réalisés sont inférieurs aux achats souhaités, alors les acheteurs sont rationnés.

Cette approche conduit Malinvaud à un réexamen de la théorie du chômage dans l'article « Réexamen de la théorie du chômage » (*Revue économique*, volume 32, n° 6, 1981, pp. 1196-1199). La « théorie du déséquilibre » rejette l'hypothèse de l'ajustement automatique des marchés par les variations des prix.

TABEAU 12.3.

Les types de déséquilibres et de chômage

Marché du travail	Marché des biens et services	
	Marché acheteur : $O > D$	Marché vendeur : $D > O$
Marché acheteur $OT > DT$	Chômage keynésien : vendeurs rationnés sur les deux marchés c'est-à-dire offreurs rationnés sur le marché des biens et offreurs rationnés sur le marché du travail Insuffisance de demande	Chômage classique : demandeurs rationnés sur le marché des biens offreurs rationnés sur le marché du travail
Marché vendeur $DT > OT$	Cas atypique très peu probable	Inflation contenue : pressions inflationnistes fortes

Il va ainsi distinguer deux types de chômage :

- Le chômage est **keynésien** quand les offreurs/vendeurs sont rationnés sur les deux marchés, car ils auraient souhaité vendre davantage. Il y a donc excès d'offre ou déficit, insuffisance de demande. Concrètement, des

entreprises sont disposées à produire davantage, mais ne sont pas incitées à le faire en raison de l'insuffisance de la demande anticipée (carnets de commandes peu remplis) et des travailleurs se présentent sur le marché du travail, mais ne trouvent pas d'emploi. Dans ce cas, une politique de soutien de la demande réduirait le chômage involontaire.

– Le chômage est **classique** quand les demandeurs/acheteurs sont rationnés sur le marché des biens et les offreurs/vendeurs sont rationnés sur le marché du travail, car le niveau des prix est trop bas et celui des salaires est trop élevé. La rentabilité est faible et les entreprises ne sont pas incitées à embaucher. Concrètement, les entreprises ne souhaitent pas produire davantage, car le niveau des salaires réels est jugé trop élevé et le niveau des profits est estimé insuffisant. L'investissement et l'emploi vont ralentir. Il convient, dans ces conditions, de restaurer les profits pour inciter les entreprises à investir et à embaucher et, *in fine*, faire baisser le chômage.

– **L'inflation** est dite « **contenue** », dans la mesure où les pressions inflationnistes sont fortes quand la demande excède l'offre sur les deux marchés. Cette théorie des équilibres à prix fixes a fourni un cadre rigoureux pour mieux cerner les diverses formes de chômage possibles et mieux définir les mesures de politique économique adaptées.

12.2.4 *Les théories de la segmentation du marché du travail*

Quand la théorie de Keynes a conduit à une critique du modèle microéconomique néoclassique à partir d'une approche macroéconomique, les théories de la segmentation du marché du travail proposent un enrichissement du modèle néoclassique standard en mettant en évidence différentes formes de segmentation du marché du travail sur la base de

la remise en question de certaines hypothèses fondamentales du modèle néoclassique.

12.2.4.1 Les trois dimensions de la segmentation du marché du travail

Le marché du travail est segmenté, car les hypothèses d'homogénéité, de fluidité du marché, de transparence du marché du travail ne sont pas respectées. Un marché segmenté est un marché cloisonné, fragmenté, compartimenté. Le marché du travail n'est pas unifié et la segmentation permet d'expliquer certaines formes de chômage. L'auteur principal de ces théories est Michael Piore avec ses articles « Notes pour une théorie de la stratification du marché du travail » en 1972 et « Le dualisme sur le marché du travail » en 1978. On peut donc mettre en évidence deux approches différentes et complémentaires de la segmentation du MT. Avec la première, on peut distinguer trois formes de segmentation. Avec la seconde, on peut observer un dualisme du MT fondé sur les caractéristiques des contrats de travail.

– Le marché du travail est hétérogène

Quand l'hypothèse d'homogénéité du marché du travail n'est pas respectée, le MT est segmenté, différencié, en fonction des qualifications des travailleurs et des postes. Par exemple, on pourrait distinguer un MT des cadres et un MT des ouvriers, selon le niveau de qualification. On peut alors expliquer la persistance d'emplois vacants quand le chômage est élevé. Il y a inadéquation entre l'offre et la demande de travail. C'est un chômage de segmentation dû à des caractéristiques qualitatives.

– Le marché du travail n'est pas fluide

Quand l'hypothèse de fluidité n'est pas respectée, il existe des barrières à l'entrée et à la sortie du MT. Il y a mobilité géographique imparfaite de la main-d'œuvre. Les actifs ne peuvent être parfaitement mobiles. Cela s'explique en partie par la barrière des distances. Dans ce cas encore, malgré l'équivalence quantitative des offreurs et des demandeurs, on

observera un chômage lié à des caractéristiques qualitatives, à une segmentation géographique du marché du travail.

– L'information sur le marché du travail n'est pas parfaitement transparente

Lorsque l'hypothèse de transparence n'est pas respectée sur le marché du travail, les offreurs et demandeurs d'emploi (ou de travail) n'ont pas un accès suffisant ni une capacité suffisante d'analyse de l'information sur les postes offerts ou les compétences des potentiels travailleurs. Cette segmentation peut se combiner avec la segmentation géographique quand les chômeurs ne sont pas forcément informés de l'existence d'emplois correspondants ailleurs.

En résumé, cette théorie fait apparaître et explique des formes de chômage qui n'étaient ni décrites ni expliquées par les théories traditionnelles du marché du travail et du chômage. Les diverses formes de chômage sont désormais dues à des caractéristiques qualitatives, car même si les nombres de demandeurs et d'offeurs d'emplois sont égaux, il y a du chômage. Ces formes de chômage font partie du chômage structurel ou chômage lié aux déséquilibres structurels de l'économie. L'inadéquation entre qualifications offertes et demandées en est un, de même que les déséquilibres géographiques ou les asymétries d'information.

12.2.4.2 La théorie du dualisme du marché du travail

Dans le cadre de cette théorie, les statuts des emplois liés à la nature des contrats de travail des salariés sont cruciaux. Selon cette théorie, il existe deux types de statuts (ou de contrats) et donc deux segments sur le MT. Le marché primaire offre au travailleur un statut favorable et le marché secondaire offre un statut défavorable. Chaque compartiment du marché correspond à un type de travailleur. Certains auteurs distinguent même un troisième segment, un marché tertiaire, celui du travail clandestin, le travail au noir du secteur informel, de l'économie de l'ombre.

12.2.5 *Les analyses contemporaines du marché du travail*

Ces théories se fondent également sur un relâchement des hypothèses de concurrence pure et parfaite (CPP). La concurrence est imparfaite sur le marché du travail lorsqu'il y a non-atOMICité du marché du travail, si l'offre de travail est

hétérogène, quand l'employabilité diffère selon les catégories de travailleurs, lorsque l'accès à l'information est sélectif (opacité de l'information) et si le facteur travail est peu mobile (entrée-sortie non libre).

Les conséquences de ces imperfections se traduisent par des rigidités des salaires et des prix, l'ajustement ne peut se faire normalement et le chômage involontaire redevient une possibilité, mais pour des raisons microéconomiques. Ces nouvelles théories du marché du travail vont permettre d'expliquer rationnellement la rigidité du salaire réel.

12.2.5.1 La théorie du capital humain

Selon Gary Becker (prix Nobel 1992), on peut rejeter l'hypothèse d'homogénéité du facteur travail, car chaque travailleur dispose d'une dotation en capital humain spécifique, un ensemble de connaissances, de compétences, d'expérience et de capacités productives. Le capital humain est un stock variable (il peut être accumulé), immatériel et inséparable de la personne qui le détient. Les facteurs d'acquisition, d'accumulation du capital humain sont l'éducation et la formation (formelle et informelle), la santé physique et mentale (vaccin = bien collectif), la possibilité de migration de l'offre de travail pour valoriser le capital humain... L'investissement en capital humain est une dépense et présente un coût d'opportunité, celui des gains immédiats auxquels on renonce en cherchant à en accumuler. Son existence pose aussi le problème des externalités positives. Enfin, des différences de capital humain justifient des différences de salaires. Le niveau de salaire peut transitoirement diverger de la productivité marginale du travail et le salaire réel peut être supérieur à l'embauche, pour attirer les meilleurs. Mais, à long terme, le capital humain implique un retour sur investissement, une productivité supérieure à la productivité marginale du travail et donc un salaire réel inférieur au salaire d'équilibre. Par rapport à une situation de concurrence pure et parfaite (CPP), les comportements d'offre et de demande de travail sont

modifiés. Par exemple, en période de ralentissement de la croissance, les salaires sont maintenus tant que la productivité marginale du travail est supérieure. La théorie du capital humain est donc une théorie de la rigidité des salaires, de l'emploi et du chômage.

12.2.5.2 La théorie des contrats implicites

Selon D. F. Gordon, M. N. Bailey et C. Azariadis (1975), il convient de rejeter l'hypothèse de flexibilité des salaires. En effet, un contrat implicite, un accord tacite entre employeurs et employés permet de privilégier l'emploi par rapport au salaire en tant que variable d'ajustement. Une « poignée de main invisible » va lier les parties par cet accord informel, car, en raison de l'aversion des salariés aux fluctuations de leurs revenus, ils sont incités à accepter un salaire réel transitoirement inférieur à celui qui résulterait des seuls mécanismes du marché et de la conjoncture. Autrement dit, la rigidité des salaires s'explique par les formes respectives d'aversion au risque des salariés et des employeurs. L'aversion des salariés à la fluctuation de leurs revenus, les incite à accepter un salaire réel en moyenne inférieur à celui, fortement variable, qui résulterait des seules forces du marché. Les salariés, cherchant à se prémunir contre de trop fortes variations de leur rémunération, troquent la stabilité de celle-ci, à un niveau inférieur au salaire d'équilibre, contre le maintien de leur emploi. L'employeur, neutre au risque, jouerait alors le rôle d'assureur. Une sorte de prime d'assurance est perçue par les employeurs. En cas de récession, il y a maintien de l'emploi, du salaire et partage du travail dans l'entreprise. Il y a faible rotation des emplois et le chômage reste stable. Ainsi, sur le marché du travail, des relations de long terme entre salariés et employeurs dominent. Cette théorie permet d'expliquer la rigidité des salaires et de l'emploi, mais elle explique mal le chômage involontaire, lorsque les emplois sont maintenus. *In fine*, le mode de formation du salaire réel n'obéit pas à la règle de

l'égalisation avec la productivité marginale, et l'économie se retrouve dans une situation sous-optimale.

12.2.5.3 La théorie de la recherche d'emploi

La théorie du "*job search*" (**Stigler, Liman, Mac Call**) met en présence deux acteurs :

- L'offreur de travail, c'est-à-dire le chômeur ;
- Le demandeur de travail, c'est-à-dire l'entreprise.

Ces deux acteurs se rencontrent donc sur le marché du travail. La théorie du "*job search*" décrit le processus menant à un accord entre offreur et demandeur, et les conditions, notamment en termes de durée du chômage, qui permettent que cet accord se réalise.

Si un salarié perd son emploi, il n'est pas forcément dans son intérêt d'accepter l'emploi qui lui est immédiatement proposé, notamment si la rémunération qui lui est offerte est inférieure à celle qu'il estime correspondre à ses qualifications. La référence que le demandeur d'emploi a en tête en termes de rémunération acceptable est directement liée à l'emploi qu'il occupait précédemment. Les individus se fixent ainsi « **un salaire de réservation** », qui est le salaire minimum qu'ils acceptent : en dessous de ce seuil, ils refusent de travailler (plus le salaire de réservation est élevé, plus la période de chômage est longue).

L'individu va donc accumuler de l'information pour sa **recherche d'emploi**, ce qui a un coût (collecte des offres disponibles, durée et conditions de travail, achat de journaux, frais de communication, etc.), tandis que les revenus qu'il perçoit se limitent aux indemnités de l'assurance chômage. Il a toutefois intérêt à prolonger sa recherche d'emploi aussi longtemps que son coût est inférieur au bénéfice qu'il en retire : c'est ce **calcul coût / bénéfice** qui va déterminer la durée de sa recherche d'emploi (à quel moment et à quelles conditions est-il alors rentable de chercher un travail et de sortir de l'inactivité ?). Cette comparaison durera tant que la valeur actualisée de la somme des revenus qu'il touchait dans son ancien emploi reste supérieure à la valeur

actualisée de la somme des revenus de l'emploi qui lui est proposé (et qu'il refuse) et des indemnités chômage. Ainsi, les exigences salariales du demandeur d'emploi vont baisser au fur et à mesure que dure sa recherche d'emploi.

Du côté de l'entreprise, des coûts de recherche d'un travailleur pour occuper un poste vacant sont également engagés : au fil du temps et de la durée de recherche, l'employeur va augmenter progressivement le salaire proposé. Il y a donc une **durée optimale** nécessaire pour que le demandeur d'emploi et l'employeur se rencontrent (*job matching*) et qu'il y ait adéquation entre leurs différentes caractéristiques.

La théorie du "*job search*" éclaire donc les **déterminants microéconomiques** du chômage (un chômage de « prospection ») : dans un environnement où l'information est imparfaite, la recherche d'emploi peut être longue et plus l'indemnisation du chômage et les aides sociales sont généreuses (allocation chômage, minima sociaux, aides familiales), plus l'incitation à allonger cette durée l'est également, et plus le taux de chômage peut être élevé.

Néanmoins, l'allongement de la durée de la recherche d'emploi peut avoir un effet positif si elle permet d'améliorer le processus d'appariement sur le marché du travail et de donner aux travailleurs l'opportunité de trouver des emplois conformes à leur niveau de capital humain.

L'**assurance chômage** est alors une forme de subvention à la recherche d'emploi et permet d'accroître la participation au marché du travail.

Dans ce cadre théorique, le rôle des pouvoirs publics consiste à optimiser l'information sur le marché du travail et à développer le capital humain de la main-d'œuvre. Pourquoi y a-t-il sur le marché du travail tant de personnes sans emploi alors que simultanément, un grand nombre de postes restent vacants ? Telle est la question à laquelle les modèles fondés sur l'**appariement** entre l'offre et la demande de travail essaient de répondre.

Cette théorie du "**matching**" (appariement) a été mise à l'honneur en 2010 lors de la remise du Prix Nobel d'économie aux trois auteurs **P. Diamond**, **D. Mortensen**, et **C. Pissarides**. Dans leur modèle, le chômeur détermine son niveau d'effort de recherche d'emploi de façon à optimiser son revenu permanent, qui est fonction de ses gains nets (allocation chômage retranchée des coûts de recherche) et des gains potentiels tirés d'un changement de situation sur le marché du travail (gains liés à la reprise d'emploi).

12.2.5.4 Les théories du salaire d'efficience

Les différents modèles du salaire d'efficience insistent sur l'impact majeur des asymétries d'information, des coûts de rotation de la main-d'œuvre liés aux indemnités de licenciement, coûts d'embauche, dépenses de formation, coûts d'adaptation... (Phelps et Salop) dans le maintien des salaires (**salaire de réservation**) au-dessus du niveau d'équilibre, les rigidités à la baisse des salaires et l'existence du **chômage involontaire**.

Selon les théories du salaire d'efficience (Leibenstein, 1957), le niveau de productivité dépend du niveau des salaires. Il y a un renversement de la causalité par rapport à la théorie néoclassique. La productivité apparente du travail (PAT = production/quantité de travail) est une fonction croissante du salaire. Plus un actif est payé, plus il sera incité à être productif, à faire des efforts. Cette théorie peut être illustrée par l'exemple historique du "5 \$ day" de Ford, lequel a augmenté les salaires de ses ouvriers pour augmenter leur productivité et leur faire accepter des conditions de travail exigeantes. En conséquence, pour être plus compétitives et rentables, les entreprises peuvent chercher à faire augmenter la productivité du travail en augmentant les salaires de leurs salariés, afin de les motiver à travailler plus efficacement et de les inciter à rester dans l'entreprise (éviter le turnover) en augmentant le coût d'opportunité d'un départ. De cette manière, elles augmentent le salaire d'efficience censé permettre d'obtenir le plus haut niveau de productivité, au moindre coût. Cela dit, au-delà d'un certain seuil, le salaire d'efficience n'a plus d'effets positifs sur la productivité des travailleurs. Ford s'en est aperçu : en effet, « on ne peut faire boire un âne qui n'a plus soif », disait-il. Enfin, l'instauration d'un salaire d'efficience peut être source de chômage si toutes les entreprises procèdent ainsi, car le salaire moyen va augmenter et se situer au-dessus du salaire d'équilibre, ce qui va créer une hausse du chômage. Mais le chômage est involontaire, car les chômeurs ne refusent pas de travailler. Ils sont au chômage, car il n'y a pas d'emploi pour eux à ce

niveau de salaire. On retrouve l'argumentation néoclassique (le chômage est dû à un niveau de salaire excessif), avec une différence majeure : ce ne sont pas les syndicats ou l'État qui créent des rigidités, mais les entreprises elles-mêmes.

Ces théories permettent de comprendre que dans des économies libérales, où l'État intervient peu sur le marché du travail, il peut y avoir du chômage. La rigidité des salaires et le chômage involontaire s'expliquent simplement par le lien étroit qui existe entre salaire réel et productivité.

L'asymétrie d'information se traduit par des comportements de type sélection adverse dans la mesure où une entreprise proposant un taux de salaire faible s'exposerait à n'attirer que les salariés les moins performants. Dès lors, en situation d'asymétrie d'information, sur les qualités réelles des postulants, les firmes vont augmenter le taux de salaire d'équilibre pour attirer les travailleurs avec les plus hauts salaires de réservation en espérant que ce soient aussi les plus efficaces.

L'asymétrie d'information va aussi provoquer des comportements de type aléa moral. En effet, dans le modèle de « tire-au-flanc », Shapiro & Stiglitz (1984) proposent de substituer aux contrôles de comportement coûteux et difficiles, des taux de salaires majorés qui accroissent le coût d'opportunité du licenciement et donc dissuadent les individus de « tirer au flanc ».

L'asymétrie d'information et l'existence d'un salaire de réservation chez les salariés se traduisent par une rigidité à la baisse des salaires afin d'attirer les meilleurs par des salaires supérieurs au salaire d'équilibre et par du chômage involontaire. De plus, la proposition de taux de salaires majorés, qui augmentent le coût d'opportunité du licenciement, est censée dissuader les individus de « tirer au flanc ». Enfin, l'existence d'un salaire équitable de l'effort, quand les employeurs proposent un salaire supérieur au salaire d'équilibre pour inciter le salarié à l'effort, à être

productif, implique la rigidité à la baisse des salaires et du chômage involontaire.

Les théories du salaire d'efficience proposent une représentation du marché du travail implique de prendre en compte la **négociation salariale et les institutions du marché du travail**. En effet, en général, les salariés sont payés au-dessus de leur salaire de réservation (celui qui les rend indifférents entre l'emploi et le chômage) et le niveau des salaires dépend des conditions du marché du travail et des institutions qui le régissent (chômage, salaire minimum, protection de l'emploi, etc.). La théorie du salaire d'efficience implique que d'une part les salariés ont un pouvoir de négociation, et que d'autre part, les entreprises acceptent elles-mêmes de payer un salaire supérieur au salaire de réservation pour des raisons de productivité.

Dès lors, le salaire nominal négocié dépend de l'anticipation du niveau général des prix, de l'état de tension du marché du travail tel que résumé par le taux de chômage ou encore le taux d'emploi, de la générosité du système de protection sociale. La négociation entre salariés et entreprises aboutira à un salaire nominal d'autant plus élevé que les agents anticipent un niveau de prix élevé, que le taux de chômage est faible (le taux d'emploi élevé), et que le système de protection sociale est généreux.

12.2.5.5 Le modèle insiders-outsiders

Le modèle insiders-outsiders est apparu dans les années 1980, développé par A. Lindbeck (1984) et D. Snower (1988). Il apporte un élément d'explication nouveau au chômage de longue durée. Au départ, la théorie distingue les salariés qui occupent un emploi, les insiders, de ceux qui n'ont pas d'emploi et qui en cherchent un, les outsiders.

L'existence de coûts de rotation de la main-d'œuvre (Stiglitz, 1974) va conduire à une sorte d'antagonisme entre les insiders et les outsiders (Lindbeck & Snower 1988).

Les coûts de rotation de la main-d'œuvre sont liés aux indemnités de licenciement, aux coûts d'embauche, aux

dépenses de formation, aux coûts d'adaptation... Dans un souci de minimisation de ces coûts, les entreprises vont chercher à fidéliser leur personnel en le rémunérant à un salaire supérieur au salaire d'équilibre.

La prise de conscience par les insiders de l'importance des coûts de rotation pour l'entreprise va les conduire à tenter d'exploiter la rente de situation dont ils bénéficient, au détriment des outsiders, en exigeant un taux de salaire supérieur au taux d'équilibre. La situation est à nouveau sous-optimale.

Les insiders peuvent exiger un salaire plus élevé que le niveau d'équilibre, car les coûts de rotation de la main-d'œuvre sont élevés (coûts d'embauche, de formation...). Cela se traduit par la rigidité à la baisse du salaire des insiders et par du chômage involontaire pour les outsiders. Cependant, il existe un salaire maximum que les employeurs acceptent de verser aux insiders. Au-delà de ce seuil, de ce plafond, le coût de rotation de la main-d'œuvre devient inférieur au coût d'embauche d'un outsider. Il vaut mieux licencier et embaucher. La situation peut donc se révéler sous-optimale.

THÉORIES DU SALAIRE D'EFFICIENCE VS MODÈLE INSIDERS-OUTSIDERS

Les théories du salaire d'efficience insistent sur le rôle des employeurs dans le maintien des salaires au-dessus de leur niveau d'équilibre, alors que les modèles insiders-outsidiers mettent davantage en évidence les comportements des salariés en place et l'acceptation du jeu par les entreprises. Ces deux types de théories se révèlent donc complémentaires, dans la mesure où la rigidité des salaires au-dessus du niveau d'équilibre et le volume de chômage involontaire qui en résulte dépendent de ce que les employeurs sont prêts à consentir (concessions) et de ce que les salariés sont prêts à accepter (revendications).

12.2.5.6 L'hypothèse d'hystérésis dans la théorie du chômage

Pour terminer ce court panorama des théories modernes du marché du travail, on peut évoquer l'hypothèse d'hystérésis dans la théorie du chômage (Phelps, 1972 ; Blanchard et Summers, 1987 ; pour le premier développement formel de la notion). D'un point de vue théorique, cette hypothèse implique

la remise en cause de la notion de taux de chômage naturel et fragilise la frontière qui sépare le chômage d'équilibre du chômage de déséquilibre, laquelle devient instable. Une variable est « frappée » d'hystérésis si l'équilibre auquel elle s'établit est sensible au temps qui passe : un effet persiste alors même que ses causes ont disparu. Ainsi, bien que les chocs à l'origine de la forte augmentation du taux de chômage dans les années 1990 aient disparu, le chômage se maintient à un niveau durablement élevé. Ce phénomène peut s'expliquer, entre autres, par le processus de déqualification (perte de capital humain) auquel sont confrontés les chômeurs de longue durée. De surcroît, selon la théorie de l'hystérésis, le taux de chômage naturel dépend de l'historique des taux de chômage effectifs passés. Il y a hausse simultanée du taux de chômage effectif et du taux de chômage naturel. Cette théorie est donc une théorie de la durée du chômage et pose le problème de l'employabilité. Le cycle économique se trouve donc accentué par ces phénomènes d'hystérésis ou d'ajustement avec retard, à la suite de chocs sur l'offre ou sur la demande. Autrement dit, une politique de désinflation est associée à un ratio de sacrifice beaucoup plus élevé et les récessions seront plus intenses. En 1995, dans son ouvrage *Le débat interdit*, Jean-Paul Fitoussi a montré comment la combinaison de politiques macroéconomiques conjoncturelles de rigueur durables a pu avoir un effet puissant sur la montée du taux de chômage d'équilibre et sur sa persistance à un niveau élevé. En Europe, l'association de politiques budgétaires restrictives, avec volonté de réduction rapide des déficits budgétaires, et de politiques monétaires restrictives a conduit l'économie sur le sentier de la croissance molle et du chômage durablement massif. La faible inflation qui en a résulté s'est accompagnée de taux d'intérêt réels durablement élevés, qui ont renforcé la spirale de la croissance molle.

12.3 LE FONCTIONNEMENT DYNAMIQUE DU MARCHÉ DU TRAVAIL ET LE PROCESSUS DE « DESTRUCTION CRÉATRICE »

L'approche théorique traditionnelle et statique du marché du travail présente des limites en matière de compréhension et d'explication du fonctionnement de celui-ci. De nouvelles analyses se sont donc développées pour proposer une vision dynamique du marché du travail fondée sur une analyse empirique.

12.3.1 *L'approche statique du marché du travail :*

une analyse en termes de stocks

Le cadre d'analyse habituel utilisé par les économistes pour expliquer le fonctionnement du marché du travail est celui du marché. Le marché du travail (MT) est le marché sur lequel les demandeurs de travail sont amenés à rencontrer les offreurs de travail afin d'échanger un service de production. Sur le MT, les offreurs de travail (OT) sont les demandeurs d'emploi (DE) et les demandeurs de travail (DT) sont les offreurs d'emploi (OE). On a donc : $OT = DE$ et $DT = OE$.

Les offreurs de travail sont les personnes qui souhaitent travailler, ils cherchent un emploi. Les demandeurs de travail sont les employeurs qui cherchent du personnel, des personnes à embaucher. Dans le cadre de cette **approche statique**, à une date donnée, le nombre de chômeurs (C) est mesuré par la simple différence entre le nombre d'actifs (PA) et le nombre d'actifs occupés (PAO). Logiquement, la différence entre deux stocks est un stock. Ainsi, le nombre de chômeurs est le résultat d'une simple différence entre le nombre d'actifs et le nombre d'actifs occupés. Soit, $C = PA - PAO = PA - E$.

Il s'agit clairement d'une **manière**, purement **comptable**, **d'envisager le marché du travail et son fonctionnement**.

Cela implique de faire, implicitement, l'hypothèse qu'à un moment donné, dans une économie, il y a un nombre fixe de personnes qui sont actives ($PA = OT = DE$), que certaines, parmi elles, occupent un emploi ($E = PAO = DE$ satisfaites = OT satisfaites) pendant que d'autres en recherchent un ($C = PAI = DE$ insatisfaites = OT insatisfaites). Elle conduit, en effet, à une analyse en termes de stocks : stock ou nombre de travailleurs, stock ou nombre d'emplois, stock ou nombre de demandeurs d'emploi, stock d'offres d'emploi.

Cependant, cette approche peut être trompeuse. En effet, imaginons qu'entre deux dates, les nombres de chômeurs, d'actifs et d'actifs occupés (d'emplois) restent inchangés, on pourrait alors en déduire que rien n'a changé sur le marché du travail. Celui-ci a été statique sur la période considérée. Ce serait pourtant commettre une magistrale erreur d'interprétation, aux conséquences non négligeables. En effet, un diagnostic erroné ne peut conduire qu'à une prescription inadaptée et le remède proposé pourrait causer un problème encore plus profond. La vérité est ailleurs, au moins en partie, car le caractère dynamique du marché du travail implique de nombreux appariements en fonction d'importants flux d'entrée et de sortie de l'emploi et de nombreux flux de créations et destructions de postes. Dans la réalité économique quotidienne, le marché du travail est un marché extrêmement dynamique. Chaque jour, des milliers de personnes entrent et sortent du marché du travail, des milliers d'emplois sont créés pendant que d'autres sont détruits, la composition de la population des chômeurs change... De facto, l'analyse du marché du travail doit, également, être une analyse dynamique, c'est-à-dire en termes de flux. Seul ce type d'étude permettra de mettre en lumière l'existence d'innombrables et perpétuels changements au sein de la population active, dans la composition de l'emploi et parmi les chômeurs.

12.3.2 *L'approche dynamique du marché du travail :*

une analyse en termes de flux

Le caractère dynamique du marché du travail s'explique par d'importants flux d'entrées et de sorties de l'emploi, de nombreux flux de créations et destructions de postes et par les appariements qui en découlent.

Prenons un exemple simple. Si en France, au 1^{er} janvier 2018, on comptabilise environ **29 millions d'actifs** et que le 1^{er} janvier 2019 on dénombre également 29 millions d'actifs, on pourrait rapidement en conclure que la population active (population présente sur le marché du travail) est inchangée, car elle paraît être rigoureusement stable, statique. Si, de surcroît, le nombre de personnes actives occupées (par un emploi) est stable à **26 millions** et que le même nombre de personnes actives inoccupées (chômeurs) reste stable à **3 millions**, alors tout semble complètement stable sur le marché du travail. Mais cette stabilité apparente et purement comptable est un mirage.

Plus précisément, début 2018, l'Insee dénombre en France (Hors Mayotte) 29,3 millions de personnes âgées de 15-64 ans qui sont actives. Parmi elles, 26,5 millions occupent un emploi. Il y a donc, à cette date, 2,8 millions de chômeurs.

En effet, l'observation et l'exploitation des millions de données statistiques désormais disponibles nous montrent, *a contrario*, l'existence de ces flux entrants et sortants de la population active, de l'emploi (de la population active occupée), du chômage... Certaines personnes ayant un emploi sont devenues chômeurs, et *vice versa*. Pour ces multiples raisons, l'étude du marché du travail doit nécessairement être dynamique. L'économiste doit procéder à un **examen minutieux des flux**. Cette **analyse**, qualifiée de **dynamique**, du marché du travail nous conduit à étudier le **processus de destruction créatrice** d'entreprises, d'activités, d'emplois... **et le processus d'appariement** entre l'offre d'emploi et la demande d'emploi qui en découle. L'objectif à atteindre est désormais d'expliquer comment rendre plus efficace ce processus de rencontre entre OT (DE) et DT (OE).

– comment faire pour qu'un demandeur d'emploi trouve un emploi conforme à sa demande ?

- comment faire pour qu'un offreur d'emploi trouve un candidat adapté à ce qu'il recherche ?
- comment faire pour qu'offreurs de travail et demandeurs de travail cessent de se croiser sans se rencontrer ?

12.3.3 *Du processus de « destruction créatrice » au processus d'appariement¹⁶*

Le processus de « destruction créatrice » va modifier la composition du marché du travail. Dans une économie développée, il existe simultanément des chômeurs, des emplois vacants et de nombreux postes sont occupés par des personnes qui ne sont pas à la « bonne place ». Cela peut sembler paradoxal au premier abord, pourtant ce paradoxe n'est qu'apparent. Il révèle les difficultés inhérentes au processus d'appariement. Mais les obstacles à la qualité du processus d'appariement sont multiples et plus ou moins prononcés selon les pays. Il est donc indispensable d'analyser méticuleusement ce processus d'appariement, car l'économie et donc le marché du travail changent tout le temps. Des activités sont en expansion, alors que d'autres sont en extinction. De manière analogue, des entreprises sont créées et d'autres disparaissent, tout comme il existe des créations d'emplois et des destructions d'emploi. Ce processus inévitable et inéluctable de destruction créatrice est indissociable du processus de croissance. Il rend indispensable la gestion d'un processus d'appariement adapté à ces mouvements.

Les économistes Pierre Cahuc et André Zylberberg ont popularisé en France cette approche peu connue du fonctionnement du marché du travail ; dans un ouvrage intitulé *Le chômage, nécessité ou fatalité*, (2004), puis en 2015, *Les ennemis de l'emploi*. Ils observent qu'en France, en moyenne, chaque jour, 10 000 emplois sont détruits et 10 000 emplois sont créés. En moyenne, chaque jour, 30 000 personnes perdent leur emploi et 30 000 personnes trouvent un emploi. Les réallocations de main-d'œuvre sont deux à

trois fois plus nombreuses que les réallocations d'emplois, en raison des départs libres et volontaires de salariés qui démissionnent, qui partent à la retraite sans qu'il y ait destruction d'emploi. Il y a simplement rotation de la main-d'œuvre. De surcroît, en moyenne, une entreprise qui crée un emploi embauche trois personnes et en licencie deux et, toujours en moyenne, la suppression d'un poste de travail s'accompagne de deux embauches et de trois départs ou licenciements.

La croissance économique est le produit d'un double processus où créations et destructions d'emplois sont en interaction avec divers autres mouvements de main-d'œuvre. Ce double processus reflète le libre choix des entreprises et des salariés. L'ampleur des mouvements observables sur le marché du travail conduit à un profond renouvellement de la conception de son fonctionnement. Ce n'est que vers la fin des années 1980 que les économistes ont pris conscience que le marché du travail fonctionne dans le cadre d'un processus de « destruction créatrice » (PDC), pour reprendre les termes de Joseph Schumpeter. Le développement de l'informatique a rendu possible l'exploitation de millions d'observations révélant l'ampleur du PDC d'emplois et ses liens avec la croissance, par la médiation des gains de productivité induits. Sans ce PDC, nous n'aurions jamais connu la croissance économique. « La prospérité provient des réallocations d'emplois. »

Les innovations y jouent un rôle moteur et leurs effets sont puissants si le fonctionnement du marché du travail permet le PDC d'emplois, c'est-à-dire les réallocations d'emplois. Mais, si le PDC apporte la croissance économique, il apporte aussi le chômage et des inégalités. La recherche d'emploi, ou le chômage, est un rouage indispensable du PDC, elle est nécessaire au bon fonctionnement de l'économie. On observe, au sein des pays industrialisés, une forte ressemblance en matière de créations et destructions d'emplois, mais de grandes différences des résultats au niveau du chômage. C'est la gestion différenciée de ce PDC

qui explique les divergences en matière de chômage, et en matière de réactivité du marché du travail (MT).

Croissance et chômage sont codéterminés par le processus de « destruction créatrice » et, dans ce cadre théorique, la mondialisation du commerce international joue un rôle secondaire. Elle n'est pas la cause principale des destructions d'emploi ni du chômage.

12.3.4 *Analyse dynamique du marché du travail et corrélation entre taux de chômage, taux d'emploi et taux d'activité*

Dans le cadre de cette analyse dynamique du marché du travail, on peut mettre en évidence une corrélation négative entre taux de chômage et taux d'emploi. Le taux de chômage est le rapport entre le nombre de chômeurs et la population active. Le taux d'emploi, rapport entre l'emploi et la population en âge de travailler (de 15 à 64 ans) est un complément utile du taux de chômage. En effet, l'existence d'un système d'allocations chômage « généreux » peut inciter les individus à déclarer chercher un emploi pour en bénéficier. *De facto*, le chômage devient en partie la conséquence d'une participation accrue au marché du travail et non d'un volume insuffisant d'emplois.

Empiriquement, de nombreuses études de l'OCDE (au cours des années 1980 et 1990) ont montré que les pays ayant un **taux de chômage** élevé sont aussi ceux qui ont un **taux d'emploi** faible et, symétriquement, les pays ayant un taux de chômage faible sont aussi ceux qui ont un taux d'emploi élevé. Il existe donc une relation décroissante entre taux de chômage et taux d'emploi. Le taux de chômage est donc un indicateur pertinent de l'abondance d'emplois dans un pays.

De surcroît, les pays confrontés à un **taux de chômage** élevé ont également un **taux d'activité** faible. Il existe une relation décroissante entre taux de chômage et taux d'activité (rapport entre population active et population en âge de travailler).

Un taux de chômage élevé ne résulte pas d'un taux d'activité excessivement important.

Donc, de manière générale, au sein de l'OCDE, certains pays étaient confrontés à un taux de chômage relativement élevé, car les créations d'emplois étaient insuffisantes, et non parce que le taux de participation au marché du travail était trop élevé. Il est important de conserver en mémoire ces faits stylisés concernant le chômage, l'emploi, l'activité dans différents pays de l'OCDE. Ces faits montrent clairement que les pays confrontés à des taux de chômage élevés durant les décennies 1980 et 1990 l'ont été à cause d'une insuffisante création d'emplois, et non à cause d'un accroissement important de l'offre de travail. **Dans ces conditions, le partage du travail, de l'emploi, n'est pas une solution viable.**

La corrélation, dans la longue durée, est clairement positive entre taux d'activité et taux d'emploi.

La corrélation négative entre taux de chômage, d'une part, et taux d'activité et taux d'emploi, d'autre part, est plus difficile à observer à chaque sous-période. Pour autant, on peut l'admettre dans la longue durée.

En résumé, les variations du taux de chômage s'expliquent par les variations de la population en âge de travailler, les variations de l'emploi et les variations du taux d'activité. La relation entre taux de chômage et emploi n'est donc pas simple. La hausse du chômage peut s'expliquer davantage par un manque de création d'emplois que par une hausse de la PA. *In fine*, le diagnostic qui sera fait de l'efficacité du fonctionnement du marché du travail à partir d'une analyse en termes de flux va conditionner la médication, c'est-à-dire les solutions éventuellement envisageables, pour remédier aux problèmes de l'emploi et du chômage. Le contenu des politiques de l'emploi à mettre en œuvre en sera, au moins, partiellement modifié. On ne peut faire l'économie des enseignements tirés des modèles d'appariement qui viennent compléter les théories traditionnelles, bien davantage que s'y substituer. Les analyses statiques et dynamiques du marché du travail sont complémentaires bien plus que substituables.

12.4 LA PLURALITÉ DES POLITIQUES DE L'EMPLOI

La diversité des formes et des analyses du chômage explique la pluralité des politiques de l'emploi afin de lutter contre le chômage :

- Les politiques macroéconomiques de soutien de la demande globale pour lutter contre le chômage keynésien
- Les politiques d'allègement du coût du travail pour lutter contre le chômage classique
- Les politiques de formation et de flexibilisation pour réduire la composante structurelle du chômage
- De la « grande récession » aux politiques pour l'emploi

La crise financière à partir de l'été 2007, qui s'est transformée en « Grande Récession » par la suite en affectant durement l'économie « réelle », a entraîné une dégradation des marchés du travail des pays développés et des destructions massives d'emplois. L'impact sur l'emploi a été très important : certains secteurs de l'économie ont été très durement touchés (immobilier, bâtiment, secteur manufacturier, automobile, services bancaires et financiers, etc.). Si la crise ne touche pas tous les secteurs de la même façon, les conséquences diffèrent selon l'âge, le sexe, le niveau d'études et le type de contrat de travail : la crise a surtout eu des conséquences sur l'emploi des jeunes, sur l'emploi temporaire, et a particulièrement pénalisé les travailleurs les moins qualifiés, ainsi que les migrations liées au travail. De plus, en raison des pertes d'emplois dans le secteur formel, la crise a fait progresser le travail informel (économie souterraine), d'où une dégradation des conditions de travail des plus pauvres et une diminution de leurs revenus.

Pour reprendre l'expression utilisée par l'OCDE dans ses « Perspectives de l'emploi » de 2012, la « résilience » des marchés du travail à la crise n'a pas été partout la même. Les situations ont été très contrastées à l'intérieur de l'Union

européenne : les pays d'Europe du Sud (Espagne, Grèce, Portugal) et l'Irlande ont subi le choc le plus dur. Si les pays du nord de l'Europe ont connu une hausse relativement faible du chômage (Pays-Bas, Finlande, Suède, hormis le Danemark), d'autres se trouvent dans une situation intermédiaire avec une hausse sensible du chômage, aux alentours de 10% (France, Italie, Royaume-Uni), tandis que l'Allemagne a connu une régression du chômage sur cette période. Les États-Unis ont connu un doublement de leur taux de chômage, de 5% fin 2007 à 9,9% fin 2009, avant d'entamer une lente décrue par la suite, pour atteindre 5,7% en janvier 2015. La hausse du taux de chômage a surtout touché les travailleurs non qualifiés.

Face à la gravité de cette crise qualifiée de « grande récession », les États ont développé deux types de stratégies complémentaires : à court terme, via des actions sur les composantes de la demande globale et à long terme, via des actions sur les déterminants de l'offre globale, des capacités de production et de leurs coûts respectifs.

– **Les différents types de politiques pour l'emploi**

Lato sensu, les politiques pour l'emploi qui ont pour objectif de soutenir, ou relancer, les créations d'emploi et de lutter contre la hausse du chômage, puis le faire baisser, articulent les politiques macroéconomiques conjoncturelles et les politiques structurelles.

Stricto sensu, la politique de l'emploi désigne l'ensemble des politiques visant à assurer le plein emploi : elle comprend les dispositifs de soutien aux chômeurs (indemnisation du chômage, formation) et les stratégies de stimulation de la création d'emplois (nouveaux).

On distingue généralement deux types de politiques de l'emploi :

- les politiques dites « passives », comme l'indemnisation du chômage et les subventions aux retraits d'activité (comme les préretraites) ;

- les politiques « actives » comme les dispositifs de soutien à la création d'emplois dans le secteur marchand ou dans le secteur non marchand ; la formation professionnelle ; l'accompagnement et le placement des chômeurs, etc.

Les politiques de l'emploi ont une pluralité d'objectifs et font toujours intervenir un arbitrage particulièrement délicat entre, d'une part, la recherche de l'efficacité économique et, d'autre part, le maintien ou l'amélioration de la justice sociale.

12.4.1 *Les politiques macroéconomiques conjoncturelles face au chômage*

En référence aux travaux de John Maynard Keynes (1883-1946) dans une perspective macroéconomique, le chômage résulte d'une insuffisance de la demande anticipée. Cette approche diffère fondamentalement de l'analyse néoclassique qui considère que c'est l'insuffisante rentabilité des entreprises, en raison des rigidités institutionnelles et d'un coût du travail excessif (et non l'absence de débouchés potentiels), qui freine l'embauche. Encore aujourd'hui, les débats restent vifs dans le champ scientifique de la macroéconomie sur les origines du chômage. Selon John Maynard Keynes, l'intervention de l'État est indispensable pour empêcher les dépressions et maintenir l'économie proche du plein emploi (ce qu'il appelle le « boom économique ») afin de « lisser » les fluctuations conjoncturelles.

Dans les économies modernes, la réactivité de la politique économique et sociale est déterminante pour maintenir le plein emploi selon les économistes keynésiens, et la recherche de l'équilibre budgétaire est présentée comme dangereuse et contre-productive lorsque la récession se prolonge.

L'État doit plutôt mobiliser les instruments de la politique conjoncturelle de relance :

- La **politique monétaire**, qui par la baisse du taux d'intérêt directeur de la banque centrale, peut stimuler l'offre de crédit et l'investissement des entreprises ;
- La **politique budgétaire**, qui, dans le cadre d'une politique d'investissements publics par exemple, permet à l'État, en contrepartie d'un creusement du déficit budgétaire, de compenser par la dépense publique l'atonie de la dépense privée. Par ailleurs, même si elle est de nature plus structurelle, la politique de redistribution des revenus (par l'impôt progressif et les dépenses sociales) peut également contribuer à la croissance en soutenant la demande globale et les bas revenus (qui ont une plus forte propension à consommer). Selon Keynes, ce sont les anticipations des entrepreneurs (demande effective) qui déterminent le niveau de l'emploi. Si ces derniers sont optimistes, ils investissent et embauchent, et les revenus distribués valident les décisions de production ; si leurs prévisions sur l'évolution des débouchés deviennent pessimistes, l'incitation à investir décroît, ainsi que la production et les revenus distribués. Contrairement à l'analyse néoclassique qui postule l'équilibre de plein emploi, Keynes démontre qu'il existe une multiplicité d'équilibres de sous-emploi aux différents niveaux de demande effective, et rien n'assure alors que cet équilibre sera celui du plein emploi (où toute la population active disponible est embauchée).

Dès lors, les politiques macroéconomiques conjoncturelles (monétaire et budgétaire) agissent à court terme afin de réduire la composante « keynésienne » du chômage (associée à une demande globale insuffisante).

En 2007, face à la crise financière puis la « Grande Récession », les **politiques budgétaires de relance** ont augmenté les dépenses publiques dites « passives » pour soutenir l'assurance chômage (indemnisation) et améliorer la situation de l'emploi, avec des dépenses dites « **discretionnaires** » d'inspiration keynésienne. Les entreprises ont pris elles-mêmes des mesures pour

sauvegarder l'emploi, en réduisant la durée du travail, en développant le chômage partiel, avec ou sans l'aide de l'État. Les gouvernements ont ainsi développé leurs aides au retour à l'emploi, mis en place des programmes de créations d'emplois dans le secteur public, soutenu les revenus des travailleurs, soutenu le secteur de la construction (travaux publics) ou agi sur la fiscalité des entreprises.

12.4.2 Les politiques économiques structurelles et l'emploi

Les politiques économiques structurelles, qui produisent leurs effets à moyen et long terme, agissent au niveau microéconomique sur les multiples modalités de fonctionnement du marché du travail, afin de réduire les composantes « classique » et « structurelle » du chômage. À plus **long terme**, les institutions économiques internationales comme l'OCDE jugent que les réformes du marché du travail ont contribué aux niveaux d'emploi élevés qui prévalaient avant l'éclatement de la crise : elles préconisent que les aides ciblées mises en place durant la crise cèdent progressivement la place à une **action plus structurelle**.

12.4.2.1 Les politiques ciblées sur la demande de travail

Les dépenses en faveur de l'emploi et du marché du travail : ciblées ou générales

On peut distinguer les dépenses ciblées ou générales sur le marché du travail

Les **dépenses « ciblées »** concernent les demandeurs d'emploi et les personnes en difficulté sur le marché du travail. Il existe trois catégories de dépenses « ciblées » :

- les services relatifs au marché du travail : dépenses pour les services publics de l'emploi (catégorie 1).
- les mesures « actives » : de formation professionnelle, d'incitations à l'emploi, d'emploi protégé et réadaptation, de création directe d'emplois, d'aide à la création d'entreprises (catégories 2 à 7).

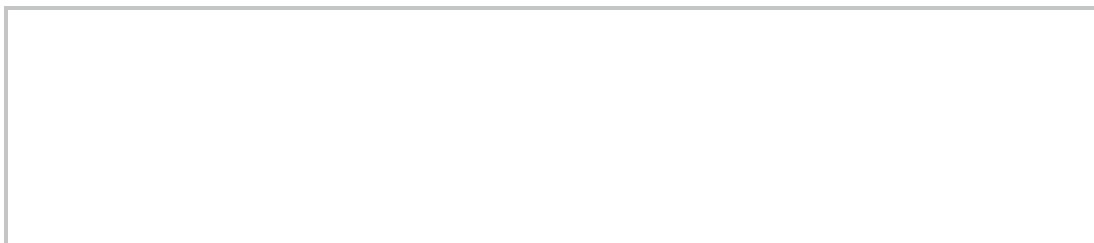
- le soutien aux revenus : maintien et soutien du revenu en cas d'absence d'emploi, préretraite (catégories 8 et 9).

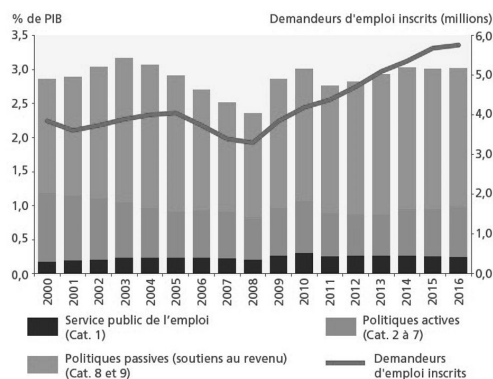
Les **dépenses « générales »** concernent la réduction du coût du travail dans certains secteurs, territoires et catégories de salariés. Il existe quatre catégories de dépenses « générales » :

- les mesures générales d'exonération dont les allègements généraux de cotisations sociales patronales sur les bas salaires, les allègements en faveur des heures supplémentaires, les aides à l'embauche aux TPE.
- les mesures d'incitation financière à l'emploi afin de rendre plus rémunérateur l'accès ou le retour à l'emploi ainsi que l'augmentation de la durée travaillée par des aides complémentaires aux revenus d'activité dont le RSA activité seul + aides personnalisées de retour à l'emploi, le revenu supplémentaire temporaire d'activité.
- les aides à l'emploi dans certaines zones géographiques dont les exonérations spécifiques aux DOM, les aides pour développer l'emploi dans des zones cibles de la politique de la ville ou de la politique d'aménagement du territoire, les aides à l'emploi dans certains secteurs d'activité.

Depuis le début des années 1990, les pouvoirs publics ont donné la priorité aux politiques du marché du travail ayant pour objectif d'agir sur la demande de travail des entreprises. Il en existe trois grandes catégories :

- les allègements de cotisations sociales sur les bas salaires
- les emplois aidés accompagnés de dispositifs ciblés pour lutter contre le chômage discriminant
- les stratégies de réduction du temps de travail.





Champ : France entière.

Source : Dares, base PMT pour Eurostat ; STMT pour les demandeurs d'emploi, catégories A, B et C (données CVS).

GRAPHIQUE 12.7.

Évolution des dépenses ciblées pour les politiques du marché du travail en % du PIB et évolution du nombre de demandeurs d'emploi inscrits

■ Les allègements de cotisations sociales sur les bas salaires

Pour aider une personne sans emploi à retrouver un emploi, un premier type d'action des pouvoirs publics consiste à agir sur la demande de travail. L'objectif est de **favoriser la création d'emplois** : en période de ralentissement économique et de contraintes de débouchés, la politique keynésienne de relance peut soutenir la demande globale à court terme (comme en 2009), mais les politiques « actives » de l'emploi ont pour objectif de stimuler la demande de travail des entreprises à long terme en agissant directement sur le coût du travail et sur la réduction des prélèvements fiscaux et sociaux. L'hypothèse théorique sous-jacente à ce type de politiques est celle d'une sensibilité de la demande de travail à son **coût**. Ainsi, dans de nombreux pays développés, les jeunes et les travailleurs les moins qualifiés ont des taux de chômage supérieurs au taux de chômage moyen de la population active. Selon la théorie économique traditionnelle de la demande de travail, cela peut résulter d'un niveau trop élevé de leur salaire *relatif* et des cotisations sociales de ces catégories de main-d'œuvre, au regard de leur productivité marginale.

Sur le marché du travail, le **salaire relatif** d'une catégorie de main-d'œuvre devrait s'ajuster en fonction des variations de son offre (déterminée par la démographie, l'évolution des taux d'activité, l'immigration) et de sa demande relatives (liée à la structure de la demande de biens, à l'intensité en travail non qualifié des importations de biens des pays émergents et au rythme du progrès technique). Si les salaires sont parfaitement flexibles, ils doivent pouvoir s'ajuster en fonction des évolutions de l'offre et de la demande de travail pour les différentes catégories de main-d'œuvre, afin de rétablir l'équilibre sur tous les marchés du travail (travail non qualifié/travail qualifié).

Cet ajustement peut s'opérer à deux niveaux :

- **au niveau microéconomique** : la réduction du coût relatif de la main-d'œuvre non qualifiée peut par exemple inciter l'entreprise à modifier sa combinaison productive (substitution du travail non qualifié au capital ou au travail qualifié) ;
- **au niveau macroéconomique** : la réduction du coût relatif de la main-d'œuvre non qualifiée peut entraîner une augmentation de la production de biens qui nécessitent une utilisation plus intensive de cette catégorie de travail dont le prix a baissé.

Ces politiques d'allègement de cotisations sociales cherchent à **élever le contenu en emplois de la croissance économique**. Les travailleurs non qualifiés sont ainsi très présents dans quelques secteurs de l'industrie (textile, cuir, habillement), ou dans un certain nombre de services (commerce de détail, hôtellerie-café-restauration), soit des secteurs où la demande de travail peut être très réactive à son prix (les économistes évoquant une forte élasticité de la demande de travail à son coût). C'est la raison pour laquelle, à partir des années 1990, de nombreuses politiques de l'emploi en France ont privilégié **la réduction du coût du travail**, et ce après la publication de nombreux rapports sur la question, en agissant par une diminution des cotisations sociales sur les salaires de la main-d'œuvre non qualifiée (exonérations de charges sociales pour les employeurs au voisinage du SMIC).

Pacte de responsabilité : CICE et réduction d'impôts pour les entreprises

À la suite du Rapport Gallois publié fin 2012, le Président de la République a décidé de privilégier une politique d'offre, basée sur la baisse de la fiscalité et des cotisations sociales sur les entreprises, afin de faire baisser le chômage et

d'améliorer la compétitivité de l'économie française dans la zone euro, en particulier. Le **pacte de responsabilité** a été présenté par le Président de la République en janvier 2014. Il **englobe** le crédit d'impôt compétitivité emploi (**CICE**), des **baisses de cotisations patronales**, et des **mesures de réduction fiscale pour les entreprises**. Le dispositif doit représenter en 2017, un coût annuel de 41 milliards d'euros pour l'État. En échange de ces aides, les partenaires sociaux (patronat et syndicats) doivent mener des négociations sur l'emploi dans les branches professionnelles. Cependant, aucun engagement chiffré et contraignant n'a été imposé aux entreprises dans la loi. Selon l'Observatoire français des conjonctures économiques, le CICE aurait permis de créer et/ou de sauvegarder environ 120 000 emplois. Effet d'aubaine oblige, les entreprises en auraient partiellement profité pour augmenter les salaires, et non pour augmenter les embauches, contrairement à l'objectif initial.

Concrètement, cela a donné lieu à la mise en place du CICE, un crédit d'impôt égal à 6% de la masse salariale correspondant aux salaires de moins de 2,5 SMIC, et du Pacte de responsabilité, correspondant à une baisse de cotisations sociales patronales pour les salaires compris entre 1 et 3,5 SMIC, ainsi qu'une baisse de la fiscalité sur les entreprises.

Les allègements de cotisations sociales sont dégressifs. En effet, le montant de la baisse se réduit à mesure que le niveau de salaire versé s'éloigne du niveau du SMIC. Cette mesure, comme toutes celles menées sur la base de cet instrument depuis 1993, vise à augmenter l'emploi des travailleurs peu ou non qualifiés et à réduire le chômage de longue durée, noyau dur difficilement compressible du chômage.

Réforme du marché du travail : les effets des ordonnances Macron sont à venir (?)

Le 22 septembre 2017, les ordonnances Macron « consacrent un engagement de campagne important et une

réforme en profondeur inédite du marché du travail indispensable à notre économie et notre société », selon le Chef de l'État. Cette réforme par ordonnance devait se traduire par plus de flexibilité pour les entreprises dans la gestion de leur personnel et par davantage de dialogue social : faciliter le licenciement pour faciliter l'embauche. Objectif : faire baisser durablement le chômage grâce à un marché du travail moins rigide.

Un an plus tard, le succès n'est pas (encore) au rendez-vous. La flexibilité dans un contexte macroéconomique fragilisé par une croissance trop molle, et des anticipations pas assez optimistes, n'incite pas à la création massive d'emplois. Quand les carnets de commandes ne sont pas suffisamment remplis, dans le présent et dans l'avenir, quand la demande globale n'est pas assez dynamique, l'embauche ne démarre pas vraiment : c'est un enseignement keynésien basique.

■ **Les contrats aidés : l'aide au retour à l'emploi**

Les contrats aidés font également partie de la panoplie des **politiques « actives »** du marché du travail : il s'agit d'emplois sur des contrats de travail spécifiques, dérogatoires au droit du travail. Ces contrats offrent aux employeurs des avantages ou des incitations particulières (subventions, exonérations de cotisations sociales pour l'employeur, aides à la formation) et cherchent à stimuler la demande de travail des entreprises pour certaines catégories de travailleurs qui, sans cela, resteraient éloignés durablement du marché du travail.

La politique de l'emploi récente constitue en effet une source de **diversification des formes du contrat de travail** : ces contrats aidés visent à diminuer par des aides directes ou indirectes les coûts d'embauche et de formation des personnes en difficulté sur le marché du travail pour faciliter leur insertion professionnelle (ou cherchent à doter les individus qui en sont dépourvus de qualifications qui améliorent leur « employabilité »). Ces contrats aidés sont donc généralement ciblés sur les jeunes (comme les contrats

d'apprentissage ou de professionnalisation), les chômeurs de longue durée (« contrat initiative emploi »), ou les titulaires de minima sociaux.

Ces contrats sont destinés soit au **secteur non marchand** (administrations, collectivités territoriales ou hospitalières, associations avec les travaux d'utilité collective – TUC –, les « contrats emploi solidarité » – CES – puis les « contrats d'avenir ») ou au **secteur marchand** des entreprises privées. Les exemples abondent depuis le début des années 80. Pourtant, force est de constater que la multiplication et le cumul de ces dispositifs ne se sont pas accompagnés d'une forte croissance de l'emploi.

Flexibilité et dualisme accrus sur le marché du travail

Les syndicats de salariés voient dans la multiplication de ces contrats un accroissement de la **flexibilité** sans garantie de **sécurité** des travailleurs. Ces nouveaux types de contrats auraient pour effet d'accroître la précarité des salariés : le droit du travail français, construit progressivement tout au long du XX^e siècle et constitutif de la norme d'emploi (fondée sur le CDI), serait menacé par cette multiplication des contrats aidés qui dégageraient par ailleurs les entreprises de leur responsabilité sociale. En outre, cette multiplication des contrats « atypiques » aurait pour effet d'accroître la **segmentation du marché du travail** entre, d'une part, des salariés en CDI bien protégés par le droit et, d'autre part, des salariés précaires et faiblement rémunérés. Ces contrats aidés font par ailleurs l'objet de deux séries de **critiques** : leur coût est jugé prohibitif pour les finances publiques (en raison du manque à gagner pour le budget de l'État à cause des exonérations fiscales) et il peut créer des **effets d'aubaine**, puisque les entreprises sont incitées à substituer ces travailleurs à d'autres embauches qui auraient de toute façon été réalisées, afin de bénéficier des aides de l'État. Par ailleurs, les titulaires de ces contrats peuvent être victimes d'un **effet de stigmatisation** si les employeurs l'interprètent

comme un signal négatif sur les compétences et la motivation du travailleur.

■ Le partage du temps de travail et de l'emploi

Les politiques de réduction du temps de travail : une « règle de trois »

Les relations entre la croissance économique, la productivité du travail, l'emploi et le temps de travail sont complexes : les gains de productivité entraînent à la fois une augmentation de la croissance potentielle, une destruction d'emplois à court terme et un bouleversement de l'organisation du travail et des qualifications. La productivité du travail favorise donc la croissance de la production à long terme et crée des besoins en emplois, mais elle incite également les entreprises à la substitution du capital au travail et à des **réallocations d'emplois** dans le système productif. Lorsque la productivité du travail augmente, cela signifie que les entreprises sont en mesure de produire autant avec une quantité de travail inférieure. C'est la raison pour laquelle, malgré la forte hausse de la production, la France utilise moins de travail qu'il y a un siècle : en effet, la productivité augmente plus vite que la production sur longue période. Toutefois, même si la quantité de travail diminue, l'emploi ne diminue pas pour autant : l'arbitrage dépend de la clé de **répartition des gains de productivité**, entre la hausse du niveau de vie, la réduction du temps de travail et la réduction du taux d'emploi.

L'évolution de l'emploi dépend alors de trois grands facteurs : la croissance de la production, de la productivité du travail et de la durée du travail. À long terme, la durée du travail a nettement diminué dans les économies développées, ce qui a favorisé la croissance de l'emploi : lorsque les gains de productivité sont élevés, il est possible de réduire à la fois la durée du travail et d'augmenter les revenus ; mais, en période de gains de productivité faibles, la stagnation des revenus freine les possibilités de réduction du temps de travail. Cette relation mécanique est cependant contestée, car le **partage du travail**, s'il peut constituer une théorie

séduisante (« travailler moins pour travailler tous »), n'en constitue pas moins un choix malthusien.

En effet, on parle souvent de « **règle de trois** » pour exprimer l'idée que, pour un niveau de production et de productivité horaire du travail donné, une baisse de la durée du travail augmente le volume de l'emploi nécessaire d'une quantité équivalente.

Mais ces trois variables, isolées dans l'équation macroéconomique, sont en réalité **interdépendantes** : le niveau de la productivité horaire est lié à la durée du travail et à la production également. En effet, la théorie économique considère que la productivité marginale du travail est décroissante : une baisse de la durée du travail peut alors stimuler la productivité horaire et freiner les besoins en création de nouveaux emplois (avec une **intensification** du travail). Les effets sont toutefois incertains, car on peut considérer à l'inverse qu'une durée du travail élevée peut avoir un impact positif sur la productivité horaire en renforçant l'expérience et l'efficacité des travailleurs.

De plus, la réduction du temps de travail diminue les **revenus individuels** des ménages :

- il faut alors mettre en œuvre une **compensation salariale** (généralement financée au moins partiellement par des aides financières de l'État), car sinon, la consommation des ménages peut faiblir et réduire l'emploi (baisse de la demande globale) ;
- de la même manière, pour une production constante, une réduction du temps de travail peut conduire à une élévation du salaire horaire et donc du **coût horaire du travail** pour les entreprises (car les salaires mensuels sont maintenus à leur niveau intérieur) ; l'ampleur de l'effet négatif sur l'emploi dépend alors de l'élasticité de la demande de travail à son coût. Répercutée intégralement sur les prix, cette hausse du coût horaire de travail peut détériorer la **compétitivité-prix** des entreprises et nuire à l'emploi dans la concurrence mondiale. Les politiques de

réduction du temps de travail s'accompagnent généralement d'une réorganisation du travail et d'un allongement de la durée d'utilisation des équipements (de manière à stimuler la productivité du capital).

La réduction du temps de travail (RTT) en France

Durant les années 1980 et 1990, la production a augmenté moins vite que la productivité du travail en France : en conséquence, les heures travaillées ont diminué fortement. Si les périodes de chômage élevé favorisent les politiques de réduction du temps de travail, au nom de considérations sociales liées au partage du travail (loi sur les 40 heures à l'époque du Front populaire, loi sur les 39 heures et cinquième semaine de congés payés en 1982, lois Aubry sur les 35 heures à la fin des années 1990, 1998 et 2000), la baisse de la durée du travail est une tendance séculaire en France et dans les économies développées (la durée du travail a été approximativement réduite de 50% de 1830 à 2011).

La durée du travail était ainsi approximativement de 2500 heures au début du xx^e siècle, et elle a constamment diminué, sous l'effet à la fois de la baisse légale de la durée hebdomadaire de travail, des congés payés et du développement du travail à temps partiel, pour atteindre autour de 1500 heures travaillées en France en 2013. Si la progression du travail à temps partiel (majoritairement contraint) explique largement cette tendance depuis les années 1980, en raison du sous-emploi persistant, une partie des gains de productivité a donc été affectée de manière volontariste à la réduction du temps de travail au titre des politiques sociales. À l'heure actuelle, la France constitue un pays (avec d'autres) où le nombre d'heures travaillées est relativement faible, mais où la productivité horaire du travail est parmi les plus élevées au monde. En Allemagne par exemple, la réduction du temps de travail a été largement due à l'essor du temps partiel. Ainsi, si les dernières lois Aubry ont créé un certain nombre d'emplois (environ 350 000), elles n'ont pas eu les effets escomptés sur la baisse du taux de

chômage, et leurs effets ont été contrés en particulier par l'augmentation de la productivité horaire du travail, car les entreprises n'ont pas embauché à proportion de la réduction du temps de travail et ont bien souvent intensifié le travail pour les salariés.

Le temps de travail et l'emploi ne se partagent pas

Des voix s'élèvent régulièrement pour revendiquer davantage de « partage du travail », afin de lutter contre le chômage, de favoriser le développement durable par le ralentissement de la croissance ou même la décroissance. D'autres, en lien avec les réformes des systèmes de retraites et le vieillissement démographique, voient dans le partage du travail une solution aux problèmes d'emploi d'une partie de la population en âge de travailler. Concrètement, l'emploi étant une denrée rare, pour que tous ceux qui sont en âge de travailler en occupent un, il faut partager l'emploi global entre un nombre plus grand d'individus.

Partager l'emploi, donc le travail, cela signifie concrètement que :

- les « vieux » qui travaillent plus longtemps, c'est moins d'emplois et plus de chômage pour les « jeunes » ;
- les femmes qui travaillent, c'est moins d'emplois et plus de chômage pour les hommes ;
- les immigrés-étrangers qui travaillent, c'est moins d'emplois et plus de chômage pour les Français ;
- les plus qualifiés qui travaillent, c'est moins d'emplois et plus de chômage pour les moins qualifiés.

Donc, dans ces conditions, pour réduire le chômage des uns, il faut réduire le temps de travail des autres :

- pour réduire le chômage des « jeunes », il ne faut pas faire travailler les « vieux » plus longtemps ;
- pour réduire le chômage des femmes, il faut que les hommes acceptent de partager ;
- pour réduire le chômage des hommes, il faut renvoyer les femmes au foyer ;

- pour réduire le chômage des immigrants, il faut que les Français partagent ;
- pour réduire le chômage des Français, il faut renvoyer les immigrants chez eux ;
- pour réduire le chômage des moins qualifiés, il faut que les qualifiés partagent...

Toutes ces affirmations semblent frappées au coin du bon sens, pourtant elles ne tiennent pas la route, car il y a souvent un abîme en économie entre le « bon sens » courant et la réalité économique, parfois beaucoup moins intuitive, mais bien mieux vérifiée par les observations statistiques d'une population dans le temps.

Le travail, donc l'emploi, ne se partage pas, tout simplement parce que l'emploi des uns, bien loin de détruire l'emploi des autres, en favorise l'essor. L'emploi global n'est pas un gâteau de taille fixe, un nombre de postes ou un nombre d'heures donné, qu'il convient de partager en un plus grand nombre de parts pour que chacun en ait une.

Si l'on regarde les évolutions statistiques à très long terme, on voit qu'avec l'augmentation de la taille d'une population totale sous l'effet de la multiplication des naissances ou sous l'impact d'une vague d'immigration, la population en âge de travailler augmente aussi. *Ipsa facto*, la population active (une partie de la population en âge de travailler) croît. Si l'emploi augmente au moins aussi vite que la population active, alors le nombre de personnes occupées par un emploi augmente et le nombre de personnes au chômage diminue. C'est bien ce que l'on observe dans tous les pays qui se développent depuis deux siècles. La population totale augmente, la population en âge de travailler augmente et l'emploi augmente aussi.

Ce qui fait que le chômage augmente, c'est l'insuffisante croissance de l'emploi relativement à celle de la population active (qui occupe un emploi ou qui est sans emploi, mais en cherche activement un – personnes au chômage). On comprend bien que, globalement, sous les effets

démographiques présentés, l'emploi se développe. Donc l'emploi des nouveaux venus ne se fait pas au détriment de ceux qui étaient déjà présents. De nouveaux emplois se créent pour absorber cet afflux de main-d'œuvre nouvellement disponible.

Comment expliquer que plus de main-d'œuvre disponible s'accompagne, globalement, de plus d'emplois, de plus d'heures de travail ?

C'est le résultat du processus de destruction créatrice, qui est un processus permanent d'innovation et qui va codéterminer les processus de croissance économique (augmentation soutenue et durable de la production globale de l'économie) et de croissance de l'emploi. C'est également la capacité de l'économie à fournir à cette nouvelle main-d'œuvre le capital nécessaire et suffisant à son emploi, c'est-à-dire à lui donner la possibilité de participer à la production du gâteau global de biens et services destinés à satisfaire les besoins de tous, le gâteau PIB, qui va déterminer la croissance de l'emploi et celle de l'économie. Chaque fois qu'une entreprise se développe, elle doit investir pour pouvoir embaucher (et réciproquement) une fois que le taux d'utilisation de ses capacités de production atteint son maximum.

L'interdépendance entre les deux facteurs de production permet de comprendre ce phénomène d'accroissement quasi simultané de l'embauche et de l'investissement. À long terme, l'emploi et la production augmentent bien sous l'effet de la croissance de la population totale, de la population en âge de travailler, de la population active ; et l'emploi des nouveaux venus s'ajoute à celui des anciens déjà présents. Dans ces conditions, plus de travail pour les vieux, les femmes, les immigrés, ne peut se traduire par moins de travail pour les jeunes, les hommes, les Français.

Les gains de productivité et la RTT

Certes, les gains de productivité obtenus grâce à l'investissement et à l'innovation permettent de produire plus en moins de temps. Ils permettent donc de réduire le temps

de travail par personne occupée (en âge de travailler), car chacune produit plus en moins de temps. Cela doit donc permettre de favoriser l'embauche d'un plus grand nombre de personnes par la baisse de la durée de travail par travailleur.

En effet, sous l'effet de la hausse de la productivité horaire, on observe bien à long terme (plusieurs siècles et décennies) une forte baisse du temps de travail par personne occupée. Mais, dans ce cas, il ne faut pas se tromper sur le sens de la causalité. Ici, c'est la croissance de la productivité qui permet, sur la longue durée, de réduire la durée du travail par emploi.

De surcroît, simultanément, cette hausse de la productivité s'accompagne d'une hausse du pouvoir d'achat de la demande (sous le double effet de la hausse des salaires et de la baisse des prix) et d'une hausse de la capacité d'offre (sous le double effet de la baisse des coûts unitaires et de la hausse des profits). Il en découle donc une croissance simultanée de la production et de l'emploi. Les gains de productivité horaire ont certes augmenté, la durée du travail par tête est plus faible, mais, comme il a fallu répondre à une hausse de la demande, il a été nécessaire d'embaucher davantage, ce qui permet d'absorber davantage d'actifs, de fournir un emploi à plus de personnes actives.

Il ne faut surtout pas se tromper sur le sens de la causalité entre gains de productivité et réduction de la durée du travail. On en déduit qu'il n'y a rien de comparable avec l'enchaînement de causalité relatif à la RTT, au sens des 35 heures en France. Dans ce cas, le passage aux 35 heures (la RTT) n'est pas une conséquence des gains de productivité à long terme, mais un moyen supposé d'augmenter le nombre d'emplois en partageant l'insuffisance du nombre d'heures de travail disponibles. L'hypothèse est que le nombre d'heures est donné et que, pour faire travailler davantage de monde, il faut réduire la durée du travail par tête. Une baisse du nombre de chômeurs en résultera. Cette vision malthusienne de l'emploi consiste tout simplement à se résigner, à partager la pénurie. Certes, la RTT (35 heures) a permis d'obtenir de faibles gains de productivité liés à la

nécessité pour les entreprises de réorganiser le travail de leurs salariés pour réduire le coût du passage de 39 à 35 heures.

Mais, bien évidemment, cela n'a pas permis d'inscrire la courbe des créations d'emplois sur une tendance durable à la hausse et celle du nombre de chômeurs sur une tendance durable à la baisse. L'emploi et les revenus des uns, encore une fois, ne se créent pas au détriment de ceux des autres. Le partage du travail, sous toutes ses formes, comme solution aux diverses difficultés d'emploi de populations différentes, n'est pas efficace à long terme. On peut admettre que transitoirement, à court terme, cela peut favoriser l'emploi, mais ce n'est qu'un pis-aller et une chimère à long terme.

À long terme, du fait des dizaines de milliers de créations et destructions quotidiennes d'emplois, d'activités, d'entreprises, en raison du progrès technique, des innovations, c'est de notre capacité de création de nouveaux emplois, de nouvelles activités, de nouvelles entreprises que dépend la création nette d'emploi. Si cette création nette est positive, cela veut dire qu'on crée plus qu'on ne détruit et le chômage baisse, les taux d'emploi et d'activité augmentent. Si cette création nette est négative, cela veut dire qu'on crée moins qu'on ne détruit et le chômage augmente, les taux d'emploi et d'activité diminuent. De surcroît, en raison de ce processus de destruction créatrice source de progrès (même si le progrès peut poser des problèmes environnementaux..., car rien n'est complètement positif), il est tout à fait logique qu'à court terme, l'emploi des uns se crée au détriment de celui des autres emplois frappés par l'obsolescence. Pourtant, il ne faut pas se tromper de diagnostic et de traitement. La meilleure chose à faire, à court et à long terme, n'est pas de défendre des activités et des emplois condamnés à disparaître inéluctablement, mais plutôt de mettre en place un accompagnement des nouveaux chômeurs vers des formations et des emplois d'avenir.

12.4.2.2 Les politiques fondées sur l'offre de travail

Les politiques de l'emploi les plus récentes ne cherchent pas uniquement à agir sur la demande du travail, et l'objectif n'est plus seulement d'obtenir une croissance économique plus forte pour favoriser la création de nouveaux postes de travail, ou d'obtenir une croissance économique plus riche en créations de nouveaux emplois. Il s'agit d'agir sur l'offre de travail pour **augmenter la participation** des individus au marché du travail. Il s'agit de mettre en place, d'une part, des incitations monétaires (financières) au retour à l'emploi (pour augmenter les revenus du travail, réduire le montant des allocations chômage) et d'autre part, des dispositifs de retour à l'emploi, de formation professionnelle pour favoriser l'appariement entre l'offre et la demande de travail (incitations non monétaires).

Ces stratégies reposent sur l'idée qu'une des causes du chômage de masse réside dans l'inadéquation des qualifications des travailleurs aux emplois proposés (chômage structurel et chômage frictionnel), et qu'une autre cause est liée au fait que les rémunérations du travail et les incitations financières sont trop faibles pour créer un **écart incitatif** suffisant avec les revenus de remplacement (prestations sociales) versés à ceux qui sont dépourvus d'emploi. Dans cette logique, les aides financières au retour à l'emploi, le soutien au revenu des travailleurs qui retrouvent un emploi ou encore les formations proposées aux demandeurs d'emploi (essentiellement) constituent l'arsenal des mesures centrées sur l'offre de travail.

– Les mesures d'incitations au retour à l'emploi

Incitation par la hausse des revenus d'activité

L'activation des politiques du marché du travail a constitué une tendance forte des réformes en faveur de l'emploi en Europe au cours des dernières décennies, avec un certain nombre de variations. Les **mesures** dites « **d'activation** » des chômeurs prennent tout d'abord la forme d'incitations financières **à la reprise d'emploi**. La dégradation du marché

du travail a renforcé le recours à ce type de dispositifs, comme la prime pour l'emploi en France (PPE), ou l'autorisation de cumuler l'indemnisation du chômage avec une activité à temps partiel et une faible rémunération. C'est par exemple le cas du revenu de solidarité active (RSA) en France, en vertu duquel une part de son montant continue à être perçue en cas de reprise partielle d'activité. De manière générale, il s'agit de revaloriser le travail en désincitant les individus à l'inactivité, et en stimulant les revenus tirés du travail ("**to make work pay**", faire que le travail rapporte). Ces politiques cherchent à augmenter le taux d'emploi (population ayant un emploi/population totale) autant qu'à réduire le taux de chômage (nombre de chômeurs/population active).

Les politiques du marché du travail qui se sont déployées à partir des années 1990 ont privilégié **les incitations monétaires** : il s'agit non seulement d'accorder des subventions aux employeurs pour réduire le coût du travail et favoriser la création d'emplois (voir *supra*), et, dans le même temps, d'offrir des subventions aux travailleurs pour qu'ils occupent les emplois créés et augmenter le gain financier tiré par les salariés.

Début mars 2015, afin d'améliorer les revenus du travail, le gouvernement français a annoncé la mise en place d'une « prime d'activité » pour les bas salaires, un dispositif qui a remplacé à partir du 1^{er} janvier 2016 la prime pour l'emploi (PPE) et le revenu de solidarité active (RSA) dans son volet « activité ». Les bénéficiaires de cette « prime d'activité » (dont les jeunes de 18 à 25 ans) doivent gagner environ 1,2 fois le SMIC et les revenus du ménage ne doivent pas dépasser un certain montant. Le 10 décembre 2018, en réponse à la crise des « gilets jaunes », le Président de la République Emmanuel Macron, a annoncé plusieurs gestes forts, dont une augmentation de 100 euros des salaires au niveau du SMIC (prime d'activité incluse).

« Je demande au gouvernement et au Parlement de faire le nécessaire afin qu'on puisse vivre mieux de son travail dès le début d'année prochaine. Le salaire d'un travailleur au SMIC

augmentera de 100 euros par mois dès 2019 sans qu'il en coûte un euro de plus pour l'employeur. » Cette hausse est rendue possible, par abondement de l'État, en complément de l'augmentation de 1,8% qui devait intervenir en janvier 2019. Cette hausse de 100 euros était prévue, mais sur l'ensemble du quinquennat.

Dans son livre *Rémunérer le travail* (1997), le prix Nobel d'économie Edmund Phelps proposait que le gouvernement américain **subventionne** les entreprises pour favoriser l'emploi et des salaires plus élevés. L'idée sous-jacente est que si l'on tient compte uniquement de la productivité du travail et des forces du marché, certains travailleurs ne peuvent gagner qu'un salaire très bas, peu incitatif. Dans ce cas, beaucoup d'individus peuvent faire le choix de ne pas travailler, préférer percevoir des allocations de chômage, voire recourir à l'économie souterraine. Dès lors, si le gouvernement met en place des subventions aux bas salaires, l'incitation à travailler sera plus grande. Dans ce cadre de pensée, la protection sociale devient un levier en faveur de **l'efficacité économique** et de la compétitivité internationale.

Incitation par la baisse des revenus de remplacement

En vertu de la théorie des incitations en économie, on assigne à la protection sociale la fonction de **modifier les comportements**, et non plus seulement de protéger les individus contre des risques sociaux traditionnels (maladie, chômage, vieillesse sans ressource). Les nouvelles politiques de l'emploi ont pour objectif le retour des individus dans l'emploi, plutôt que le versement d'un revenu de remplacement dans le cadre de politiques dites « passives ». Néanmoins, ces politiques du marché du travail n'ont pas empêché la progression du phénomène des travailleurs pauvres.

- **Les problèmes d'appariement sur le marché du travail et le chômage structurel**

Le chômage structurel résulte d'un déséquilibre durable dans le fonctionnement du marché du travail : il est l'expression d'une inadéquation entre l'offre et la demande travail, c'est-à-dire entre les compétences des demandeurs d'emploi et les compétences exigées par les offreurs d'emploi (entreprises, administratives...).

Les politiques de l'emploi agissent isolément sur la demande et sur l'offre de travail, mais elles cherchent également à faciliter leur rencontre : c'est ce que l'on appelle les politiques d'appariement. La rencontre entre l'employeur et le salarié suppose du temps et des coûts de recherche en raison des imperfections de l'information sur le marché du travail. Afin de faciliter cet appariement, les pouvoirs publics peuvent déployer un service public de l'emploi (ANPE en 1967, Pôle emploi depuis 2008).

Il s'agit de développer des systèmes d'information sur l'existence des postes vacants, afin de fluidifier le marché du travail. On sait par exemple que les jeunes issus des zones urbaines sensibles (ZUS) sont pénalisés par l'éloignement des zones dynamiques créatrices d'emplois, en raison du coût des transports et des pressions à la hausse sur les prix de l'immobilier dans les centres-villes.

Pour réduire ce **chômage de mobilité (ou frictionnel)**, l'État peut alors réduire le coût des transports et mettre en œuvre des politiques d'incitations fiscales aux entreprises pour qu'elles s'implantent dans les zones franches urbaines. Ce chômage frictionnel correspond à la période nécessaire au passage d'un emploi à l'autre et dépend du degré de **mobilité géographique et professionnelle** des individus.

Le **niveau de formation** constitue un critère de différenciation important vis-à-vis du chômage. Il pénalise fortement les jeunes lors de leur entrée sur le marché du travail : quel que soit leur niveau de formation et de diplôme, leur situation tend à s'aggraver par rapport à la fin des années 1970, et le chômage des jeunes est avant tout celui des faiblement diplômés et des non-diplômés.

Depuis le début des années 1980, la **sélectivité** du marché du travail en termes de niveau de qualification a ainsi nettement progressé. Loin d'avoir un caractère statique, le marché du travail est **dynamique** et fonctionne selon un processus de créations et destructions d'emplois, qui peut, en période de mutations technologiques rapides et d'intense concurrence mondiale (qui implique une baisse de la demande relative pour le travail non qualifié et un biais en faveur du travail qualifié au sein des pays développés), entraîner un désajustement entre les qualifications des travailleurs et les emplois proposés.

Il peut alors exister un **chômage d'inadéquation** qui résulte de la coexistence de structures rigides d'offre et de demande de travail par types de qualifications qui ne coïncident pas. Ce type de chômage apparaît lorsque les salariés ont des compétences peu utiles pour les entreprises et ne peuvent disposer rapidement de ces compétences utiles, lorsque les entreprises ont des demandes précises pour lesquelles il n'y a pas de salariés disponibles, et si elles ne peuvent substituer d'autres salariés à ceux qu'elles recherchent. Cette forme de chômage n'est ainsi pas négligeable pour les seniors (ayant des difficultés à adapter leurs qualifications aux exigences du marché) ou pour les chômeurs de longue durée durablement déqualifiés et qui ont perdu de leur « employabilité ».

Faciliter l'appariement : améliorer l'information et la formation professionnelle et continue

Sur le marché du travail, les **difficultés de recrutement** sont plus ou moins fortes selon les métiers. Dans certains secteurs, la main-d'œuvre disponible est rare (ingénieurs en informatique, techniciens de l'industrie mécanique ou de l'énergie). *A contrario*, dans d'autres secteurs d'activité, la main-d'œuvre disponible est abondante (cuisiniers, coiffeurs, esthéticiens, ouvriers non qualifiés du bâtiment...). Les systèmes de formation professionnelle doivent tenir compte de ces situations hétérogènes.

Ainsi, dans les secteurs où la main-d'œuvre disponible est rare (offre de travail très faible relativement à la demande de travail), la formation de la main-d'œuvre s'impose comme une solution incontournable permettant d'adapter les qualifications et compétences des travailleurs potentiels aux exigences des nombreux postes proposés par les entreprises. À titre d'illustration, en France les pouvoirs publics ont mis en place le contrat de professionnalisation (réaménagé à compter du 1^{er} janvier 2019) ; c'est un contrat à durée déterminée ou un contrat à durée indéterminée ayant pour objectif l'insertion professionnelle ou le retour à l'emploi de jeunes ou d'adultes par l'acquisition d'une qualification professionnelle reconnue par l'État et/ou la branche professionnelle. Il s'adresse aux jeunes de 16 ans à 25 ans révolus (26 ans moins un jour) pour compléter leur formation initiale, aux bénéficiaires du RSA, aux bénéficiaires de l'allocation de solidarité spécifique (ASS), aux bénéficiaires de l'allocation aux adultes handicapés (AAH), aux personnes ayant bénéficié d'un contrat unique d'insertion (CUI), aux demandeurs d'emploi d'au moins 26 ans. Cette formule associe l'occupation d'un emploi dans une entreprise et une formation dispensée par un établissement spécifique, qui permet aux personnes peu qualifiées de disposer d'une formation « sur mesure ».

Plus généralement, la réforme de la formation professionnelle et le plan d'investissement associé doivent permettre la mise en place des formations adaptées aux nouvelles compétences apparues avec le processus de destruction créatrice. Cependant, les effets de ces mesures structurelles ne peuvent être que différés et progressifs.

Pour la seconde catégorie de métiers où la main-d'œuvre est moins qualifiée et pour laquelle les conditions de travail sont plus difficiles (contraintes physiques, tâches répétitives...), l'instabilité de l'emploi (recours aux CDD, à l'intérim...) est élevée. Le taux de chômage structurel pourra rester à un niveau durablement haut.

Dans ce deuxième cas de figure, la formation n'est pas la mesure prioritaire. Il convient, d'abord, de rendre les

travailleurs plus mobiles, de faciliter l'embauche sous contrats courts, mais, simultanément, de renforcer l'accompagnement afin de rendre ces emplois plus attractifs et moins instables.

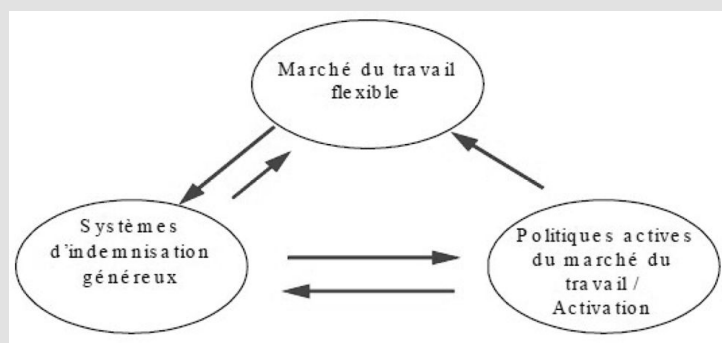
12.5 LA FLEXICURITÉ : POUR UN ÉQUILIBRE ENTRE FLEXIBILITÉ DE L'EMPLOI ET SÉCURISATION DES REVENUS ET DES PARCOURS PROFESSIONNELS

La dynamique économique se traduit, *ipso facto*, par de nombreuses destructions et créations d'emplois, d'activités, d'entreprises, chaque jour, chaque semaine, chaque mois, chaque année. On estime qu'en France, en général 10 000 emplois sont créés et détruits en moyenne par jour. À court terme, en période de récession, il y a davantage de destructions, mais moins de créations. Le chômage augmente. À long terme, on explique ce processus de destruction créatrice d'activités, d'entreprises, d'emplois... par l'innovation (et la mondialisation). Dans un tel environnement, pour faire face et survivre à la concurrence, une entreprise doit pouvoir s'adapter, soit par une stratégie de parts de marché à court terme (jeu à somme nulle), soit par une stratégie de croissance par l'innovation à long terme (accroissement de la taille globale du marché).

D'avantage de souplesse, de flexibilité, semble indispensable pour que les entreprises puissent éviter la faillite. Pourtant, plus de flexibilité peut nuire aux salariés. En effet, on observe une réelle perte de sécurité en matière d'emploi, de revenu, de solvabilité, de retour à l'emploi. La flexibilité semble s'accompagner de davantage de précarité. Se pose alors la question la suivante : est-il possible de gérer efficacement cette incompatibilité apparente entre le besoin de flexibilité des entreprises et celui de sécurité des salariés ?

C'est une question à laquelle Jérôme Gauthié a tenté d'apporter des réponses en 2006. Pour cet économiste, spécialiste reconnu du marché du travail, la réponse est

affirmative dans la mesure où le modèle de flexicurité est adapté aux particularités de l'économie et de la société françaises. Depuis plusieurs décennies, le modèle de flexicurité de référence est le système danois d'organisation du marché du travail et des relations entre partenaires sociaux, car il semble apporter une réponse équilibrée à ce dilemme. La flexicurité correspond à une combinaison harmonieuse des deux exigences, devenues complémentaires, consistant à protéger les salariés, sans protéger l'emploi, dans le cadre d'un marché du travail plus flexible pour les entreprises. Pour y parvenir, il est impératif de combiner : un marché du travail plus flexible en matière de licenciement (et donc d'embauche), un système d'indemnisation du chômage plus généreux, et une politique plus active d'aide au retour à l'emploi (recherche, appariement et formation) grâce à un service public de l'emploi très efficace et adapté.



GRAPHIQUE 12.8.

Le triangle d'or de la flexicurité

Jérôme Gautié, « La flexicurité : nouvel horizon du modèle social français ? », Cahiers français n° 330, consacrés au « modèle social français » (La Documentation française, 2006)

12.6 CROISSANCE ET CHÔMAGE : CYCLE DE PRODUCTIVITÉ ET VITESSE D'AJUSTEMENT DE L'EMPLOI

La nature et l'ampleur des chocs économiques depuis 2008 ont provoqué une récession dans de nombreuses économies occidentales, notamment. Les ralentissements observés de la croissance économique dans différents pays, a fortiori les baisses du PIB, se sont logiquement accompagnés d'un ajustement de l'emploi aux besoins de l'économie dans son ensemble. Mais, si l'emploi a été très touché, durant la crise, selon les pays et les secteurs d'activité, son évolution n'a pas toujours reflété celle de la croissance du PIB, en diminuant plus ou moins que prévu.

12.6.1 *Vitesse d'ajustement de l'emploi et cycle de productivité*

Ce décalage entre variations du taux de croissance du PIB et variations de l'emploi, en réponse à ces variations de la croissance peut se comprendre en observant le cycle de productivité.

En effet, la relation de causalité entre croissance économique et croissance de l'emploi, que ce soit en phase d'accélération ou de ralentissement de la croissance (voire de baisse du PIB), est très étroitement liée à l'évolution de la productivité par travailleur, à court terme.

Ainsi, en période de ralentissement de la croissance du PIB, les entreprises licencient davantage et embauchent moins. Pour autant, l'ajustement de l'emploi au ralentissement de la croissance économique, ou à la baisse du PIB n'est ni immédiat ni strictement proportionnel. Pourquoi ?

Tout simplement en raison de ce que les économistes appellent le « cycle de productivité ».

Ce décalage temporel entre variations de l'emploi induites par les variations de la croissance du PIB, s'accompagne donc

d'un cycle de la productivité par tête ou productivité par travailleur.

Le cycle de productivité sera d'autant plus marqué que la vitesse d'ajustement de l'emploi sera faible et que le cycle de productivité sera d'autant moins marqué que la vitesse d'ajustement de l'emploi sera forte. Pour le dire autrement, plus l'emploi s'ajuste rapidement et plus que proportionnellement aux variations de la croissance et moins la productivité par tête augmente (en période de ralentissement) ou baisse en phase de reprise. *A contrario*, plus l'ajustement de l'emploi sera lent et faible relativement aux variations de la croissance (par exemple, une récession) et plus la productivité par tête sera variable (baissera quand on maintiendra en emploi de nombreuses personnes alors que la demande et la production des entreprises seront plus faibles, et augmentera quand on favorisera les destructions d'emplois en réponse aux fluctuations de la croissance).

12.6.2 *La relation entre croissance économique, productivité et emploi*

On peut comprendre ce point à partir d'une arithmétique simple :

Soit la productivité apparente par travailleur (PAT) et TV le taux de variation

On parle de productivité apparente car la productivité d'un travailleur dépend aussi du capital productif qu'il utilise. On ne peut donc pas lui attribuer l'ensemble des gains de productivité.

PAT = Production / Emploi

Donc, Emploi = Production / PAT

En variation, on obtient,

TV Emploi = TV Production - TV PAT

Ou

TV PAT = TV Production - TV Emploi

On déduit (induit) donc clairement de ces simples équations qu'une accélération de la croissance économique (le taux de croissance du PIB augmente) entraînera une augmentation d'autant plus forte de l'emploi (l'ajustement de l'emploi sera important), que le taux de croissance de la productivité par tête sera faible.

Pour approfondir, à partir de relations arithmétiques, en décomposant la productivité par tête en nombre d'heures travaillées et productivité horaire, on peut observer le comportement de marge des entreprises à travers les variations des coûts salariaux unitaires réels, lesquelles dépendent du rapport entre les variations de salaires réels par tête et les gains de productivité.

Cela permet de commencer à entrevoir les impacts que peuvent avoir les différents degrés d'ajustement de l'emploi aux fluctuations de la croissance, sur la productivité horaire, sur les coûts salariaux unitaires réels, sur le comportement de marges et de prix des entreprises ...

À court terme, cette relation entre croissance du PIB, croissance de la PAT et croissance de l'emploi, permet d'expliquer que dans une économie frappée par une grave récession, au sens où le PIB va baisser pendant au moins deux trimestres consécutifs comme en 2008-2009 :

- Les économies dans lesquelles l'emploi s'ajuste avec lenteur (faible flexibilité du marché du travail) vont voir la productivité du travail chuter dans un premier temps et le chômage augmenter lentement. En revanche, dans un second temps (période de reprise), l'ajustement retardé de l'emploi se traduira d'une part par une remontée de la PAT avec la forte hausse du chômage. Autrement dit, le déficit excessif et transitoire de productivité annonce une hausse future du chômage et une hausse de la productivité.
- Les économies dans lesquelles l'emploi s'ajuste avec rapidité (forte flexibilité du marché du travail) vont voir la productivité du travail rester plus stable, ou baisser plus

modérément, dans un premier temps, et le chômage va augmenter fortement...

Logiquement, l'ampleur et la vitesse d'ajustement de l'emploi (et celles des variations de la productivité) en début de crise vont impacter les ajustements en sortie de crise.

Ainsi, lors de la grande récession de 2009, on a pu observer d'importantes divergences de comportement en matière d'ajustement de l'emploi dans les différents pays :

« – l'Allemagne, l'Italie, le Royaume-Uni (dans le secteur marchand) ont sous-ajusté l'emploi

– le Danemark, l'Espagne, les États-Unis ont surajusté l'emploi

– la Suède et le Royaume-Uni (dans le secteur manufacturier et la construction) ont d'abord plutôt sous-ajusté l'emploi, mais ont engendré des gains de productivité en fin de période

– la France, à l'exception du secteur de la construction où l'emploi a été globalement préservé sur l'ensemble de la période, a sous-ajusté en première période puis, en seconde période, a réalisé des gains de productivité (secteur manufacturier) ou a renoué avec les niveaux de productivité d'avant la crise (secteur marchand). »

Source : « L'ajustement de l'emploi pendant la crise », Centre d'analyse stratégique, septembre 2012

Le degré général de flexibilité du marché du travail via les différentes formes de flexibilité à l'œuvre dans un pays donné et les politiques de l'emploi à l'œuvre permettent de discriminer entre les modèles d'ajustement de l'emploi aux variations de l'activité économique. Dans certains cas, les formes de flexibilité interne étant privilégiées (faible flexibilité externe), la productivité sera très variable. Le cycle de productivité est donc plus marqué ! Autrement dit, quand les créations/destructions d'emplois seront moins que proportionnelles aux fluctuations de la croissance économique, la productivité par tête sera beaucoup plus fluctuante. Dans d'autres cas, la flexibilité quantitative externe (ajustement rapide et important des effectifs) étant forte, les

variations de la productivité seront plus faibles. Le cycle de productivité est donc moins marqué ! Quand les créations/destructions d'emplois seront proportionnelles ou plus que proportionnelles aux fluctuations de la croissance économique, la productivité par tête sera moins fluctuante.

16 La réflexion sur l'appariement sur le marché du travail remonte à la courbe de William Beveridge (1944), qui relie le taux de chômage (u , en abscisse) et le taux d'emplois vacants (v , en ordonnée. Il s'agit des emplois vacants en part de la population active). Cette courbe est décroissante et convexe. Elle permet de montrer qu'il existe un chômage structurel lié aux difficultés d'appariement sur le marché du travail entre emplois vacants (demandeurs de travail/offreurs d'emploi) et chômeurs (demandeurs d'emploi/offreurs de travail). La formalisation des modèles d'appariement est due à Christopher Pissarides (1979 et 2000) et à Dale Mortensen et Christopher Pissarides (1994). En octobre 2010, Peter A. Diamond, Dale T Mortensen, et Christopher A. Pissarides se sont vus décerner le Prix Nobel d'économie pour leur méthode d'analyse des marchés portant sur les difficultés à harmoniser offre et demande, en particulier pour l'emploi.

INÉGALITÉS ET JUSTICE SOCIALE

« La justice consiste à traiter de manière identique les gens égaux et de manière inégale les inégaux. » - Aristote, 384 av. J.-C. - 322 av. J.-C.

« Dans la mesure où le système [capitaliste] marche, il est source d'efficacité pour l'économie. Mais la poursuite de l'efficacité crée nécessairement de l'inégalité et par conséquent la société est confrontée à un arbitrage entre efficacité et égalité. » - Arthur Okun, 1975.

« Les éthiques de l'organisation sociale qui ont résisté à l'épreuve du temps ont quasiment toutes en commun de vouloir l'égalité de quelque chose [...]. Non seulement il existe des "égalitaristes du revenu" qui veulent les mêmes rentrées d'argent pour tous et des "égalitaristes du bien-être social" qui revendiquent un accès égal à ce bien-être, mais les utilitaristes classiques préconisent, eux aussi, que l'on accorde une importance égale aux "utilités" de tous, et les libertariens purs que l'on reconnaisse à tous l'égale jouissance d'une classe entière de droits et de libertés.

Tous sont "égalitaristes" sur un point crucial – ils prônent résolument l'égalité de quelque chose que tout le monde devrait avoir, et qui est absolument vital dans leur approche particulière. Lorsqu'on perçoit le combat d'idées comme un affrontement entre partisans et adversaires de l'égalité – ce que font souvent les écrits théoriques –, on manque donc une caractéristique essentielle du sujet. » - Amartya Sen, 1992

« En tout salarié, il y a maintenant un exclu qui sommeille. » - Jean-Paul Fitoussi, 1995.

SOMMAIRE

[13.1 La dynamique des inégalités économiques](#)

[13.2 Les inégalités et la croissance économique : une relation complexe](#)

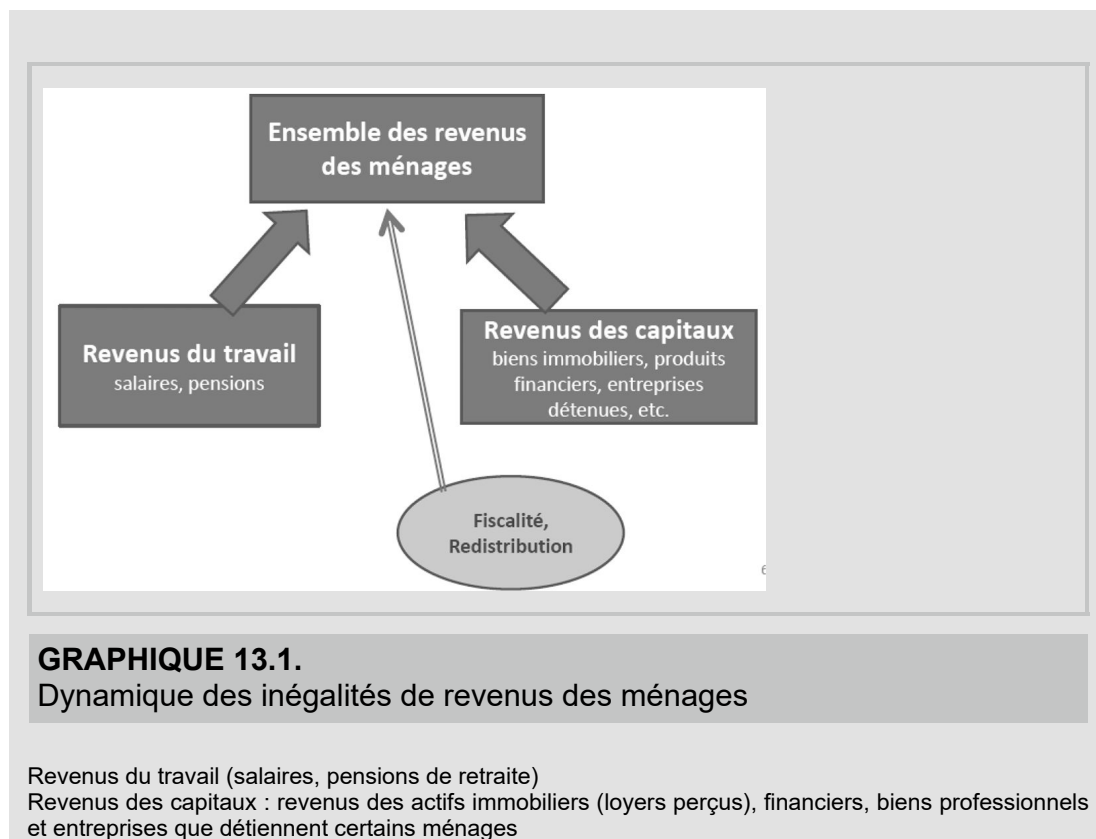
[13.3 Système de protection sociale : architecture, fonctionnement et réformes](#)

13.1 LA DYNAMIQUE DES INÉGALITÉS ÉCONOMIQUES

13.1.1 *Les différentes formes d'inégalités*

L'activité économique conduit à la création de valeur ajoutée, et, ensuite, la répartition des revenus a pour objet le partage de ces mêmes richesses entre les agents économiques. La question des inégalités de la répartition est très importante, mais c'est une question complexe d'un point de vue méthodologique et statistique. Les inégalités se situent de plus sur différents plans, au niveau du salaire tout d'abord, mais surtout au niveau du patrimoine. En effet, les inégalités peuvent être saisies suivant les deux grandes familles de revenus : les revenus d'activité et les revenus de la propriété.

13.1.1.1 Les inégalités de revenus



Source : Inégalités et croissance : Quelles perspectives ? Frédérique Savignac (Banque de France). « Quels emplois et quels revenus pour demain ? », Colloque DFC à BDF, 21 mars 2017 (organisé et modéré par David Mourey)

La mesure des inégalités porte d'abord sur les salaires, car ils constituent l'essentiel de l'activité des ménages français et sont mieux connus que les revenus des indépendants. Le salaire est la somme versée mensuellement par un employeur à son employé, en contrepartie d'un effort productif défini par le contrat de travail (il s'agit donc d'un flux monétaire). Même si les deux sont reliés, on peut distinguer les salaires des revenus, les seconds englobant davantage de flux monétaires (les salaires, mais aussi les revenus d'actions, les revenus immobiliers, etc.) Pour chacun de ces critères, on observe des inégalités : ainsi, un cadre touche un salaire en moyenne 2,5 fois supérieur à celui d'un ouvrier ou d'un employé (critère des professions et catégories socioprofessionnelles), ou encore un salarié de sexe masculin à temps complet perçoit en moyenne une rémunération nette supérieure de près de 25% à celle d'une femme. L'interprétation de ces inégalités salariales pose certains problèmes, car les facteurs interagissent les uns avec les autres et les interprétations peuvent être multiples. Ainsi, les inégalités entre salaires masculins et féminins tiennent à un phénomène de discrimination à emploi équivalent, mais aussi à l'occupation par les femmes de postes en moyenne moins qualifiés, ou encore à une fréquence plus grande du chômage et du temps partiel chez les femmes, etc.

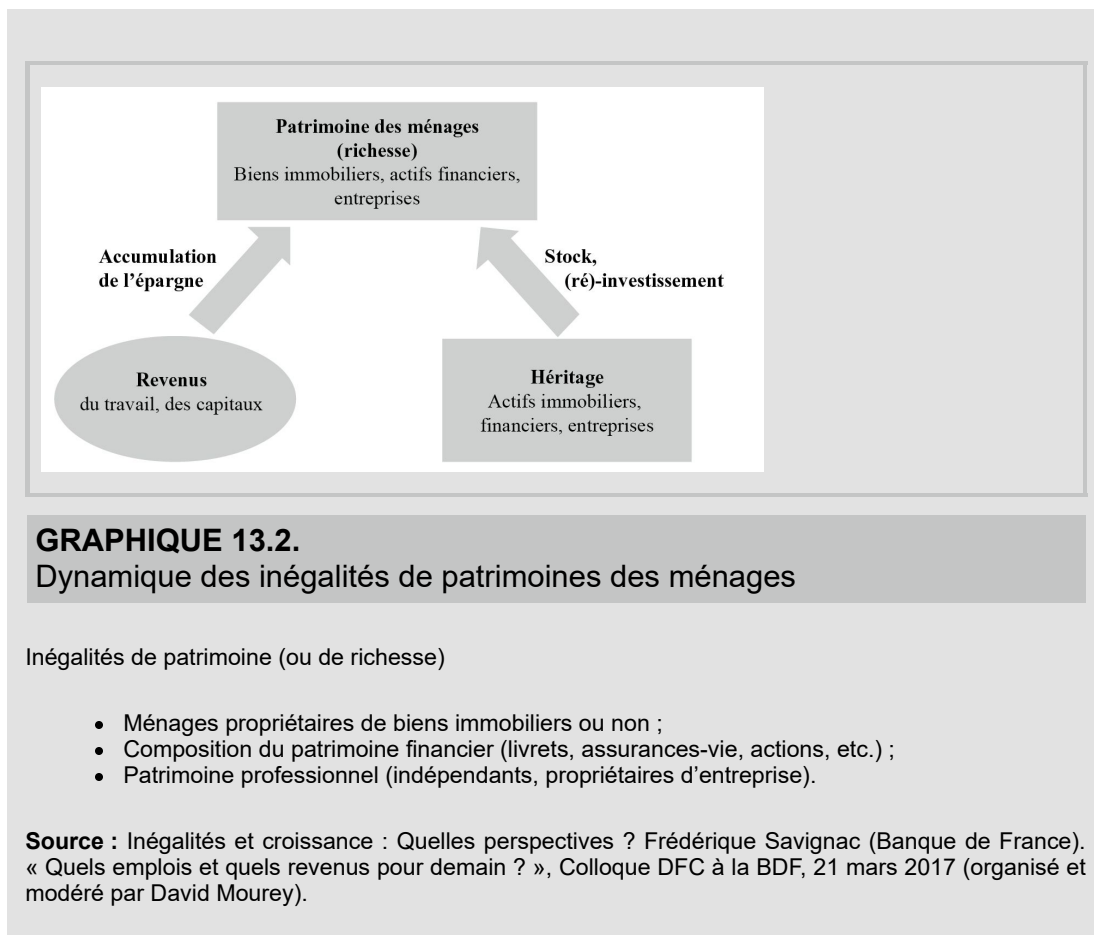
13.1.1.2 Les inégalités de patrimoine

Le patrimoine est l'ensemble des richesses (biens immobiliers, actifs financiers, biens professionnels, etc.) possédées par un agent économique (il s'agit donc d'un stock). Les inégalités de revenus de la propriété sont très importantes, car elles proviennent elles-mêmes de patrimoines très inégalement distribués : ainsi, en France, les 10% des ménages les plus riches possèdent environ la moitié

du patrimoine français. Le phénomène est d'autant plus important qu'il comporte une dimension cumulative dans le temps, qui va le plus souvent de pair avec des revenus salariés et qu'il combine tous les ressorts permettant l'accumulation (diversification du patrimoine, fiscalité favorable, héritages). Il existe des liens réciproques entre inégalités de salaire et de patrimoine. Un salaire élevé va, par exemple, permettre à un agent de dégager chaque mois de l'épargne et d'investir dans l'immobilier. Son salaire lui permet ainsi d'accroître son patrimoine. Mais en retour, ce patrimoine accumulé peut gonfler ses revenus : s'il loue par exemple l'appartement qu'il vient acheter, il augmente son revenu, il devient encore plus riche, donc peut acheter à terme un nouveau bien immobilier. Il existe donc une relation positive entre la progression du salaire et la progression du patrimoine, qui peut conduire à de très fortes inégalités.

Il existe une pluralité de déterminants des inégalités, parmi lesquels on peut citer l'âge, le sexe, la taille de la firme, la région, la profession et catégorie socioprofessionnelle (PCS) ou la qualification. Ainsi, certains économistes comme Gary Becker, dans le cadre de sa théorie du capital humain (1964), expliquent les différences de rémunération par des différences d'investissements en formation. Adam Smith au XVIII^e siècle affirmait déjà que les hommes ont toujours consacré des efforts importants à l'accroissement de leur habileté et de leur savoir, et que ces efforts sont pour la plupart volontaires. L'élévation de cette forme de capital immatériel (dont fait aussi partie l'effort dans la recherche) serait même la marque des systèmes économiques modernes, dans lesquels l'accumulation de capital technique et la création de connaissances nouvelles ne pourraient se faire sans une élévation parallèle des compétences des individus au travail. La prise en compte du capital humain (connaissances, aptitudes, compétences et savoir-faire acquis par l'individu) permet d'ailleurs d'expliquer l'augmentation de la productivité globale des facteurs (PGF). Par ailleurs, la hausse des revenus du travail, inégale selon

les individus, ne serait pas due à des mécanismes de rente, mais à un retour légitime sur investissement fait par les êtres humains. On identifie généralement cinq activités qui améliorent les capacités humaines : la santé, l'apprentissage par la pratique, la formation scolaire, la formation pour adultes et les migrations (diversité des expériences).



13.1.2 Les instruments statistiques de mesure des inégalités

■ Les quantiles

Une série statistique recense, pour différentes observations, la valeur que prend (ou prennent) la ou les variable(s) étudiée(s). Il est par exemple possible de recenser la valeur du revenu de l'ensemble des Français. Les quantiles (appelés aussi fractiles) sont des paramètres de position qui divisent la

distribution statistique en un certain nombre de parties égales : on peut regrouper les Français selon leur revenu, en formant par exemple quatre, cinq ou dix ensembles de personnes dont les revenus sont proches. Afin de calculer les quantiles, il faut commencer par classer la distribution par ordre croissant. Pour reprendre notre exemple, il faudrait ainsi classer les individus du plus pauvre au plus riche.

On peut choisir de diviser la population totale en parts plus ou moins grandes :

- Les quartiles divisent la distribution en quatre parties égales, chacune comprenant un quart des effectifs. Ainsi, Q_1 est la valeur qui divise la distribution telle que 25% des valeurs observées lui sont inférieures et 75% lui sont supérieures. Q_2 est appelée la médiane : c'est la valeur telle que la moitié de l'effectif a une valeur inférieure et la moitié une valeur supérieure. Pour garder l'exemple du revenu, le revenu médian désigne ainsi la valeur du revenu telle que 50% des Français ont un revenu inférieur à ce montant, et 50% un revenu supérieur. Pour un groupe de cinq élèves dont les notes (de zéro à dix, classées par ordre croissant) sont : zéro, deux, six, sept, huit, la médiane sera de six, car la moitié des élèves a une note inférieure à six et l'autre moitié une note supérieure. Si le nombre d'élèves avait été pair, six par exemple, la médiane aurait été la moyenne de la troisième et de la quatrième note.
- Les quintiles divisent la distribution en cinq parties égales, chacune comprenant un cinquième des effectifs. Par exemple, Q_4 est le quintile supérieur tel que 80% des valeurs lui sont inférieures et 20% supérieures.
- Les déciles divisent la distribution en dix parties égales. D_1 est la valeur qui divise la distribution telle que 10% des valeurs lui sont inférieures et 90% supérieures. D_5 est la médiane.
- Les centiles divisent la distribution en cent parties égales, chacune comprenant 1% de l'effectif total. Chaque

décile est ainsi divisé en dix parties égales.

– Les **quartiles**

25 % de l'effectif	25 % de l'effectif	25 % de l'effectif	25 % de l'effectif
Q1	Q2 (médiane)	Q3	

– Les **quintiles**

20 % de l'effectif				
Q1	Q2	Q3	Q4	

– Les **déciles**

D1	D2			D5				D9	

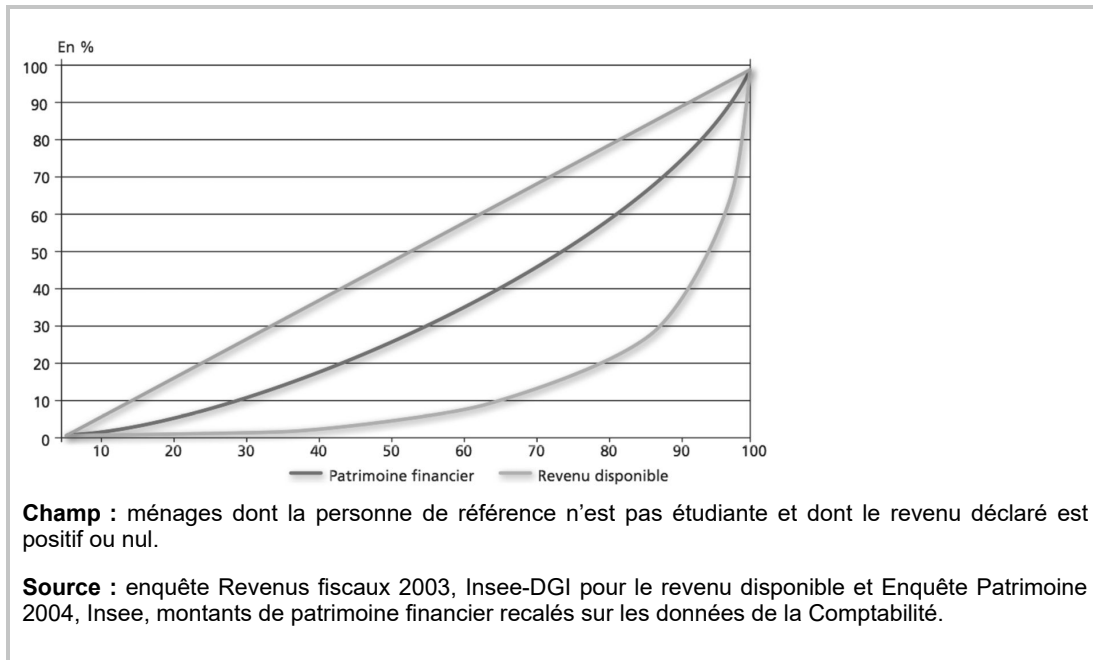
– Les **centiles** divisent la distribution en 100 parties égales, chacune comprenant 1% de l'effectif total. Chaque décile est ainsi divisé en 10 parties égales.

■ La courbe de Lorenz

La courbe de Lorenz (ou courbe de concentration) est une représentation graphique qui permet de visualiser le degré de concentration d'une variable statistique. Il s'agit de visualiser la façon dont est réparti un certain ensemble (par exemple, le revenu global des ménages). Pour la construire, il faut préalablement classer les observations par ordre croissant (par exemple du plus pauvre au plus riche) et calculer la part des observations pour lesquelles la valeur de la variable est inférieure à un certain niveau, et ce, pour différents montants (par exemple, la part des individus gagnant moins de 1 000 euros par mois, puis la part des individus gagnant moins de 1 500 euros par mois). Cette valeur définit la valeur de l'abscisse des différents points de la courbe.

Pour déterminer la valeur de l'ordonnée des points, il faut calculer pour chaque groupe considéré (par exemple, le groupe des personnes dont le revenu est inférieur à 1 000 euros par mois) la part de l'ensemble (par exemple, le revenu global des Français) détenu par les membres du groupe. La bissectrice (ou diagonale) représente la droite d'équipartition, illustrant une répartition parfaitement égalitaire de la variable, chaque partie de la population disposant de la même part de l'ensemble (par exemple, si les 10% les plus

pauvres possèdent 10% du revenu total des ménages). Plus la courbe de Lorenz s'éloigne de cette droite, plus le degré d'inégalité dans la répartition est élevé (plus la variable est concentrée).



GRAPHIQUE 13.3.

La courbe de Lorenz du revenu disponible et du patrimoine financier en France en 2003

La courbe de Lorenz sert souvent, en économie, à rendre compte de la répartition des revenus et des patrimoines à l'intérieur d'une population. On trouve en abscisses les % cumulés du patrimoine des ménages et en ordonnées les % cumulés du revenu disponible, ou du patrimoine financier.

Dans le cas de la France (comme pour les autres pays développés), la courbe du patrimoine est plus éloignée de la bissectrice que la courbe du revenu. Ainsi, les revenus sont distribués de manière moins inégalitaire que le patrimoine (ce dernier étant plus concentré).

En France, d'après une étude de l'Insee datant de 2004, la moitié des ménages aux revenus les plus faibles possède 27% de la masse des revenus disponibles (donc les 50% les plus aisés en possèdent 73%), tandis que la moitié des

ménages les moins bien dotés possède environ 4% de la masse totale du patrimoine financier (les 50% les mieux dotés en possèdent donc 96%). Les 90% les moins dotés en patrimoine financier en possèdent environ 37% (donc les 10% les mieux dotés en possèdent 63%).

■ L'indice de Gini

Le degré de concentration est mesuré par un coefficient appelé indice de Gini. Il se calcule en faisant le rapport entre l'aire de la surface entre la courbe de Lorenz et la bissectrice et l'aire du triangle sous la bissectrice (demi-aire du graphe). Plus la surface entre la courbe de Lorenz et la diagonale est élevée, plus la distribution est inégalitaire et plus l'indice de Gini est proche de 1.

Plus la distribution est égalitaire, plus l'indice de Gini est proche de 0. Ainsi, dans l'exemple précédent :

$$0 < \text{indice de Gini du revenu} < \text{indice de Gini du patrimoine} < 1$$

13.1.3 *L'évolution à long terme des inégalités économiques*

À l'échelle mondiale, l'analyse économique montre que les inégalités ont récemment décru entre les pays depuis les années 1980, en particulier sous l'effet du rattrapage économique des pays émergents (Chine et Inde notamment), mais qu'elles se sont creusées au sein des pays eux-mêmes. Les inégalités économiques demeurent considérables entre les nations : le niveau de vie des 10% les plus pauvres du monde (par exemple l'Éthiopie) reste en moyenne quatre-vingt-dix fois inférieur au niveau de vie des 10% les plus riches du monde (les États-Unis par exemple). Les inégalités de niveau de vie avaient continûment progressé au cours du XIX^e siècle et tout au long de la première moitié du XX^e siècle, sous l'effet de la révolution industrielle : l'écart de richesse entre les 10% les plus riches et les 10% les plus pauvres a été multiplié par trois de 1820 à 1980, même si les inégalités ont légèrement décru après la Seconde Guerre mondiale,

sous l'effet des politiques de redistribution mises en place par les pays d'Europe continentale. La réduction des inégalités mondiales a été spectaculaire depuis les années 1980, avec le développement économique des BRICS (Brésil, Russie, Inde, Chine, Afrique du Sud), même si un cinquième de la population mondiale reste confronté à la pauvreté. La crise depuis 2007 n'a que peu affecté l'évolution des inégalités à l'échelle mondiale, puisque le rattrapage économique des pays émergents, installés sur un *trend* de croissance vigoureux, ne s'est pas interrompu. La mondialisation de l'économie a entraîné des bouleversements gigantesques : dans les pays émergents, les effets bénéfiques du développement et de l'ouverture internationale se sont conjugués pour entraîner une restructuration profonde des appareils productifs, ce qui a favorisé l'enrichissement d'une partie de la population, grâce à l'élévation progressive des rémunérations, mais également une concentration accrue des revenus et des patrimoines.

Sur la période récente, l'accroissement des inégalités est dû à l'accroissement de la part des plus hauts revenus

La divergence de la dynamique des inégalités est très forte entre l'Europe de l'Ouest et les États-Unis, qui avaient des niveaux d'inégalité comparables en 1980, mais se trouvent aujourd'hui dans des situations radicalement différentes. Alors que la part de revenu du centile supérieur était proche de 10% dans les deux régions en 1980, elle est un peu montée en Europe en 2016 (12%), mais elle s'est envolée à 20% aux États-Unis. Durant la même période, aux États-Unis, la part des 50% les plus pauvres est passée de plus de 20% en 1980 à 13% en 2016.

La mondialisation a joué un rôle considérable dans l'évolution des inégalités : elle a hissé plusieurs centaines de millions de personnes au-dessus du seuil de pauvreté, même si elle a contribué directement (concurrence avec les pays à bas salaires, prime donnée aux détenteurs de capitaux) ou indirectement (libéralisation des économies, désindustrialisation des économies développées) à une

augmentation des inégalités au sein des nations. Le mouvement de réouverture des inégalités internes constitue une tendance générale dans les pays à hauts revenus de l'OCDE depuis les années 1980, en particulier aux États-Unis. Au sein des pays développés, dans certains secteurs (sport, culture, *show business*), « l'effet superstar » (*The Winner Takes All*) et les mutations technologiques se sont conjugués pour générer une croissance très forte des hauts revenus (sur la base d'un pourcentage perçu sur un gros volume de transactions). Par ailleurs, la globalisation financière et la mobilité du capital ont entraîné une forte progression des revenus financiers, et permis une élévation des rémunérations des décideurs et des travailleurs les plus qualifiés. Les inégalités se sont creusées au sein des pays développés en raison d'une série de causes : l'automatisation et d'autres changements technologiques, l'affaiblissement des syndicats, la financiarisation de l'économie et l'évolution des normes concernant les écarts de salaires au sein des entreprises. Dans les pays émergents, le recentrage sur le marché intérieur et l'intégration des masses paysannes au marché du travail ont favorisé une certaine diminution de la pauvreté et un resserrement des inégalités (comme l'avait observé l'économiste Simon Kuznets pour les pays développés depuis la révolution industrielle), même si le processus n'est pas acquis, notamment pour les pays les plus pauvres (comme en Afrique), par exemple si la rente tirée de l'exploitation des matières premières reste accaparée par une oligarchie dirigeante. L'enrichissement des élites économiques dans les pays émergents a toutefois accru les écarts entre une classe sociale insérée dans la mondialisation et les classes populaires.

D'une façon générale, au niveau national, on peut considérer qu'après une période de réduction sensible des inégalités qui va jusqu'au milieu des années 1980, on assiste à une stabilisation et même à un accroissement des écarts, en raison pour l'essentiel des revenus tirés de la propriété et du patrimoine qui bénéficient mécaniquement aux ménages les

plus aisés. Ainsi, l'ampleur des inégalités conduit-elle à poser la question de l'équité de la redistribution des revenus, et de l'opportunité de corriger la répartition due aux mécanismes du marché. C'est la raison pour laquelle les administrations publiques sont amenées à intervenir pour redistribuer les revenus, dans l'idée de corriger les inégalités les plus contestables, ou d'accroître l'efficacité du fonctionnement de l'économie de marché. Aujourd'hui, dans la zone OCDE, le revenu des 10% les plus riches est 9,5 fois plus élevé que celui des 10% les plus pauvres. À titre de comparaison, dans les années 1980, ce rapport allait de 7 à 1. Le coefficient de Gini est un autre indicateur du degré d'inégalité. Sa valeur est égale à 0 si l'égalité de revenu est parfaite (concentration minimum) et à 1 si tout le revenu est concentré entre les mains d'un seul individu (inégalité et concentration maximum). Dans les pays de l'OCDE, la valeur de ce coefficient était de 0,29 au milieu des années 1980, et elle s'établissait à 0,318 en 2014 (la valeur du coefficient de Gini a progressé dans seize des vingt et un pays de l'OCDE).

13.1.4 *Inégalités et justice sociale*

13.1.4.1 Une question controversée en économie

Régulièrement au cœur de l'actualité médiatique, les inégalités sont une question majeure en science économique. Elles constituent également un fait de société, qui renvoie à notre conception du « vivre-ensemble » et aux différentes conceptions de la justice sociale : il s'agit de savoir si les inégalités sont légitimes ou pas, ce qui fait référence à des choix citoyens. La justice sociale peut se définir comme l'ensemble des principes qui définissent la répartition équitable des droits sociaux et des devoirs au sein d'une collectivité, c'est-à-dire une répartition tenant compte de la situation personnelle des individus. Plusieurs conceptions de la justice sociale peuvent alors être mises en œuvre : traiter les individus de la même manière, ou bien les traiter inégalement. Dans le deuxième cas, on peut choisir de « donner plus à ceux qui ont moins » (pour corriger une

inégalité de départ) dans le cadre de politiques de « discriminations positives » qui créent volontairement une inégalité de droit pour mieux améliorer l'égalité des chances et l'égalité des situations. Au contraire, il peut être juste de ne pas taxer un héritage reçu par un individu, car la constitution d'un patrimoine plus important peut être bénéfique à la croissance de l'économie.

Il existe une diversité de critères qui permettent de définir ce qu'est une société « juste », mais les travaux des économistes n'ont pas vocation à trancher entre ces différentes conceptions de la justice. Toutefois, ils peuvent cependant éclairer les débats sur l'action de l'État en mesurant l'efficacité comparée des différents instruments disponibles pour atteindre les objectifs choisis par les pouvoirs publics. Dans l'histoire des idées, on oppose souvent deux grands courants :

- Les héritiers de Marx considèrent que c'est le fondement même de la société capitaliste qui est injuste et que seul un changement radical dans le mode de production permettrait qu'advienne une société plus juste ;
- Les penseurs du libéralisme comme von Hayek, qui dénoncent « le mirage de la justice sociale » : dans ce cadre, seule une décision individuelle peut être décrétée juste ou injuste. La notion de justice sociale n'a donc aucun sens dans cette conception libérale et elle est, de plus, dangereuse, puisqu'elle sert à justifier une intervention publique, qui risque toujours d'empiéter sur les libertés individuelles et de perturber l'allocation des ressources sur le marché.

Selon Amartya Sen, prix Nobel d'économie 1998, la société doit permettre à chacun d'exploiter pleinement ses potentialités, c'est-à-dire ses « capacités » (*capabilities*), pour pouvoir mener une vie digne et exercer une liberté réelle, concrète. Les « capacités » sont les chances laissées à un individu de réaliser ses objectifs ou de choisir entre différents modes de vie possibles : dans une société ouverte,

le respect des libertés fondamentales et l'égalité des chances sont ainsi nécessaires au plein exercice de ces « capacités ». Dans la conception d'Amartya Sen, la pauvreté n'est donc pas que privation de ressources monétaires, elle implique plus globalement une absence ou une insuffisance d'opportunités qui crée une situation d'infériorité sociale et de marginalisation. La pauvreté est alors inhérente à l'inégalité : le pauvre est ainsi privé d'une partie de sa liberté, celle de choisir sa façon de vivre.

Selon Jean-Paul Fitoussi, dans son ouvrage *La Démocratie et le Marché* (Grasset, 2004), il existerait dans nos démocraties occidentales contemporaines une tolérance plus grande aux inégalités, ainsi qu'une lente inversion des priorités entre les libertés politiques et les libertés économiques dans la phase actuelle de mondialisation de l'économie, c'est-à-dire d'expansion de la sphère du marché à la fois à l'intérieur de chaque société et à l'échelle de la planète. Il rappelle notamment que John Maynard Keynes se posait déjà la question des inégalités dans les années 1930 : « *Les deux vices marquants du monde économique où nous vivons sont que le plein emploi n'y est pas assuré et que la répartition de la fortune et du revenu y est arbitraire et manque d'équité* », écrivait-il. Jean-Paul Fitoussi plaide alors pour retrouver les réponses que Keynes apportait à son époque, c'est-à-dire une assurance collective d'activité par le moyen d'une politique macroéconomique de relance, et une réduction des inégalités par le moyen de la politique fiscale pour soutenir la demande. Selon lui, le marché et la démocratie apparaissent comme complémentaires plutôt que substituables et opposés : par l'élévation des niveaux de vie qu'il doit permettre, le système économique a pour effet d'accroître l'adhésion au régime politique ; et la démocratie, en réduisant les insécurités économiques (par la redistribution des richesses), doit rendre les résultats de l'économie de marché acceptables. Jean-Paul Fitoussi plaide, au final, pour l'idée que l'extension de la sphère du marché et celle du champ de

la démocratie politique et sociale se renforcent mutuellement en se limitant l'une l'autre.

Dans le cadre des démocraties, il est difficile pour des pouvoirs publics de contribuer à une société plus juste sans affirmer une égale considération pour tous les citoyens. Mais cette aspiration à l'égalité peut revêtir des contenus différents et la poursuite de l'idéal d'égalité dans une de ses dimensions peut coexister avec le maintien d'inégalités fortes dans une ou plusieurs autres dimensions. On peut aussi garantir le caractère effectif de cette égalité des droits, en luttant contre les discriminations, c'est-à-dire en traitant de manière différenciée des individus en fonction de leur appartenance à des catégories spécifiques, quand ces traitements conduisent à des inégalités avec les membres d'autres catégories. Ainsi dans les années 1990 en France, un débat important s'est établi sur les vertus comparées du principe d'égalité et celui du principe d'équité pour orienter les politiques publiques. Pour des auteurs comme Alain Minc, dans un rapport tiré de la Commission « La France de l'an 2000 », l'équité est une condition de l'efficacité, et le principe d'égalité peut dès lors être trop uniforme. Selon lui, *« dans un tel contexte, il faut identifier les inégalités inefficaces et injustes qui doivent être combattues, développer les formes d'équité qui sont un facteur d'efficacité, et trouver le niveau acceptable d'inégalités nécessaires pour assurer le dynamisme de l'économie. Cela suppose une réflexion sur l'optimum que la société souhaite »*.

Jean-Paul Fitoussi propose quant à lui, dans le cadre de ses différents ouvrages, de définir l'équité comme une propriété du critère d'égalité qu'on choisit. Selon lui, en effet, *« il apparaît donc vain de vouloir opposer égalité et équité, comme on l'a fait récemment. Ce serait vouloir opposer, à une conception, le jugement moral qu'on porte sur elle. L'égalité devant la vie, si elle implique qu'on accepte de grandes inégalités de ressources, exige pourtant qu'on cherche à en limiter l'ampleur, c'est-à-dire qu'on en contrôle, par des interventions publiques, le degré d'inégalité.*

Personne, je crois, ne pourrait contester une telle affirmation, et je n'ai entendu aucun discours - d'intellectuels ou d'hommes politiques - qui remît en cause la nécessité d'un système de redistribution, même de la part de ceux qui considèrent le système actuel comme trop généreux. Et encore, s'ils contestent le système actuel, la plupart de ces derniers le font-ils fréquemment parce qu'au nom d'une égalité plus formelle que réelle, il ne réduit pas suffisamment le degré spontané d'inégalité dans la répartition des revenus et des richesses » (Le débat interdit, Arléa, 1995).

En 2012, le prix Nobel d'économie Joseph Stiglitz a évoqué « *le prix de l'inégalité* » aux États-Unis en montrant que la dynamique inégalitaire des revenus échappait à l'analyse économique, qui fait de la rémunération l'indicateur d'une productivité marginale du travail et d'un investissement dans le capital humain supérieur. Les inégalités américaines seraient d'abord liées, selon lui, à l'importance disproportionnée d'activités liées à la capture de rentes : monopoles, lacunes de la gouvernance des entreprises (qui permettent à leurs dirigeants de s'arroger une part considérable des bénéfices), capacité à profiter de la bienveillance de l'État (comme pour l'industrie bancaire et financière). Selon Stiglitz, une partie du patrimoine des financiers proviendrait aussi de l'exploitation des plus démunis par le truchement de prêts usuriers et de pratiques abusives du financement des achats par cartes de crédit. Des auteurs comme Paul Krugman, également prix Nobel d'économie, notent que la critique d'une justice « *distributive* » a été utilisée, depuis le début des années 1980, pour justifier un retrait général de l'État-providence : la seule tâche de l'État serait de s'assurer que les actions individuelles respectent les règles abstraites qui se sont dégagées de l'évolution des sociétés humaines. La justice sociale serait vue comme une quête contre-productive, puisqu'elle voudrait redistribuer une richesse qui ne pourrait être créée dans une société matériellement égalitaire. Dans

ce cadre de pensée dominant, il faut laisser les riches s'enrichir si on veut que les pauvres sortent de la pauvreté.

13.1.4.2 La question du partage capital/travail

La répartition des richesses fait l'objet d'un vif débat politique, puisque la question s'inscrit dans le conflit traditionnel capital/travail au sein du capitalisme. En France, la controverse renvoie à la tolérance plus ou moins forte aux inégalités, dans un pays marqué par la « *passion pour l'égalité* », évoquée jadis par Alexis de Tocqueville. Des pères fondateurs de l'économie politique (Malthus, Ricardo, Marx) aux théoriciens de l'école néoclassique ou postkeynésienne, les économistes ont fait du partage de la valeur ajoutée une problématique centrale. Pourtant, dans les faits, la stabilité du partage capital/travail reste frappante sur le long terme, malgré les transformations structurelles gigantesques qu'ont connues les économies développées (tertiarisation, mondialisation de la production) : le partage reste proche d'un tiers pour le capital et deux tiers pour le travail. En France, le débat politique s'est focalisé sur la baisse de la part du travail perceptible à partir des années 1980 (et donc la hausse du taux de marge), après une augmentation dans la deuxième moitié des années 1970. Les enjeux macroéconomiques de ce partage sont considérables, avec le freinage de la demande qui persiste, malgré une meilleure rentabilité des entreprises (la hausse de la part des profits implique une baisse de la part des salaires), et le risque d'aggravation des inégalités et de déformation du partage des revenus, responsable d'un endettement croissant et risqué des classes populaires pour compenser la faiblesse des revenus salariaux. Cette déformation du partage des richesses peut être due aux changements technologiques et à la mondialisation de l'économie, mais il soulève aussi des questions politiques et sociales sur la gouvernance des entreprises centrée sur la création de valeur pour les actionnaires et sur le capitalisme financiarisé.

Dans l'ensemble des pays avancés, la part de la valeur ajoutée distribuée aux salariés a diminué au cours des quarante dernières années. Entre 2008 et 2015, les États-Unis et l'Europe ont ainsi connu une baisse assez nette de cette part salariale : de 68% à 60% dans les pays européens, et de 65% à 58% pour les États-Unis, et dans tous les secteurs de l'économie. Fait marquant, on observe également une baisse de cette part dans les pays émergents. On note également que sur la période 1980-2015, la part des profits dans la valeur ajoutée a augmenté de 7 points aux États-Unis et de 19 points en Europe, contribuant fortement à l'augmentation de la part revenant au capital.

Le partage de la richesse créée par les entreprises, entre actionnaires et salariés, suscite le débat aujourd'hui : certaines études comme celle d'OXFAM en mai 2018 sur la France, intitulée *CAC 40 : des profits sans partage*, critiquait l'augmentation des rémunérations des actionnaires depuis la crise financière, puisqu'entre 2009 et 2016, ces entreprises ont distribué 67% de leurs bénéfices aux propriétaires du capital, avec une part restante de 27% pour l'investissement et de seulement 5% pour les primes aux salariés. À la même période, d'autres travaux montraient que les entreprises américaines du S&P 500, l'indice boursier aux États-Unis, consacraient 91% de leurs bénéfices aux actionnaires, avec une part résiduelle de 9% pour l'investissement et les primes aux salariés.

Par ailleurs, lorsqu'on examine plus en détail la part du travail, aux États-Unis, au Japon et en France, on constate que la part du travail peu qualifié baisse nettement dans le coût du travail, tandis que la part revenant au travail qualifié progresse (avec sur le long terme une stagnation du salaire réel moyen).

Au sein du capital, la part croissante distribuée aux actionnaires est le signe d'une forte rentabilité, mais pourrait fragiliser la capacité des entreprises à investir et à soutenir leur compétitivité structurelle à long terme. Elle explique aussi une aggravation des inégalités au sein du salariat. La

croissance des inégalités, tant aux États-Unis qu'en Europe, s'explique donc par la croissance des revenus du capital : par exemple, aux États-Unis depuis les années 2000, ces revenus du capital ont augmenté 22 fois plus vite que les revenus du travail. On sait que cette croissance des revenus du capital alimente une progression cumulative des inégalités économiques : elle permet une accumulation patrimoniale plus rapide pour les titulaires de hauts revenus, dont la capacité d'épargne est supérieure, et leur permet ensuite une plus grande accumulation patrimoniale, etc.

La crise financière internationale de 2008 a également suscité de nouvelles réflexions en économie, notamment sur la détermination d'un partage assurant à la fois l'équilibre macroéconomique et la stabilité du système financier.

13.2 LES INÉGALITÉS ET LA CROISSANCE ÉCONOMIQUE : UNE RELATION COMPLEXE

13.2.1 *L'arbitrage efficacité/égalité : un débat central en économie*

Les économistes analysent l'arbitrage (en anglais *trade-off*) entre l'efficacité économique et la justice sociale : ce dilemme s'impose dans les sociétés capitalistes qui adoptent un double objectif, à savoir maintenir des incitations à la création de richesses, tout en défendant l'idéal démocratique d'égalité.

- L'efficacité consiste à obtenir le maximum des facteurs de production utilisés et suppose que les individus sont prêts à acheter cette production supplémentaire.
- L'égalité concerne, elle, les disparités de niveaux de vie qui reflètent les inégalités dans la distribution des revenus et du patrimoine.

L'égalité des niveaux de vie n'implique pas pour autant une uniformité ou une monotonie dans les choix des individus et ne signifie pas non plus une égalité des chances. L'arbitrage

entre l'efficacité et l'égalité ne veut pas dire que tout ce qui accroît la première réduit la seconde, mais les deux buts entrent souvent en conflit. L'argument central de ce partage efficacité-égalité est que la richesse doit être produite *avant* d'être redistribuée, et que la manière dont est partagé le gâteau peut avoir une incidence sur sa taille. Économiste keynésien, Arthur Okun analyse au milieu des années 1970 l'arbitrage (*trade-off*) entre l'efficacité et l'égalité, qui s'impose dans les sociétés capitalistes adoptant ce double objectif. Ce compromis, selon lui, est « *l'arbitrage socio-économique le plus important* », car « *on ne peut avoir le gâteau procuré par l'efficacité économique et le partager également* ».

Dans son ouvrage publié en 1975, intitulé "*Efficiency and equality, the big trade off*", il considère que « *dans la mesure où le système [capitaliste] marche, il est source d'efficacité pour l'économie. Mais la poursuite de l'efficacité crée nécessairement de l'inégalité et par conséquent la société est confrontée à un arbitrage entre efficacité et égalité* ». L'efficacité consiste à obtenir le maximum des inputs utilisés ; elle implique que le « *plus* » est le « *mieux* », sous réserve, note Okun, que les individus soient prêts à acheter cette production supplémentaire. Cette définition suppose que les individus sont les meilleurs juges de leur choix et laisse de côté la question de l'impact de la croissance sur l'épuisement des ressources. L'égalité concerne, elle, les disparités de niveaux de vie qui reflètent les inégalités dans la distribution des revenus et de la richesse. L'égalité des niveaux de vie n'implique pas pour autant une uniformité ou une monotonie dans les choix des individus et ne signifie pas non plus une égalité des chances. L'arbitrage entre l'efficacité et l'égalité ne veut pas dire que tout ce qui accroît la première altère la seconde, mais, selon lui, les deux buts entrent souvent en conflit. Arthur Okun évoque ainsi un effet de « *seau percé* » (*leaky bucket*) qui comprend à la fois les effets désincitatifs de l'impôt redistributeur sur l'incitation à épargner et à travailler et les coûts administratifs des transferts lorsqu'ils se multiplient. Il

reste cependant convaincu qu'il « *ne faut pas mettre tous ses œufs dans le panier du marché* » ce qui justifie, malgré la perte d'efficacité, le maintien de l'impôt.

13.2.2 Les inégalités comme facteur de dynamisme économique

Les économistes soulignent généralement l'acceptation d'un certain niveau d'inégalités considéré comme légitime, en laissant une large initiative aux individus dans une société libre, en stimulant la compétition et le désir d'entreprendre. Ainsi, même si le marché est soumis à une multitude d'interactions et de coordinations entre acteurs, dont l'issue est incertaine, il est alors probable que le marché récompense les acteurs en fonction de leur effort fourni. Par généralisation, les inégalités sont bénéfiques et nécessaires, car elles stimulent l'initiative privée : l'intervention de l'État peut donc être néfaste à la croissance économique si elle décourage l'incitation à la création de richesses. Les économistes évoquent ainsi les effets négatifs de l'impôt sur l'incitation à épargner et à travailler et les coûts administratifs des transferts sociaux lorsqu'ils se multiplient. L'aphorisme « *trop d'impôt tue l'impôt* » est ainsi souvent cité : déjà repéré par Adam Smith, ce principe suggère que les recettes fiscales s'élèvent logiquement avec les taux d'imposition, mais que, passé un certain seuil, le rendement de l'impôt commence à diminuer jusqu'à s'épuiser. L'arbitrage ("*trade-off*") entre le volume des recettes fiscales et le taux d'imposition va être matérialisé dans les années 1970 par une courbe célèbre, proposée par l'économiste américain Arthur Laffer, chef de file de « l'économie de l'offre ». L'explication de cette courbe repose sur le mécanisme suivant : jusqu'à un certain taux d'imposition, les individus augmentent leur activité pour compenser la baisse de revenu provoquée par la taxation (effet revenu). À partir de ce taux, ils ne sont plus incités à travailler et à épargner ou ils cherchent, par le travail au noir ou l'évasion fiscale, à échapper à l'impôt (effet de substitution). Ceci provoque une réduction de l'assiette de

l'impôt (l'ensemble des revenus taxables) et des recettes fiscales. Ainsi, la taxation des plus riches produirait-elle des effets pervers (fuite des capitaux, désinvestissement) qui pénaliseraient au final les plus pauvres, puisque le niveau de vie moyen stagne, voire régresse. Par ailleurs, les individus seraient désincités à travailler, soit parce que leurs revenus seraient réduits par les impôts, soit parce que les allocations les enfermeraient dans des « *trappes à inactivité* ». Enfin, l'importance des prélèvements sur le travail renchérirait son coût et conduirait à une délocalisation de la production dans les pays à bas salaires. Mais pour la majorité des économistes, une société ne peut laisser uniquement le marché déterminer ses équilibres sociaux. Le creusement des inégalités peut créer des tensions sociales importantes et accentuer la ségrégation et l'exclusion sociales. Par ailleurs, des politiques de redistribution mal ciblées et non centrées sur les outils les plus efficaces peuvent se solder par un gaspillage de ressources et être source d'inefficiences : si l'imposition des entreprises rentables et des revenus élevés amoindrit les incitations à créer des richesses, le résultat sera une plus faible croissance économique et des efforts individuels moindres.

13.2.3 Le prix de l'inégalité et la réduction des inégalités comme stimulant de la croissance

Des études récentes de l'OCDE démontrent que les inégalités entravent l'accumulation du capital humain : les inégalités de revenus compromettent les possibilités de s'instruire pour les populations les plus défavorisées, ce qui limite ainsi la mobilité sociale et le développement des compétences. John Maynard Keynes a insisté sur l'importance de l'investissement dans une économie pour stimuler la croissance, lui-même dépendant du niveau de consommation. Or, ce dernier est fonction de la propension moyenne et marginale à consommer : plus les agents économiques consacrent une part importante de leur revenu à la consommation, plus la demande, l'investissement et, *in*

fine, la croissance sont stimulés. Or, en présence de fortes inégalités de revenu, les plus riches ont tendance à épargner plus en proportion de leurs revenus, et donc à bloquer ce processus. C'est pourquoi une politique fiscale adaptée, en visant à réduire les inégalités de revenu initiales par l'impôt, peut permettre, grâce à la redistribution en faveur des plus défavorisés, qui ont une propension moyenne et une propension marginale à consommer plus fortes, de stimuler la consommation. Ainsi, la réduction des inégalités dans ce cadre est un bon moyen de favoriser l'activité économique. Le philosophe John Rawls, dans son ouvrage *Théorie de la justice* (1971) explique que les inégalités économiques, de revenu notamment, ne sont acceptables que si elles bénéficient à l'ensemble de la société. Il parvient à cette conclusion en raisonnant à partir d'une situation théorique fictive, où chaque individu ne connaîtrait pas sa position économique (« voile d'ignorance »), et s'interrogerait sur sa tolérance à l'égard des inégalités. D'après lui, et à condition que la société garantisse les principes d'égalité des chances et de liberté maximale, les inégalités sont acceptées si l'amélioration du sort des plus favorisés participe à celle des plus défavorisés (principe dit du « maximin »). Dans ce cadre, l'impôt progressif (qui taxe davantage les hauts revenus) peut à la fois se justifier et être efficace, car il correspond à une redistribution des plus riches vers les moins riches, sans pour autant conduire à une « préférence pour le loisir », c'est-à-dire décourager le travail.

13.2.4 Inégalités, prélèvements obligatoires et justice fiscale

L'impôt reste le moyen privilégié de donner à l'État les moyens de financer des actions croissantes. Les grands changements politiques (Révolution française, Révolution américaine) se sont aussi accompagnés d'une transformation du pouvoir fiscal de l'État (suppression de privilèges, choix démocratique de l'impôt). Si l'État peut se financer en empruntant, il le fait donc surtout en usant de sa souveraineté

(contrainte légitime) sur les citoyens par des « *prélèvements obligatoires* ». Les individus ne sont pas libres de déterminer leur contribution, mais y consentent à travers leurs représentants, qui votent chaque année une « *loi de finances* ». Ces prélèvements sont sans contrepartie : les prestations publiques ne sont pas proportionnelles aux versements des contribuables ou ne sont pas individualisables. Ils sont destinés aux administrations publiques (État central, collectivités locales, Sécurité sociale ainsi que les institutions de l'Union européenne). Si on distingue analytiquement, au sein des prélèvements obligatoires, les impôts des cotisations sociales, l'usage commun fait des « *impôts* » un synonyme des « *prélèvements obligatoires* ». Certains économistes français comme Thomas Piketty, Camille Landais et Emmanuel Saez soulignent que les impôts sont en réalité des « *suppléments obligatoires* ». S'ils financent d'abord un certain nombre de services collectifs (justice, police, défense, état civil), ils sont avant tout utilisés pour redistribuer de la richesse sous forme de transferts financiers (prestations sociales, subventions, bourses, revenu minimum). Cette redistribution permet d'assurer une cohésion sociale (plus de la moitié des dépenses publiques sont en France des dépenses sociales) et une certaine redistribution des revenus. Thomas Piketty a ainsi montré que l'impôt progressif sur le revenu, créé au début du ^{xx}^e siècle, a considérablement réduit les inégalités de fortune au cours du ^{xx}^e siècle en France. Ces impôts sont enfin, lorsque leur volume est suffisant, un moyen de la politique économique, soit par la réduction de leur montant (*tax cut*) pour relancer l'activité, soit par leur relèvement pour la ralentir. Cantonnés au niveau de 10% du PIB jusqu'au début du ^{xx}^e siècle, les taux d'imposition se sont depuis progressivement élevés jusqu'au niveau de 30 à 40% du PIB suivant les pays. L'impôt sur le revenu a été l'outil privilégié de réduction des inégalités au ^{xx}^e siècle puisqu'il a freiné l'accumulation patrimoniale des grandes fortunes grâce à des taux marginaux supérieurs relativement élevés. La réduction

de ce taux marginal au cours des années 1980-1990 aurait alors favorisé la réouverture des inégalités.

En France notamment, le niveau élevé des prélèvements obligatoires (soit 45,4% du PIB en 2017) a été critiqué, au nom des effets pervers qu'il peut exercer sur l'offre productive, en désincitant à l'épargne, au travail et à l'innovation. Loin d'être un facteur d'accroissement de la richesse, l'impôt détournerait aussi une partie croissante de la richesse pour des usages improductifs (dépenses sociales, « *train de vie* » de l'État) et serait manipulé pour satisfaire une multitude d'intérêts privés. Pourtant, l'élévation constante des besoins de financement de l'État et le creusement des déficits publics ces dernières décennies a déplacé la question vers celle de la structure de la fiscalité. La nécessité de l'impôt étant admise, il s'agit alors de trouver une fiscalité qui ne pénalise pas l'usage de certains facteurs (travail) et qui retrouve une progressivité pour assurer une certaine « *justice fiscale* ».

La question de la justice fiscale est très vive dans le débat public, puisqu'elle renvoie à la répartition des revenus et à l'arbitrage politique entre efficacité économique et cohésion sociale. La distribution des revenus est essentiellement liée au mérite, aux talents et au travail de chacun, mais les sociétés décident à des degrés divers de tenir compte des inégalités de départ et du poids de l'environnement social des individus, et d'opérer un certain niveau de redistribution verticale (des plus riches vers les plus pauvres). Les pouvoirs publics décident d'une certaine progressivité de l'impôt, c'est-à-dire que le taux moyen de celui-ci augmente quand le revenu de l'assujetti augmente (en France, les principaux impôts progressifs sont l'impôt sur le revenu, l'impôt de solidarité sur la fortune et l'impôt sur les successions). L'impôt demeure en effet l'instrument majeur de la lutte contre les inégalités : il permet le versement de prestations en espèces, mais aussi le renforcement de l'accès aux services publics – entre autres à des services d'éducation, de formation et de soins de qualité – qui sont autant d'investissements sociaux

qui doivent se traduire, à long terme, par davantage d'égalité des chances. L'État est alors en quête de la fiscalité dite « *optimale* » : il s'agit de déterminer un niveau de prélèvement suffisamment incitatif pour les agents privés, tout en conservant une certaine progressivité, afin de réduire les inégalités les plus choquantes. La complexité de cet arbitrage se trouve accentuée dans une économie mondialisée où la mobilité du capital est très forte (et sa taxation plus difficile), tandis que la mobilité du travail est nettement plus faible. Les acteurs plus « *nomades* » (firmes, détenteurs de capitaux, travailleurs très qualifiés) sont en mesure d'échapper à la progressivité de l'impôt, tandis que les États peuvent se trouver engagés dans une concurrence fiscale pour renforcer l'attractivité des territoires, notamment en Europe.

13.3 SYSTÈME DE PROTECTION SOCIALE : ARCHITECTURE, FONCTIONNEMENT ET RÉFORMES¹⁷

13.3.1 *Les logiques et principes des systèmes de protection sociale*

13.3.1.1 L'architecture du système de protection sociale

L'étude du système de protection sociale s'inscrit dans le prolongement logique des chapitres sur la répartition et la redistribution des revenus, sur les inégalités et la justice sociale et sur la place de l'État dans l'économie. Dès la présentation des trois problèmes économiques fondamentaux de l'économiste Paul Samuelson (Que produire ? Comment produire ? Pour qui produire ?), la question de la protection sociale est abordée : quels services de protection produire ? Comment les produire ? Pour quelles populations et selon quelles modalités d'accès ?

Ensuite, lors de la présentation des trois types d'organisation de l'économie (économie de marché, économie planifiée, économie mixte), les interrogations relatives à la protection

sociale sont présentes, dans la mesure où, selon l'organisation économique, la protection sociale reposera sur des logiques et principes très différents. Enfin, avec les trois fonctions économiques de l'État distinguées par R. Musgrave (allocation des ressources rares, répartition des revenus et régulation économique), la problématique est abordée encore plus directement dans la seconde fonction.

D'une certaine manière, les réflexions sur le système de protection sociale s'inscrivent dans l'espace plus large des réflexions sur la place de l'État (au sens large) dans l'économie et la société.

La protection sociale désigne l'ensemble des mécanismes de prévoyance collective permettant aux individus de faire face à des risques sociaux multiples et à leurs conséquences financières. Ces risques sont dits « sociaux », car ils sont étroitement liés au fonctionnement de l'économie et de la société. Ils ne sont pas naturels au sens strict. Le risque de perdre un emploi est un risque social. Avoir des enfants est un risque social, dans la mesure où cela va avoir des impacts économiques, sociaux... pour les parents, mais également pour la société dans son ensemble. Être malade est un risque social, car des maladies peuvent avoir des origines purement économiques et des conséquences économiques et sociales... Pour un ménage, un risque social peut se traduire par une baisse de ses revenus, une hausse des dépenses (vieillesse et retraite, santé et maladie, santé et invalidité, emploi et chômage, famille et maternité, éducation), mais les effets, dans une économie et une société d'échanges divers, ne resteront pas circonscrits aux ménages concernés, d'autant plus que les mêmes risques affectent simultanément de nombreux ménages, de nombreuses personnes.

Pour ces raisons, les pouvoirs publics ont mis en place progressivement un ensemble de dispositifs destinés à couvrir une population donnée contre certains aléas inhérents à la vie économique et sociale. Si cette protection sociale est obligatoire, c'est en raison de l'imprévoyance, par insouciance, par égoïsme, par calcul de l'individu face aux

risques auxquels il pourrait faire face. En effet, l'imprévoyance des uns, quand elle est largement partagée dans la population, peut avoir des effets externes négatifs sur de nombreux autres qui pourraient être prévoyants. Tout se passe comme si la protection sociale est un bien public à l'origine de nombreuses externalités positives (quand elle est efficace) ou négatives (quand elle est insuffisante). Dans ces conditions, la somme de comportements rationnels au niveau individuel peut conduire à un comportement collectif irrationnel.

Pour le dire autrement, pour des raisons économiques et sociologiques, tout le monde a intérêt à ce que les autres puissent avoir un emploi stable, des revenus stables et élevés, à ce que chacun puisse être complètement intégré à la société, soit en bonne santé pour travailler plus et mieux consommer, puisse avoir des enfants, etc.

13.3.1.2 Les logiques de l'État-providence

Le terme État-providence est souvent utilisé pour désigner le système de protection sociale. Mais, pour être précis, retenons d'emblée que, au sens large, l'État-providence correspond à l'ensemble des interventions économiques et sociales de l'État (donc politiques de régulation macroéconomique comprises). En revanche, au sens strict, l'État-providence désigne l'intervention de l'État dans le domaine social, particulièrement à travers le système de protection sociale. Cette conception du rôle de l'État dans l'économie et la société se distingue fondamentalement de celle de l'État gendarme et de ses fonctions exclusivement régaliennes (armée, justice, police).

Un système de protection sociale fonctionne selon trois logiques principales : une logique d'assurance sociale, une logique d'assistance et une logique de protection universelle.

- Dans la logique d'assurance sociale, l'objectif est d'obliger les individus à se protéger contre des risques potentiels de pertes de revenus pour différentes raisons. Comme pour une assurance privée, les prestations

sociales sont les contreparties des cotisations versées préalablement, qui les financent. Seuls ceux qui versent les cotisations et leurs ayants droit sont couverts en cas de réalisation du risque.

– Selon la logique de l'assistance, l'objectif est de mettre en place une autre forme de solidarité entre les individus pour protéger ceux qui ne sont pas ou plus couverts par des assurances sociales. Les prestations sociales sont davantage la contrepartie d'impôts et sont versées sous conditions de ressources.

– Dans le cadre de la logique de protection universelle, le but est de couvrir certaines catégories de dépenses pour tous les individus (universalité). Les prestations sont versées sans contreparties et sans conditions de ressources. Elles sont identiques pour tous ceux qui sont concernés par le risque couvert.

La combinaison de ces logiques va faire émerger deux types de solidarité et de redistribution. L'assurance sociale crée une forme de solidarité entre les assurés et entraîne une redistribution horizontale. L'assistance sociale implique surtout une autre forme de solidarité entre les citoyens de la société, indépendamment de contributions préalables, et se traduit essentiellement par de la redistribution verticale, même si des formes de redistribution horizontale persistent.

Les assurances sociales regroupent l'ensemble des dispositifs, privés ou publics, destinés à protéger les travailleurs (ou leurs ayants droit) contre des risques sociaux. La protection sociale, selon cette logique d'assurance, est née avec la Révolution industrielle, la salarisation du travail, l'institutionnalisation de la relation de travail dans le cadre de l'emploi. L'assuré doit cotiser à une caisse de protection sociale. En contrepartie, il recevra des prestations lors de la réalisation d'un risque. La cotisation donne un droit de bénéficier de la protection collective *via* une redistribution horizontale qui va de ceux qui n'ont pas subi le risque vers ceux qui en ont été les victimes (de ceux qui sont en bonne

santé vers les malades, de ceux qui travaillent et produisent vers les jeunes et les personnes âgées, des travailleurs vers les chômeurs).

Les logiques d'assistance sociale vont renforcer ou compléter la solidarité, selon des mécanismes de redistribution verticale, mais aussi horizontale. Les prestations sont versées sans contrepartie préalable (cotisations sociales) à des personnes qui ne disposent plus de protection sociale *via* l'assurance. Il est davantage tenu compte de la situation particulière de la personne que de sa capacité contributive et de l'obtention d'un droit en contrepartie de celle-ci. C'est la situation particulière de l'individu qui confère ce droit. De fait, le financement de ces prestations sans contrepartie est fondé sur l'impôt. Ce n'est plus le statut de travailleur qui prévaut, mais celui de citoyen. Des minimas sociaux sont versés aux plus démunis, selon une logique de redistribution verticale et des allocations familiales sont versées aux familles nombreuses en enfants, selon une logique de redistribution horizontale.

L'organisation du système de protection sociale en France opère une distinction claire entre deux types de protection sociale dont les logiques de solidarité et de financement diffèrent : les régimes qui versent des prestations contributives, dont l'objectif est d'assurer des revenus de remplacement aux revenus d'activité, et les systèmes de protection sociale qui offrent des prestations à tous les citoyens en fonction de leurs besoins, sans rapport avec leurs contributions.

13.3.2 *Les modèles de l'État-providence*

13.3.2.1 **Modèle « bismarckien » et modèle « beveridgien »**

Sur la base de ces grandes logiques de financement de la protection sociale, on différencie habituellement deux modèles d'État-providence : le modèle « universaliste » et le modèle « corporatiste ». En effet, historiquement, les systèmes de protection sociale en Europe ont deux origines principales : le système « bismarckien », fondé par les lois du chancelier de Prusse Bismarck dans les années 1880 et le système « beveridgien », créé sur la base du rapport Beveridge en Angleterre (1942).

– Le modèle **fondé sur l'assurance** (corporatiste ou **bismarckien**) propose une protection basée uniquement sur l'assurance et repose sur l'emploi. Les cotisations et prestations sociales sont fixées en pourcentage du salaire et le financement est assuré par l'employeur et le salarié. Les droits sont donc accordés à celui qui travaille et à ses ayants droit (conjoint et enfants). En Allemagne, les trois lois fondamentales votées en 1883 (assurance maladie), 1884 (assurance accidents du travail) et 1889 (assurance invalidité et vieillesse) ont abouti à la mise en place d'un système d'assurance sociale qui garantit un revenu de remplacement ou de compensation en cas de perte de revenu liée à la réalisation d'un risque social. Le système bismarckien va avoir une forte influence sur les systèmes mis en place notamment au Luxembourg, aux Pays-Bas, en Autriche, Norvège, Suède, Italie, Belgique et France.

– Le modèle **fondé sur l'assistance** (universaliste ou **beveridgien**) propose une protection sociale à tous, en contrepartie de la citoyenneté. Le système de protection sociale est universaliste et repose donc sur la logique de l'assistance pour renforcer la solidarité. Chaque citoyen, contribuable, a droit à la protection sociale sans contrepartie en termes de cotisations sociales. Ce modèle universaliste est caractérisé par une protection sociale de haut niveau ouverte à tous en raison de la citoyenneté. C'est le rapport Beveridge « Social Insurance and Allied Services », publié en 1942, qui établit les fondements de cette nouvelle forme de protection sociale. Sir William Beveridge avait eu pour mission en 1941 de réfléchir à l'organisation d'un système britannique de sécurité sociale. Son rapport est considéré comme la base de ce qu'on appelle communément le *Welfare State* ou « État du bien-être », acception spécifique de l'État-providence. Beveridge, qui a été fortement influencé par la théorie keynésienne du sous-emploi et du chômage involontaire en raison d'un déficit de demande, est convaincu qu'une politique de sécurité sociale ne peut être efficace que dans

le cadre plus large d'une politique de lutte contre le chômage. Il publiera en 1944 un ouvrage devenu célèbre, *Full employment in a free society* (*Du travail pour tous dans une société libre*).

Le système beveridgien, à l'origine et en principe, s'appuie sur la règle des « 3 U » : universalité, uniformité et unité. En ce qui concerne le système de protection sociale *stricto sensu*, Beveridge refuse autant la logique de l'assistance que celle de l'assurance sociale réservée aux seuls travailleurs. Il préfère l'instauration d'un système généralisé et uniformisé de prestations sociales forfaitaires financées par des contributions forfaitaires. L'objectif est d'éradiquer la pauvreté en versant une protection à tous les citoyens (universalité), *via* une prestation forfaitaire (uniformité), indépendamment de la réalisation d'un risque social. Par souci de simplification et afin d'éviter toute forme de discrimination, l'organisation même du système doit être unifiée dans le cadre d'un service public unique de protection sociale (unité), placé sous l'autorité directe du gouvernement. Il est généralement admis que, pour être justes, les systèmes de protection universels, dont le financement repose sur l'impôt, doivent garantir des prestations élevées.

Le modèle français d'État-providence est un système hybride sur lequel nous reviendrons plus loin. Retenons juste qu'en France, le système de protection sociale est à l'origine très bismarckien et donc principalement fondé sur la logique de l'assurance et le financement par cotisations sociales. Mais, progressivement, il devient de plus en plus beveridgien et la part de prestations sociales financées par l'impôt (et donc par la collectivité des citoyens) croît relativement à la part de celles qui sont contributives, et donc financées par des cotisations sociales assises sur le travail.

13.3.2.2 La typologie des modèles d'État providence selon

Gosta Esping-Andersen

Le sociologue Gosta Esping-Andersen proposera une typologie en quatre modèles pour les pays européens : les modèles social-démocrate, corporatiste, libéral et méditerranéen. La distinction des systèmes nationaux est basée sur quatre critères : le mode de financement des prestations (impôts ou cotisations sociales), le mode de calcul (forfaitaires, proportionnelles), la population destinataire (prestations universelles ou sous conditions de ressources) et la forme de ces prestations (prestations en espèces ou prestations en nature, par la mise à disposition gratuite ou semi-gratuite de services sanitaires (soins) ou sociaux (crèches, maisons de retraite).

– Les systèmes « **sociaux-démocrates** » concernent des pays scandinaves (Danemark, Suède et Finlande). Ces pays sont ceux qui s'inscrivent le plus dans la logique universelle de Beveridge. La protection sociale est un droit pour tous les citoyens. Les prestations sont financées essentiellement par des impôts, même si les cotisations sociales s'y sont développées (Danemark). Les prestations, forfaitaires et élevées, sont versées automatiquement en cas de réalisation d'un risque. Les prestations d'assistance sont très limitées. La protection sociale, complètement publique, est sous l'autorité directe des pouvoirs publics. L'assurance chômage peut être assurée dans le cadre d'un système séparé.

– Les systèmes « **libéraux-résiduels** » de la Grande-Bretagne et de l'Irlande ne retiennent que certaines recommandations de Beveridge. L'accès à la protection sociale n'est pas lié à l'emploi, mais seul le service national de santé (*National Health Service*) est véritablement universel. Les prestations sont forfaitaires, mais leur niveau est bas. De fait, les assurances privées et les régimes de protection complémentaire d'entreprise s'y développent. Des prestations sous condition de ressources sont versées aux plus démunis. Le financement de la protection sociale publique repose largement sur l'impôt, même si quelques prestations peuvent être financées par

des cotisations sociales. Ces systèmes de protection sociale se traduisent par une forte dualisation au sein de la population entre ceux qui sont bien protégés et ceux qui ne le sont pas ou presque pas.

– Les systèmes professionnels « **corporatistes-conservateurs** » sont ceux qu'on peut retrouver en Allemagne, France, au Benelux et en Autriche. La logique de l'assurance y est prépondérante. Le versement de cotisations sociales donne droit à des prestations sociales. L'emploi, et donc le travail, est au cœur du système. Les assurances sociales sont obligatoires dans la plupart des cas. Les prestations en espèces sont donc contributives et proportionnelles aux revenus. Les cotisations sociales sont respectivement versées par les employeurs et par les travailleurs. Une partie faible, mais croissante, des prestations est financée par l'impôt ; on parle de fiscalisation de la protection sociale. En effet, la couverture des populations les plus en difficulté par rapport à l'emploi et donc au revenu implique le versement de prestations d'assistance. L'impôt n'a plus vocation à financer l'universalité, mais plutôt la solidarité à l'égard des plus fragiles. Ces systèmes sont gérés par les représentants des employeurs et des travailleurs sous contrôle de l'État.

– Le modèle « **méditerranéen** » concerne les pays d'Europe du Sud, Espagne, Grèce, Italie, Portugal. Le système est bismarckien, car il existe des assurances sociales pour les prestations de garantie de revenu, mais il existe de très grandes hétérogénéités entre les différents régimes, car l'assurance sociale repose sur une base professionnelle. Les systèmes de protection sociale sont en partie beveridgiens, car certaines branches (comme les systèmes de santé) sont souvent universalistes. Enfin, au-delà des spécificités entre ces systèmes nationaux, la faiblesse de l'État et son fonctionnement clientéliste, la persistance de fortes inégalités, en sont des caractéristiques durables.

13.3.3 *Le fonctionnement de l'État-providence*

13.3.3.1 Le fonctionnement vertueux de l'État-providence

En période de croissance soutenue et durable, de plein emploi avec une population active relativement abondante dans la population totale, les ressources du système de protection sociale sont abondantes, grâce aux fortes rentrées sociales et fiscales. Symétriquement, de nombreuses dépenses sont faibles. Le système peut donc verser des prestations de niveau élevé sur de longues périodes. La prospérité économique permet une couverture importante face aux différents risques sociaux. Elle favorise aussi une certaine insouciance dans la gestion des comptes du système. Quand les caisses se remplissent si facilement, pourquoi compter, s'interroger sur l'efficacité de la redistribution ?

Mais il est clair qu'un retournement des tendances économiques et démographiques changera les contraintes financières et imposera des changements plus ou moins drastiques au niveau des recettes et des dépenses. C'est ce qui sera observé à partir des années 1970 en France, en particulier après la période faste des « Trente Glorieuses » (1945-1973). Pour bien comprendre ce retournement tendanciel et ses conséquences, on peut s'appuyer sur l'analyse proposée en 1981 par Pierre Rosanvallon, dans un ouvrage intitulé *La crise de l'État-providence*. Pour l'auteur, « *la crise de l'État-providence se situe à trois niveaux "ceux du financement, de l'efficacité et de la légitimité de l'action publique"* » (*La crise de l'État-providence*, Seuil, 1981, p. 110). La crise financière de l'État-providence s'accompagne selon lui d'une crise d'efficacité et d'une crise de légitimité. Quatorze ans après son diagnostic sur cette crise, Pierre Rosanvallon a élargi son analyse dans *La nouvelle question sociale* (1995), en vue de repenser ses fondements et ses modalités. Sa thèse centrale est qu'à la crise financière de l'État-providence, apparue dans les années 1970 avec la crise économique et le net

ralentissement de la croissance (qui creuse les déficits des comptes publics), se sont superposées successivement une crise d'efficacité dans les années 1980 et une crise, d'ordre philosophique, de légitimité dans les années 1990.

13.3.3.2 La triple crise de l'État-providence

La *crise financière* ou crise de financement du système de protection sociale s'explique simplement par un « effet de ciseaux ». Avec la croissance économique durablement molle, le chômage durablement élevé, le montant des prestations sociales à verser augmente, alors que les recettes de cotisations sociales et fiscales baissent, la protection sociale reposant essentiellement sur l'emploi. La crise de l'emploi met *de facto* en danger la protection sociale. La part des prestations sociales dans le PIB croît plus rapidement que la part des cotisations sociales, d'où le creusement du déficit. Pour résoudre ces problèmes de financement d'un système essentiellement fondé sur la logique de l'assurance (le plein emploi), l'État est obligé de financer par l'impôt (fiscaliser) une part croissante de la protection sociale (création de la CSG). Avec le temps, et les effets du vieillissement démographique, les déficits des comptes sociaux sont de plus en plus difficiles à réduire. Dans le même temps, les inégalités traditionnelles persistent et de nouvelles inégalités apparaissent. L'efficacité de l'État-providence est alors remise en question.

La *crise d'efficacité* du système est donc une conséquence de la crise de financement et l'expression de ses déséquilibres et dysfonctionnements structurels. Malgré le niveau des dépenses, les inégalités se multiplient et la pauvreté, forme extrême d'inégalité, s'accroît. Or, à l'origine, la protection sociale a pour fonction de protéger efficacement contre les risques sociaux et de prévenir la croissance des inégalités et de la pauvreté. Non seulement la hausse du taux de prélèvements obligatoires (PO) ne permet pas de résoudre la crise financière, *via* une réduction des déficits sociaux, mais, de surcroît, le coût élevé de cette protection sociale

manque sérieusement d'efficacité. Cette hausse des PO peut d'ailleurs devenir contre-productive en pesant sur la croissance économique effective et potentielle (selon la logique de la courbe de Laffer, selon laquelle « trop d'impôt tue l'impôt »). Mais le problème le plus fondamental est celui de l'intégration pour les personnes qui échappent à la protection sociale. Cette crise d'efficacité de la protection sociale se transforme peu à peu en crise du lien social par défaillance de la capacité d'intégration sociale. Inexorablement, cette crise d'efficacité devient aussi une crise de légitimité.

La *crise de légitimité* exprime le fait que la protection sociale ne fonctionne pas dans la mesure où des inégalités persistent, augmentent, changent de formes ; d'autres inégalités apparaissent et se cumulent aux anciennes. Et, cerise sur le gâteau, ces nouvelles inégalités semblent profiter aux plus favorisés. Dans une démocratie, le citoyen a tendance à trouver une forme de légitimité dans les PO. Mais quand leurs contreparties se réduisent comme une peau de chagrin et que, de surcroît, la perte en efficacité vient s'ajouter à la perte de recettes, alors les prélèvements obligatoires peuvent être de plus en plus ressentis comme illégitimes aux yeux des contribuables, et notamment ceux des catégories moyennes. Il devient tout simplement de plus en plus difficile d'accepter de payer très cher un système de plus en plus inefficace. Avec la crise de légitimité de l'État-providence, c'est, au-delà de la question des inégalités, celle de la justice sociale qui est posée avec force.

13.3.4 *Le système de protection sociale en France*

13.3.4.1 Les quatre volets du système de protection sociale

Le système de protection sociale français est un système qui souhaite marier, avec une certaine harmonie, la logique de Bismarck à celle de Beveridge, celle de l'assurance et celle

de l'assistance. Le modèle français est surtout bismarckien, mais il devient progressivement plus beveridgien, dans la mesure où une partie croissante de la protection sociale relève de l'aide sociale fondée sur l'assistance. Désormais, l'impôt finance près de 20% (un cinquième) de la protection sociale. Cette « fiscalisation » s'inscrit dans une logique universaliste. En effet, étant donné que tout le monde doit en profiter, alors tout le monde doit participer au financement.

La Sécurité sociale est l'instrument privilégié de la protection sociale destinée aux individus et aux familles. Toutefois, la notion de protection sociale est plus générale et couvre un champ plus vaste que celui de la Sécurité sociale. La protection sociale fait l'objet d'une définition européenne et d'un système statistique harmonisé, le Système européen de statistiques de la protection sociale (SESPROS), piloté par EUROSTAT. Selon ce système, la protection sociale recouvre tous les mécanismes institutionnels, prenant la forme d'un système de prévoyance collective ou mettant en œuvre un principe de solidarité sociale, par lesquels les membres de la société sont protégés contre la réalisation d'un certain nombre de risques ou voient leurs besoins sociaux fondamentaux satisfaits. Elle implique le versement de prestations sociales aux ménages confrontés à la réalisation de ces risques. La protection sociale comprendra donc l'ensemble de ces mécanismes qui garantissent le versement de prestations : l'assurance sociale et l'assistance sociale d'une part, des régimes privés et des régimes publics d'assurance sociale, d'autre part.

Institutionnellement, le système de protection sociale comporte quatre volets :

- La Sécurité sociale, qui est composée d'organismes prélevant des cotisations obligatoires et assurant des prestations légales déterminées en fonction des cotisations ou des situations personnelles, et versées automatiquement ;

- L'assurance chômage : régime séparé pour couvrir contre le risque de perte d'emploi et de revenus ;
- L'aide sociale, qui désigne des prestations légales, mais versées de façon conditionnelle à des personnes dont les ressources sont trop faibles pour faire face aux besoins liés au handicap, à la maladie, à la vieillesse ou à des difficultés sociales ou économiques, dispense des prestations et des services répartis en trois grandes catégories : l'aide aux personnes âgées, l'aide aux personnes handicapées et l'aide sociale à l'enfance. Depuis le 1^{er} janvier 1984, la quasi-totalité de l'aide sociale a été transférée aux départements. La gestion de l'allocation du Revenu minimum d'insertion (RMI) puis du Revenu de solidarité active (RSA) l'a été au 1^{er} janvier 2004 ;
- L'action sociale, qui désigne des prestations facultatives versées par des organismes multiples (comités d'entreprise, mutuelles, etc.), parmi lesquels peuvent figurer des organismes de Sécurité sociale.

La notion de Sécurité sociale peut prêter à confusion, car elle désigne, au sens large, l'ensemble des organismes chargés de mettre en œuvre la protection sociale obligatoire, et dans un sens intermédiaire, l'ensemble des régimes de protection contre les risques de maladie, maternité, accidents du travail, vieillesse et famille et, au sens étroit, le régime de l'assurance maladie (Sécu).

13.3.4.2 La structure du système de protection sociale en France

L'Insee définit le champ de la protection sociale dans le Système de comptabilité nationale (SCN), conformément au Système européen de comptabilité (SEC), sur la base des notions de risques et de prestations sociales.

Sous l'influence du Système européen de statistiques de la protection sociale, la liste des risques sociaux a été formulée et appliquée de la manière suivante :

- La santé (prise en charge totale ou partielle de frais liés à la maladie, à l'invalidité, aux accidents du travail et aux maladies professionnelles) ;
- La vieillesse et la survie (pensions de retraite, pensions de réversion, prise en charge de la dépendance) ;
- La maternité et la famille (prestations familiales, prestations liées à la maternité, allocations familiales, aides pour la garde d'enfants) ;
- L'emploi (indemnisation du chômage) et les difficultés d'insertion ou de réinsertion professionnelle ;
- Le logement (aides au logement) ;
- La pauvreté et l'exclusion sociale (minima sociaux : revenu minimum d'insertion : RSA, minimum vieillesse, etc.).

Les prestations sociales sont des transferts courants destinés à préserver le revenu disponible des ménages contre la réalisation d'un certain nombre d'aléas de la vie. Ces transferts aux ménages, en espèces (pensions de retraite, allocations familiales, RMI, allocations chômage...) ou en nature (remboursements de soins ou de médicaments...), sont destinés à alléger la charge financière que représente la protection contre différents risques liés à la maladie, la vieillesse, le logement, etc. Ils sont effectués, dans le cadre de l'assurance sociale, par l'intermédiaire de régimes (publics ou privés) organisés de façon collective ou, dans le cadre de l'assistance sociale, par des administrations publiques ou des ISBLSM. Il existe un compte satellite de la protection sociale qui introduit la notion de « prestation de protection sociale » et qui comprend en plus des prestations de services sociaux telles que l'accès gratuit ou à tarif réduit à des services comme l'hospitalisation publique, l'hébergement des personnes handicapées, la réduction des tarifs des transports en commun pour les familles nombreuses ou les personnes âgées.

Une des originalités des systèmes de comptabilité nationale est de distinguer, afin de couvrir ces risques, l'assurance

sociale de l'assistance sociale. L'assurance sociale est caractérisée par l'existence de « régimes » mettant en œuvre un mécanisme d'assurance. Dans les régimes d'assurance sociale, « des cotisations sociales sont versées, par les salariés ou d'autres particuliers ou encore par les employeurs pour le compte de leurs salariés, afin de garantir aux salariés ou autres cotisants ainsi qu'à leurs personnes à charge ou à leurs survivants un droit à des prestations d'assurances sociales [...]». Contrairement aux prestations d'assistance sociale, l'octroi de prestations d'assurances sociales est subordonné à l'affiliation à un régime ». À l'inverse, l'assistance sociale n'implique ni affiliation du bénéficiaire ni versement de contributions. Elle est mise en œuvre par les administrations publiques au sens habituel du terme (État, ODAC, administrations locales) ou, à une moindre échelle, par des institutions sans but lucratif au service des ménages (ISBLSM).

Les quatre premiers risques (santé, vieillesse et survie, maternité et famille, emploi) sont typiquement des risques couverts par les régimes d'assurance sociale, et donc de Sécurité sociale, tandis que les deux derniers (logement, pauvreté et exclusion sociale) sont plutôt caractéristiques de l'assistance sociale.

13.3.4.3 Les dépenses et les recettes du système de protection sociale en France

■ Les ressources de la protection sociale

Les ressources de la protection sociale sont de natures diverses : cotisations sociales, impôts et taxes affectées (notamment la contribution sociale généralisée, les contributions publiques, etc.) En 2016, selon la DREES, les ressources de la protection sociale progressent de 2,2% et à 758,7 milliards d'euros. En 2016, les cotisations sociales représentent 61% du total des ressources de la protection sociale (461,3 milliards d'euros). Les impôts et taxes affectés représentent 24% du total des ressources de la protection sociale (184,1 milliards d'euros). La CSG en représente à elle

seule un peu plus de la moitié, soit 13% du total des ressources, auxquels il faut ajouter la contribution au remboursement de la dette sociale (CRDS) et la compensation des exonérations de cotisations patronales, assurée par l'État. Les autres ressources représentent 15% du total des ressources de la protection sociale (113,3 milliards d'euros). Les contributions publiques, avec 93,7 milliards d'euros, représentent la majorité de ces autres ressources de la protection sociale. En 2016, la hausse des ressources de la protection sociale est surtout portée par la forte augmentation des contributions publiques. En 2016, le déficit de la protection sociale se réduit de nouveau significativement et atteint son niveau le plus bas depuis la crise de 2008. Le déficit de la protection sociale se réduit de nouveau en 2016 pour s'élever à 0,4 milliard d'euros, contre 3,3 milliards d'euros en 2015. C'est le déficit le plus bas observé depuis le début de la crise économique et financière de 2008, l'année 2009 ayant marqué, sous l'effet de la récession, le retour à des déficits importants de la protection sociale.

■ Les dépenses de la protection sociale

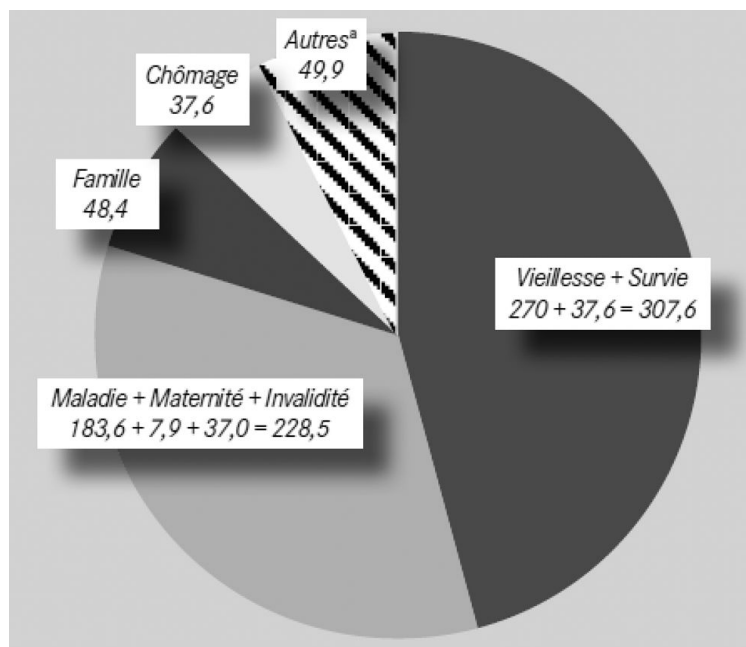
Les dépenses de protection sociale s'élèvent à 759,1 milliards d'euros en France en 2016 selon les chiffres de la DREES. Les prestations de protection sociale représentent 94,1% de l'ensemble des dépenses de protection sociale, soit 714,5 milliards d'euros en 2016. Les autres dépenses sont constituées de frais de gestion, de frais financiers, etc. Les prestations de protection sociale progressent de +1,8% en 2016, contre +1,7% en 2015. Si les prestations de protection sociale ont progressé à un rythme soutenu entre 2000 et 2010 (+4,3% par an en moyenne), leur croissance est nettement plus modérée depuis : elles augmentent de 2,8% par an entre 2010 et 2014, puis de moins de 2% par an en 2015 et 2016. On peut considérer que les mesures de maîtrise des dépenses, en particulier la montée en charge des réformes des retraites et le respect de l'Objectif national de dépenses d'assurance maladie (ONDAM), depuis 2010,

ont contribué au ralentissement de la croissance des prestations de protection sociale en France.

La France et les pays du nord de l'Europe consacrent plus de 30% du PIB à la protection sociale. Toutefois, l'importance des systèmes de protection sociale varie fortement entre les pays de l'UE-15 d'une part, où les prestations représentent en moyenne 28,3% du PIB, et les 13 autres États membres d'autre part, où la moyenne atteint 18,0% du PIB. En France, en 2016, les prestations de protection sociale représentent 32,1% du PIB.

La santé et la vieillesse-survie constituent de loin les principaux postes de dépenses, en France comme en Europe. La répartition des dépenses de prestations sociales entre les six grands risques sociaux reflète les caractéristiques sociodémographiques des pays et les priorités retenues par les États membres de l'Union européenne. Les dépenses au titre du risque vieillesse-survie, qui comprennent principalement le versement des pensions de retraite, constituent la part la plus importante du total des prestations dans la quasi-totalité des pays européens (46% au sein de l'UE-28). Le vieillissement démographique contribue en effet largement à la croissance des dépenses du risque vieillesse-survie, avant comme après la crise de 2008. En consacrant 46% du total des prestations à ce risque, la France se situe ainsi dans la moyenne européenne. Le risque santé est le 2^e poste de dépenses de prestations. Il représente en moyenne 37% du total des prestations de l'UE-28 et 35% du total des prestations en France. En France comme en Europe, les autres risques sociaux représentent des parts nettement moindres de la dépense. Ainsi, en France en 2016, le risque famille constitue 8% des prestations sociales, le risque emploi 6% et les risques logement et pauvreté-exclusion sociale représentent chacun 3% du total des prestations sociales.





Note : a = (Autres) = Pauvreté-exclusion sociale (16,8) + Logement (17,7) + Insertion-réinsertion professionnelle (3,4) + Accidents du travail (12,2).

Source : Données issues des Comptes de la protection sociale 2015, données 2013. Gouverner la protection sociale : transparence et efficacité, Les notes du conseil d'analyse économique, n° 28, janvier 2016.

GRAPHIQUE 13.4.

Décomposition de la dépense de protection sociale, en milliards d'euros

17 Cette partie 13.3 consacrée à la redistribution s'appuie sur les éléments présentés dans le chapitre 4 : parties 4.7 et 4.8.

L'ÉCONOMIE OUVERTE : COMMERCE INTERNATIONAL ET MONDIALISATION

« Le commerce guérit des préjugés destructeurs ; et c'est presque une règle générale, que partout où il y a des mœurs douces, il y a du commerce ; et que partout où il y a du commerce, il y a des mœurs douces. » - Charles de Montesquieu, 1689-1755

« Aucune nation n'a jamais été ruinée par le commerce. » - Benjamin Franklin, 1774.

« Supposons deux ouvriers sachant l'un et l'autre faire des souliers et des chapeaux : l'un d'eux peut exceller dans les deux métiers ; soit en faisant des chapeaux, il ne l'emporte sur son rival que d'un cinquième ou de 20%, tandis qu'en travaillant à des souliers, il a sur lui un avantage d'un tiers, ou de 33%. Ne serait-il pas de l'intérêt de tous les deux que l'ouvrier le plus habile se livrât exclusivement à l'état de cordonnier, et le moins adroit à celui de chapelier ? » - David Ricardo, 1817.

« Lorsqu'on a suffisamment démontré que les gains attendus de l'intervention sont faibles, l'économie politique peut être invoquée pour justifier que l'on renonce à toute intervention. Le libre-échange peut être le noyau d'un accord entre les pays pour éviter la guerre commerciale » - Paul Krugman, 1998.

[14.1 les fondements du commerce international](#)

[14.2 Les théories du commerce international](#)

[14.3 L'avenir de la mondialisation économique
et commerciale](#)

14.1 LES FONDEMENTS DU COMMERCE INTERNATIONAL

14.1.1 *La mondialisation contemporaine : une définition*

Dans son ouvrage intitulé *La Grande Désillusion*, le prix Nobel d'économie Joseph Stiglitz a défini la mondialisation comme « *l'intégration plus étroite des pays et des peuples du monde qu'ont réalisée d'une part la baisse des coûts de transport et de communication, et d'autre part la destruction des barrières artificielles à la libre circulation des biens, des services, des capitaux et, dans une moindre mesure, des personnes* ». La mondialisation est un processus qui implique une interdépendance croissante des économies nationales dans le cadre d'un marché planétaire : traduction du terme anglais "*globalization*", apparu dans la presse financière américaine au milieu des années 1980, il s'agit d'un phénomène multidimensionnel qui recouvre les échanges internationaux de biens et de services, la mondialisation de la production sous l'effet des stratégies des firmes multinationales, mais aussi l'unification des marchés de capitaux et la circulation des informations et des idées. Tous ces éléments se combinent, se cumulent et font système pour édifier une économie mondialisée.

Le processus d'intégration économique internationale s'est accéléré ces dernières années, notamment sous l'effet de nombreux accords internationaux passés dans le cadre de l'OMC. Il s'agit tout à la fois de progrès en matière d'intégration commerciale, mais aussi d'intégration financière, tandis que les lieux de concentration des échanges ont profondément évolué, avec le développement des grandes économies émergentes (Brésil, Russie, Inde, Chine et Afrique du Sud).

La globalisation de l'économie a également posé avec plus d'acuité des problèmes globaux : les questions traitées jusque-là dans la limite des territoires nationaux ne peuvent plus dès lors trouver de solutions viables que dans le seul cadre des États nations, qu'il s'agisse des menaces sur l'environnement mondial, de la dégradation des ressources naturelles, de la croissance de la population mondiale, des risques de contamination sanitaire et de l'instabilité des marchés financiers. Une certaine ouverture des économies est aujourd'hui nécessaire, car aucun pays ne peut prétendre disposer de toutes les dotations factorielles et des ressources nécessaires pour produire l'ensemble des biens désirés par les consommateurs. C'est la raison pour laquelle l'accroissement des échanges mondiaux est la caractéristique principale de la mondialisation (on ne peut importer sans exporter quelque chose) : lorsque l'on parle de « commerce extérieur », on élargit le champ pour passer de celui de « l'entreprise » à celui de la « nation » (ce qu'on appelle par convention le « reste du monde » dans la comptabilité nationale).

L'intrication des différentes économies est désormais très forte, et on évoque aujourd'hui l'émergence d'une économie mondialisée, notamment grâce à la baisse des coûts de transport et aux progrès des moyens de communication. La mondialisation est souvent associée au seul commerce international. Si celui-ci est naturellement l'une de ses manifestations, l'interdépendance des économies créée par la mondialisation dépasse largement le simple domaine des échanges de biens et services. Cette intrication des économies comprend également la circulation des facteurs de production, le capital financier, les taux de change, c'est-à-dire les devises, le capital humain, les différents types de main-d'œuvre et la technologie. L'ouverture croissante des économies a également accru la synchronisation des cycles économiques et une transmission rapide et violente des dérèglements économiques (on parle parfois de « contagion »). La recherche récente montre que la politique

monétaire américaine influence le cycle du crédit partout dans le monde.

14.1.2 *Une brève histoire du commerce international depuis 1945*

14.1.2.1 L'essor du libre-échange après 1945

Sous l'effet combiné du progrès technique, des cadres institutionnels et politiques favorables aux échanges internationaux (GATT puis OMC en 1995) et du dynamisme des firmes multinationales, le commerce mondial a connu une croissance exceptionnelle au cours de la seconde moitié du XX^e siècle.

Les exportations de marchandises ont augmenté de 6% par an en moyenne durant la période de l'après-guerre. En 1970, le commerce extérieur rapporté au PIB mondial a retrouvé le niveau élevé qu'il atteignait au début du XX^e siècle, soit 8%. Entre 1955 et 1975, la valeur des exportations mondiales a été multipliée par plus de neuf, alors que la production mondiale a « seulement » été multipliée par quatre. Malgré une certaine baisse de régime de croissance à la fin des « Trente Glorieuses », la contribution du commerce international au PIB mondial a globalement augmenté, ce qui est le signe d'une intégration croissante des économies de la planète.

Dans les années 1960, la révolution des transports maritimes s'accélère (essor de la flotte marchande de porte-conteneurs et de superpétroliers, pour pallier notamment les obstacles liés aux conflits régionaux qui bloquent certaines routes, comme le canal de Suez en 1956). Cette mutation impose l'aménagement des ports maritimes et réduit le coût du fret et l'obstacle de la distance. Dans un contexte de décolonisation et de suppression des préférences impériales en matière de commerce, les négociations multilatérales internationales qui se succèdent (Dillon Round puis Kennedy Round) favorisent une **baisse spectaculaire des protections douanières** : si

elles correspondaient à 40% du prix des produits importés juste après la guerre, elles chutent aux environs de 10% en 1967.

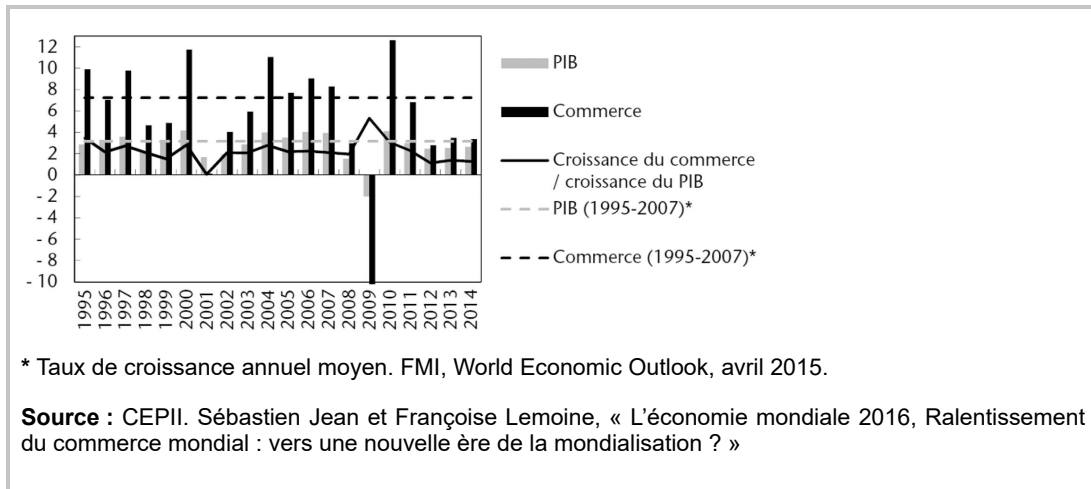
De 1948 à 1973, le commerce international en volume progresse de 8% par an en moyenne et, durant cette période des « trente glorieuses », la croissance annuelle des échanges mondiaux est deux fois plus forte en moyenne que celle de la production. Le commerce international stimule aussi bien la demande (nouveaux débouchés pour les firmes) que l'offre (incitation des firmes à innover face à une compétition plus vive, dépenses accrues de recherche et développement), et l'expansion commerciale a largement contribué à l'essor économique de pays comme le Japon ou l'Allemagne fédérale.

L'intensification de la concurrence mondiale a poussé les entreprises européennes à réaliser des efforts de modernisation pour améliorer leur compétitivité prix (en diminuant les coûts unitaires de production et en réalisant des gains de productivité), autant que leur compétitivité hors prix (innovation, qualité des produits, marketing, réputation de la marque, etc.)

14.1.2.2 Les mutations de la division internationale du travail à partir des années 1970

Le début des années 1970 est marqué par une période d'instabilité pour les économies des pays développés, en raison d'une accélération de l'inflation, de la montée du taux de chômage, de la montée du prix du baril de pétrole (quadruplement en 1973, suivi d'une nouvelle forte hausse en 1979), tandis que le prix de nombreuses matières premières progresse également, même si dans des proportions plus limitées. Dans les pays industrialisés, la croissance du PIB a significativement diminué, en passant de 2,8% par an en moyenne de 1973 à 1979, à 2,2% en moyenne de 1980 à 1985. Dans un contexte d'intensification des échanges internationaux et de dépendance énergétique de nombreux

pays occidentaux, l'augmentation de la facture pétrolière conduit à une hausse des coûts de production des entreprises, qui a érodé les marges bénéficiaires et incité les firmes à relever leurs prix de vente.

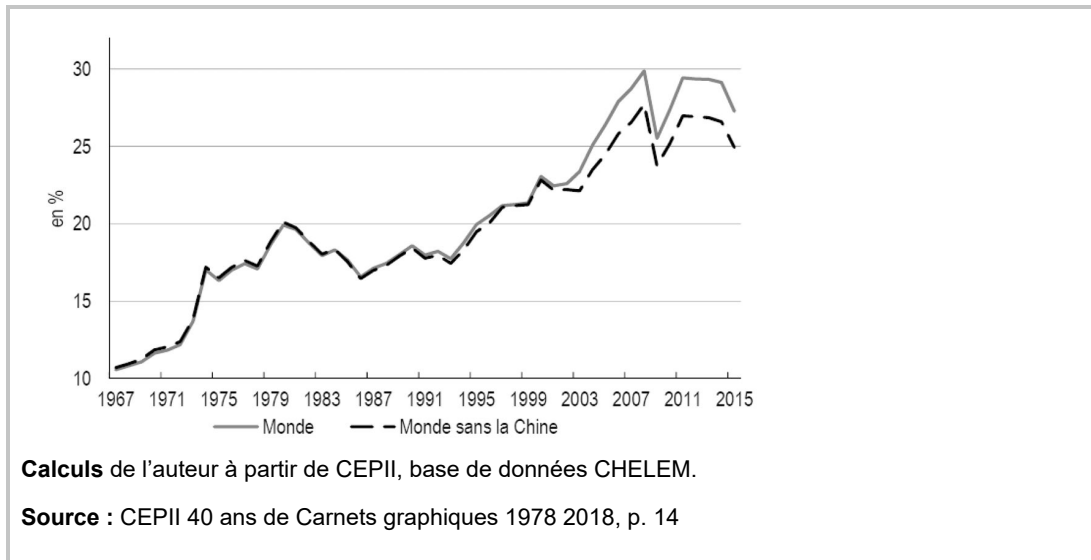


GRAPHIQUE 14.1.

Croissance en volume du PIB et du commerce de biens et services, au niveau mondial (en %)

Le commerce mondial reste cependant dynamique durant cette période, et poursuit une progression plus rapide que la production mondiale : les droits de douane continuent de diminuer (avec une baisse de 35% durant les années 1980), tandis que les pays développés adoptent des engagements réciproques qui visent à réduire les obstacles non tarifaires, comme l'harmonisation des normes techniques et des règlements sanitaires ou l'ouverture des marchés publics. Malgré les crises conjoncturelles et les incertitudes économiques nées de divers « chocs » (chocs pétroliers des années 1970, krach boursier de 1987, crise monétaire européenne et crise des monnaies asiatiques dans les années 1990, éclatement de la bulle internet en 2000), le commerce international a continué à croître à un rythme soutenu, en dépit de l'essoufflement de la croissance économique. La crise et la récession mondiale de 2009 ont certes provoqué une chute du commerce mondial, mais la croissance des échanges s'est poursuivie par la suite.

L'OMC prévoyait ainsi pour 2018 une croissance de 4,4% du volume des échanges de marchandises, mesuré par la moyenne des exportations et des importations, soit une progression stable par rapport à la croissance de 2017 (4,7%). Mais pour l'OMC, le taux de croissance devrait retomber à 4,0% en 2019, en dessous du taux moyen de 4,8% enregistré depuis 1990, mais toujours nettement au-dessus de la moyenne d'après crise (3,0%). Toutefois, certains signes indiquent que l'escalade des tensions commerciales, notamment entre la Chine et les États-Unis, pourrait déjà affecter la confiance des entreprises et les décisions d'investissement, ce qui pourrait compromettre les perspectives actuelles.



GRAPHIQUE 14.2.

Taux d'ouverture commerciale mondiale

Malgré les différentes récessions qui ont affecté les principales économies du monde depuis le début des années 1970, le taux d'ouverture commerciale mondiale reste clairement sur une tendance à la hausse : ce taux d'ouverture passe d'environ 15 % en 1974 à environ 27 % en 2015.

14.1.2.3 L'accélération de la mondialisation : vers une économie globalisée

À partir de la seconde moitié des années 1990, l'économie mondiale a été profondément remodelée par une nouvelle vague d'innovations dans les télécommunications, l'informatique et les réseaux mondiaux d'information auxquels elles ont donné naissance. Grâce à la fibre optique, aux satellites et à la technologie numérique, le coût des télécommunications à longue distance a diminué drastiquement pour devenir quasiment nul. Les puces électroniques de plus en plus puissantes (selon la « loi de Moore », la puissance des circuits intégrés double à peu près tous les deux ans) ont aussi réduit de façon spectaculaire le coût de la puissance de calcul. Les technologies de l'information et de la communication (TIC) ont permis d'élever la productivité des facteurs de production et de stimuler le commerce : c'est vrai des ventes aux consommateurs (*business to consumer, B to C*), mais l'enjeu est encore plus grand pour les transactions interentreprises (*business to business, B to B*). Par ailleurs, le commerce international est indissociable d'une division internationale du travail (DIT), c'est-à-dire d'une spécialisation des économies nationales dans des activités particulières. C'est cette spécialisation qui donne lieu à des échanges commerciaux. La spécialisation économique internationale désigne le fait que chaque État se spécialise dans la production de certains biens pour en exporter une partie et importe les biens dont il a abandonné la production à d'autres pays.

Historiquement, l'économie mondiale est passée **d'une DIT traditionnelle** au XIX^e siècle, dans le cadre de laquelle les pays pauvres demeurent spécialisés dans les produits primaires (richesses minières et agricoles), tandis que les pays industrialisés (européens pour la plupart) fabriquent des produits manufacturés, **à une « nouvelle » division internationale du travail**. Dans le cadre de cette dernière, les pays en développement ne se contentent plus d'exporter des produits primaires, mais vont monter progressivement en gamme en fabriquant des biens de plus en plus élaborés

(textile et petite électronique, mais aussi automobiles, informatique, téléviseurs, etc.).

Avec l'**émergence des BRICS** (Brésil, Russie, Inde, Chine et Afrique du Sud), la constitution de cette nouvelle DIT s'explique par le développement des échanges intra-branches (où les pays s'échangent des produits similaires, mais différenciés) et des échanges intra-firmes (soit l'ensemble des échanges à l'intérieur du réseau même d'une entreprise multinationale).

Le **commerce des services** a connu lui aussi une véritable explosion, du fait de la réduction spectaculaire du coût des communications. Même s'il demeure minoritaire en proportion des échanges de biens, son potentiel de progression est très important, puisque les services représentent en moyenne 70% du PIB dans les pays développés. Avec les technologies de l'information et de la communication (TIC), de nouveaux services dits « intermédiaires » ont pu être externalisés : travail de bureau, opérateurs informatiques, traitement des données, programmation, recherche scientifique et ingénierie. Leur internationalisation récente a aussi pu se déployer grâce à l'émergence dans les pays à bas coût d'une main-d'œuvre plus qualifiée (c'est le cas de l'Inde avec les services informatiques, par exemple dans la région de Bangalore). D'importants secteurs, qui demeuraient abrités de la concurrence internationale, tels que les services bancaires, le commerce de détail, la médecine ou l'enseignement, sont en mesure aujourd'hui de devenir les secteurs les plus ouverts à l'échange international, grâce au développement des services en ligne notamment. Concurrencés sur l'industrie et les biens manufacturés par les pays émergents, les pays développés ont eu tendance à se positionner sur le commerce des services à forte valeur ajoutée, avec des progrès de la déréglementation du secteur, notamment dans le cadre de l'Union européenne (transports, énergie, télécommunications). Aujourd'hui, toutefois, les marchés des services et du travail restent encore majoritairement

nationaux et cela s'explique en particulier par l'importance des facteurs culturels dans ces secteurs.

L'essor des échanges internationaux est avant tout le résultat de **progrès de la régionalisation du commerce**. Elle désigne la tendance des pays à commercer avec d'autres pays qui appartiennent à la même zone géographique (que cette zone soit organisée en union régionale ou non). Il existe donc une tendance à la concentration géographique des flux commerciaux. L'analyse économique permet de montrer que le commerce international entre deux pays est proportionnel aux montants de leur PIB respectif (effet lié à la taille des économies) et inversement proportionnel à la distance géographique entre ces pays. Pour mieux rendre compte du commerce international, il faut prendre en compte un « effet frontière ». Des études ont montré par exemple que les échanges étaient plus intenses entre régions voisines d'un même pays (par exemple les États des États-Unis situés le long de la frontière canadienne) qu'entre régions situées de part et d'autre d'une frontière. À distance égale, un État des États-Unis commerce davantage avec un autre État des États-Unis qu'avec une province canadienne.

Le **régionalisme** commercial repose sur la coopération – en matière économique et commerciale – entre États appartenant à une même aire d'échanges. À la différence de la **régionalisation** (qui est lié à la proximité géographique) le régionalisme suppose une volonté politique et la mise en place d'institutions qui favorisent le commerce entre pays voisins. Depuis la fin de la Seconde Guerre mondiale, une centaine d'organisations économiques régionales ont été créées.

De façon générale, le mouvement d'intégration régionale accompagne l'ouverture au commerce et il a souvent pour motivation une volonté des pays membres de chaque union de peser sur l'organisation de l'économie mondiale en faisant davantage entendre leurs voix dans les négociations internationales.

L'analyse du commerce international permet ainsi de dégager l'existence de **trois zones dominantes** :

- La zone américaine structurée autour de l'ALENA
- La zone européenne
- La zone asiatique

La zone asiatique, et notamment l'Asie de l'Est et du Sud Est, connaît une croissance importante de sa part dans le commerce mondial

14.2 LES THÉORIES DU COMMERCE INTERNATIONAL

14.2.1 *La doctrine libérale : l'ouverture aux échanges comme gain mutuel*

14.2.1.1 La théorie des avantages absolus selon Adam Smith

Contre la conception mercantiliste qui préconise un protectionnisme au nom de la défense des intérêts nationaux, les économistes classiques (libéraux) considèrent à partir du XVIII^e siècle que le libre commerce crée des liens d'intérêt et d'échanges mutuellement profitables pour les États nations, qui deviendront ainsi les membres d'une « société universelle ». Dans la tradition des philosophes des Lumières, le « doux commerce » sert de substitut à la guerre et constitue un vecteur de civilisation des mœurs : « *C'est presque une règle générale que partout où il y a des mœurs douces, il y a du commerce ; et que partout où il y a du commerce, il y a des mœurs douces [...]* L'effet naturel du commerce est de porter à la paix », écrit Montesquieu en 1748 dans *De l'esprit des lois*. Des auteurs classiques comme John Stuart Mill ou Frédéric Bastiat ont fait la promotion de l'échange international comme facteur de paix, et ont dénoncé les mesures protectionnistes, qui exacerbent les antagonismes commerciaux et sont porteuses de risques de conflits. Le libre-échange mondial devrait permettre, à

l'instar d'un marché classique où échangent des individus rationnels, de bénéficier aux producteurs et aux consommateurs de tous les pays engagés dans un jeu à somme positive. La division internationale du travail sur la base de la spécialisation productive permet alors de créer suffisamment d'interdépendances pour garantir l'harmonie naturelle des intérêts, au sens de la métaphore de la « main invisible » employée en 1776 par Adam Smith. La division du travail est au cœur de la dynamique de croissance à travers les gains de productivité, mais également les apprentissages et les innovations qui en découlent. Adam Smith développe en 1776 la théorie des avantages absolus : un pays dispose d'un avantage absolu dans la production d'un bien s'il est plus efficient dans la production de ce bien. En conséquence, Adam Smith préconise que chaque pays se spécialise dans la production des biens pour lesquels il dispose d'un avantage absolu, abandonnant ainsi la production des autres biens. Il démontre qu'en divisant ainsi le travail, il existe un gain mutuel à l'échange : la spécialisation engendre en effet des gains de productivité qui permettent, à quantité de facteurs de production donnée, de produire plus qu'en situation d'autarcie (c'est-à-dire en situation où il n'y a pas d'échange). Chaque individu cherche à satisfaire son intérêt personnel, mais, indirectement et involontairement, « conduit par une main invisible », il permet aux autres d'augmenter également leur bien-être.

Avec cette division internationale du travail vertueuse, Milton et Rose Friedman font ainsi référence au gain à l'échange et à « l'harmonie naturelle des intérêts » qu'évoquait Adam Smith en 1776 avec sa métaphore de la « main invisible » : l'échange volontaire et le mécanisme des prix permettent de satisfaire les besoins des consommateurs.

Le principe de spécialisation, selon l'avantage absolu, que propose Smith peut conduire à une impasse et à l'impossibilité d'échanger. En effet, si un pays n'a que des désavantages absolus, il ne peut échanger, car comment

importer sans rien exporter ? C'est une faille majeure que va relever Ricardo afin d'y apporter une solution.

14.2.1.2 La théorie des avantages comparatifs : le modèle de David Ricardo

Dans le chapitre VII de ses *Principes de l'économie politique et de l'impôt* (1817), David Ricardo défend les analyses anti-mercantilistes d'Adam Smith, fondées sur la théorie des avantages absolus selon laquelle chaque pays a intérêt à se spécialiser dans la ou les production(s) où il est le plus efficace. Mais, ayant davantage recours à l'abstraction théorique qu'Adam Smith, il les prolonge avec sa « loi » des avantages comparatifs (ou relatifs) : dans le cadre d'un modèle simplifié à deux pays (le Portugal et l'Angleterre) et à deux biens (le vin et le drap), il démontre que l'échange international peut être mutuellement avantageux, même dans l'éventualité où l'une des deux nations est plus efficiente que l'autre dans la production des deux marchandises échangées, dès lors que chacune se spécialise dans les produits pour lesquels elle est la plus avantagée ou la moins désavantagée (coûts unitaires, productivité du travail).

Autrement dit, même en l'absence d'avantage absolu, un pays peut se spécialiser là où il dispose d'un avantage comparatif, là où son avantage absolu est le plus grand ou là où son désavantage absolu est le plus faible. L'échange international redevient alors possible. Chaque pays, quel que soit son niveau de développement, gagne à se concentrer dans le ou les produit(s) en vertu de ses avantages relatifs, et a intérêt à abandonner les autres productions : si chaque nation ne produit que ce pour quoi elle est (relativement) la plus douée, elle produit une quantité de biens et services supérieure et l'échange international devient un jeu à somme positive (le gain mondial est accru), et permet de reculer l'avènement de l'état stationnaire, lorsque la croissance économique s'arrête. Si le théorème de Ricardo est fondé sur des hypothèses restrictives qui en limitent le champ de validité (valeur des produits déterminée par la quantité de

travail utilisée, immobilité internationale des facteurs de production et rendements constants), la rigueur scientifique de sa démonstration en fait la clé de voûte de la théorie libérale du commerce international.

TABLEAU 14.1.
Les avantages comparatifs selon David Ricardo

	Drap	Vin
Angleterre	100	120
Portugal	90	80

Pour expliquer sa théorie, David Ricardo donne l'exemple suivant : supposons que le Portugal doit utiliser 90 heures de travail pour produire une pièce de drap et 80 heures pour un litre de vin, contre 100 heures pour une pièce de drap en Angleterre et 120 heures pour le vin.

Le Portugal dispose d'un avantage absolu dans la production de vin et de drap. Pour autant, le Portugal a intérêt à abandonner la production de drap pour laquelle l'Angleterre dispose d'un avantage comparatif et à se spécialiser dans la production de vin pour laquelle le Portugal est relativement plus efficient (par rapport au drap).

En effet, en se spécialisant selon ces principes, l'Angleterre utilisera les 120 heures libérées par l'abandon de la production de vin à la production de drap. L'Angleterre disposera ainsi de 220 heures au lieu de 100 pour produire des draps, et pourra ainsi en produire 2,2 unités.

De même, le Portugal utilisera les 90 heures libérées par l'abandon de la production de drap, pour les utiliser à la production de vin : il disposera ainsi de 170 heures pour la production de vin et celle-ci sera alors de 2,125 unités. Au total, 2,2 draps sont produits contre 2 auparavant et 2,125 litres de vin contre 2 auparavant. La spécialisation et l'échange ont entraîné un gain. La façon dont ce gain est réparti va dépendre du prix des biens. Les gains à l'échange peuvent ainsi être répartis de façon très inégale entre les pays.

14.2.1.3 La théorie néoclassique du commerce international

Les théoriciens néoclassiques (les Suédois Heckscher et Ohlin et l'Américain Paul Samuelson) prolongent les analyses de David Ricardo au cours des années 1930 en montrant que les pays doivent se spécialiser en fonction de leurs dotations factorielles relatives. Un pays qui disposerait relativement de plus de capital que de travail devrait, dans cette optique, se spécialiser dans la production de biens qui demandent relativement plus de capital que de travail pour être produits. Selon le théorème « H.O.S », l'extension du libre-échange permet la convergence mondiale de la rémunération des facteurs de production (salaires, profits).

En effet, dans les pays spécialisés dans des productions nécessitant l'emploi d'une main-d'œuvre abondante (comme aujourd'hui la Chine), la hausse des exportations permet une hausse de la production et un accroissement de la demande de travail dans le pays. Ceci doit engendrer une hausse des salaires, qui se rapprochent alors de ceux pratiqués dans les pays développés (rémunération des ouvriers américains).

Ils justifient le libre-échange en avançant que si chaque pays se spécialise dans la production du bien demandant relativement beaucoup de facteurs abondant dans ce pays, l'ensemble des partenaires à l'échange bénéficiera d'une situation meilleure que l'autarcie, c'est-à-dire l'absence d'échanges. Les prix des biens et services baisseront, et les prix des facteurs de production (travail et capital) convergeront, avec une tendance à l'égalisation des rémunérations entre les pays.

HO (HECKSCHER-OHLIN) EST LE MODÈLE NÉOCLASSIQUE DU COMMERCE INTERNATIONAL : LA « THÉORIE PURE » DE LA PROPORTION DES FACTEURS.

Alors que dans le modèle de Ricardo, l'avantage comparatif ne dépend que des différences de productivité, les Suédois ont montré qu'il pouvait résulter de la différence de rareté relative de facteurs de production entre les pays. La rareté (ou l'abondance) relative d'un facteur ne dépend pas des quantités de facteurs, mais du rapport entre ces quantités. Le théorème HO s'énonce de la manière suivante : les pays exportent les produits dont la production requiert l'utilisation intensive du facteur abondant et importent ceux qui utilisent intensivement les facteurs rares. La démonstration d'Heckscher et Ohlin repose des hypothèses : pas d'obstacles aux échanges de marchandises, absence de coûts de transport, immobilité des facteurs de production, concurrence parfaite sur le marché des produits et des facteurs, les technologies et les facteurs de production sont identiques dans les deux pays, les techniques de production utilisées sont différentes et dépendent du prix relatif des facteurs, les préférences des consommateurs sont identiques dans les deux pays.

HOS (HO-Samuelson) représente la **théorie de l'égalisation internationale du prix des facteurs**.

Le commerce international conduit sur un marché concurrentiel à une affectation optimale des ressources au niveau international, ce qui débouche sur une égalisation du prix des facteurs. Cette théorie prédit donc un phénomène de convergence du prix des facteurs de production sous l'effet du commerce international. C'est le théorème de Stolper-Samuelson : dans un pays qui se spécialise dans la production utilisant le facteur le plus abondant, la demande de ce facteur augmente et son prix s'élève. Au contraire, le prix du facteur rare, moins demandé, tend à augmenter (théorème de Stolper-Samuelson). Selon Samuelson, si le libre-échange conduit à l'unicité du prix des biens entre les pays, il tend à égaliser le prix des facteurs dans les différents pays.

HOV (HO-Vanek) représente la **théorie de la généralisation du raisonnement HO à plusieurs pays et plusieurs produits**.

La théorie contemporaine de l'économie internationale considère que les gains du commerce international sont cumulatifs : l'ouverture internationale entraîne des

avantages comparatifs qui permettent ensuite une plus grande ouverture, etc.

Pour la majorité des économistes, l'échange international procure en effet trois types d'avantages : un effet de dimension, un effet de diversification et un effet de concurrence.

- L'ouverture aux échanges crée un avantage comparatif, car elle permet la spécialisation et engendre donc un effet de dimension : chaque nation peut produire en plus grande quantité certains biens, ce qui offre des avantages comme les économies d'échelle (l'entreprise réduit ses coûts unitaires en produisant davantage) ou les effets d'apprentissage.

- L'ouverture aux échanges crée un effet de diversification. Grâce à l'ouverture internationale, le consommateur peut choisir entre un nombre plus important de produits pour satisfaire un même besoin. Cette diversité de biens disponibles profite non seulement aux consommateurs, mais aussi aux producteurs, qui auront un choix supplémentaire en biens de production.

- L'ouverture aux échanges crée un effet de concurrence. L'ouverture internationale permet à de nouvelles entreprises d'entrer sur les marchés nationaux, ce qui accentue la concurrence ; les prix et le niveau de production deviennent donc plus efficaces, car plus proches de ceux qui résulteraient d'un équilibre de marché.

14.2.2 *Les coûts économiques et sociaux du libre-échange*

14.2.2.1 L'ouverture internationale peut aggraver les déséquilibres internes

Certains économistes ont insisté sur les coûts économiques et sociaux de la libéralisation des échanges, à l'origine d'une mise en concurrence des systèmes productifs, et la constitution d'un marché mondial du travail, par catégories de

travailleurs, et des salaires fixés à l'échelle planétaire. Si l'on suit le raisonnement de David Ricardo, la spécialisation internationale est globalement favorable à tous les participants aux échanges : mais si la Chine se spécialise dans les biens de consommation et les États-Unis dans les biens de production, l'ouvrier américain travaillant dans l'habillement ou l'électronique grand public perd son emploi et peut éprouver des difficultés à en retrouver un dans sa région et selon sa qualification. La mondialisation intensifie le processus de création-destruction d'emplois et provoque des vagues de délocalisations destructrices d'emplois dans les secteurs soumis à la pression de plus en plus forte de la concurrence des pays à bas salaires. La concurrence se fait aussi par le niveau de la fiscalité, et la mondialisation pourrait entraîner dès lors un alignement par le bas et une paupérisation des États. Face à l'essor des pays émergents à bas salaires et les pressions exercées par leur important réservoir de main-d'œuvre (qui exerce une forte pression sur les coûts salariaux des pays développés), certains économistes préconisent aujourd'hui la mise en œuvre de nouvelles mesures protectionnistes (réglementations, taxes, droits de douane). Le prix Nobel d'économie Maurice Allais dénonçait ainsi le « dogme du libre-échange » et préconisait de recréer des ensembles régionaux plus homogènes (c'est-à-dire unissant des pays présentant les mêmes conditions de revenus et de protection sociale), qui auraient le droit de prendre des mesures protectionnistes plus raisonnables pour se protéger contre les écarts trop importants de coûts de production, tout en maintenant une concurrence réelle à l'intérieur de la zone.

14.2.2.2 Vertus et limites de la mondialisation

Dans le cadre du libre-échange mondial, la concurrence de la main-d'œuvre à bas salaires des pays émergents constitue-t-elle une menace pour le volume de l'emploi et le niveau des salaires dans les pays du nord ? Ce débat est devenu très vif avec le thème de la désindustrialisation et des vagues de

délocalisations de certaines activités des pays développés vers ces pays à bas salaires. Une délocalisation est une stratégie d'entreprise consistant à fermer une ou plusieurs unités de production dans un pays donné et à implanter une ou des unités de production équivalentes dans un ou plusieurs autres pays, pour bénéficier de conditions de production jugées plus favorables (coûts de la main-d'œuvre, exonérations fiscales, etc.).

La concurrence avec les pays émergents et les effets des délocalisations (sous la forme d'investissements directs étrangers) suscitent des craintes dans les pays développés quant à la mondialisation, régulièrement accusée d'être responsable de la hausse du chômage et du creusement des inégalités salariales :

- Les entreprises délocalisent pour acquérir un avantage de compétitivité, ce qui se traduit par de nombreuses pertes d'emplois sur le sol national (avec un risque de réaction similaire des concurrents de la branche) et une hausse du chômage, notamment pour les travailleurs les moins qualifiés, dont les emplois sont déjà souvent remplacés par du progrès technique. La hausse du chômage entraîne la baisse des revenus, l'affaiblissement de la demande intérieure et celle de la croissance économique. Les délocalisations de sites de production peuvent alors, par ricochet, déstabiliser l'ensemble d'une région (fermetures de commerces, chute du marché de l'immobilier). Les vagues de délocalisations ont surtout touché l'industrie des biens à faible valeur ajoutée (textile, jouet, électronique grand public, appareils électroménagers), en concurrence directe avec les pays à bas salaires, mais aussi désormais les services. On estime qu'aujourd'hui, dans les pays de l'OCDE, un travailleur sur cinq occupe une activité de service potentiellement « délocalisable ». Le volume total d'emplois perdus du fait de la concurrence internationale est difficile à établir, car le lien entre les deux est souvent indirect et diffus.

– La mondialisation est également accusée d'aggraver les inégalités salariales au sein des pays du nord, entre les travailleurs peu qualifiés et qualifiés, puisque la concurrence internationale favorise les destructions d'emplois dans les secteurs intensifs en travail peu qualifié, et donne une prime au travail très qualifié à haut niveau de rémunération.

Mais il convient toutefois de nuancer la responsabilité de la mondialisation dans les destructions d'emploi :

– Les écarts de coûts horaires du travail à l'avantage des pays émergents sont réduits lorsque l'on calcule le coût salarial unitaire : on intègre alors la productivité horaire du travail, soit l'efficacité des heures de travail, c'est-à-dire le rapport entre les richesses créées et les ressources en travail mobilisées pour créer ces richesses.

Si l'on suppose que le coût horaire de la main-d'œuvre est de 150 dans le pays A et de 30 dans le pays, autrement dit dans un rapport de 1 à 5, mais que chaque heure travaillée en A permet de produire 15 unités d'un bien, contre 3 en B, alors la productivité horaire en A est cinq fois supérieure à celle mesurée en B.

Le coût salarial unitaire est alors de dix euros par unité produite dans les deux pays. Ainsi :

TABLEAU 14.2.
Coût horaire, productivité horaire et coût salarial unitaire

	Coût horaire	Productivité horaire	Coût salarial unitaire
Pays A	150	15	$150/15 = 10$
Pays B	30	3	$30/3 = 10$

Dans cet exemple, le différentiel de productivité horaire a annulé le différentiel de coût horaire initial.

– Les firmes qui délocalisent augmentent leur rentabilité, ce qui leur permet de développer de nouvelles activités. Par ailleurs, de nombreuses entreprises ne délocalisent

qu'une partie de leur production dans les pays émergents : les emplois détruits sont souvent ceux des branches à forte intensité de main-d'œuvre et à faible niveau de qualification, comme les activités d'assemblage des biens. Sous la pression de la concurrence, l'entreprise peut améliorer sa productivité et sa compétitivité : les économies qu'elle réalise lui permettent d'effectuer de nouveaux investissements et, au total, ses ventes et son chiffre d'affaires s'accroissent, ce qui peut aussi lui permettre d'embaucher de nouveaux salariés qualifiés dans des activités à forte valeur ajoutée.

– Il faut intégrer l'effet positif du développement des pays à bas salaires à moyen et long terme, grâce aux implantations des firmes occidentales et aux investissements directs étrangers, et donc l'effet de la hausse des exportations des pays émergents, du développement de leur production, de l'amélioration des revenus de leurs consommateurs. Celles-ci entraînent des importations de produits en provenance des pays développés, ce qui crée des emplois et stimule la production dans d'autres secteurs de leurs économies.

– Les produits en provenance des pays émergents sont achetés à des prix plus faibles, ce qui augmente le pouvoir d'achat des consommateurs des pays développés : ce revenu ainsi libéré crée une demande solvable pour d'autres secteurs de l'économie abrités de la concurrence internationale, lesquels sont susceptibles de créations d'emplois, dans les services notamment. Les activités commerciales de services aux entreprises (comptabilité, ressources humaines, centres d'appel, *web design*, logistique, etc.), l'hôtellerie et la restauration, les soins de santé et l'aide aux personnes ont ainsi créé de nombreux emplois sur le sol national. De plus, la mondialisation, avec l'intensification de la concurrence sur le marché des biens, a permis de maintenir une inflation faible et stable pour les consommateurs au sein des pays développés.

– Si des firmes installées en France délocalisent une partie de leur production vers les pays à bas salaires, les firmes étrangères s'implantent également en France, ce qui a des effets bénéfiques pour la croissance et l'emploi. Les investissements directs étrangers (IDE) créent directement des emplois sur le sol national, mais aussi indirectement, en générant des emplois et des revenus dans la zone régionale d'accueil de ces IDE. L'implantation de firmes étrangères accroît aussi la concurrence, ce qui incite les firmes nationales à réaliser davantage d'efforts d'innovation, de productivité et de compétitivité que si elles restaient à l'abri de barrières protectionnistes.

– Enfin, les études statistiques disponibles montrent l'impact limité des délocalisations sur le volume de l'emploi des pays développés (le commerce international demeure concentré entre les pays à niveaux de développement comparable), tandis que l'évolution du taux de chômage est peu corrélée au taux d'ouverture (un certain nombre d'économies très ouvertes du nord de l'Europe ont ainsi de meilleures performances en termes d'emploi que des économies moins ouvertes). De plus, la montée des coûts salariaux dans les BRICS liée à leur développement économique, et la hausse des coûts des transports (en raison de la hausse du prix du pétrole), en annulant le gain financier des délocalisations d'activités, pourraient accélérer le mouvement de relocalisation des industries dans les pays développés, déjà perceptible.

En définitive, la concurrence des pays à bas salaires n'expliquerait au final qu'une modeste part du chômage dans les pays développés : les économistes observent peu de corrélation entre le taux d'ouverture de l'économie et l'évolution du taux de chômage. Le progrès technique et les restructurations internes expliqueraient davantage que le commerce international la réduction en besoin de main-d'œuvre peu qualifiée, comme l'a montré le prix Nobel d'économie Paul Krugman dans son ouvrage intitulé *La mondialisation n'est pas coupable*. D'autres explications

peuvent aussi être mobilisées comme les rigidités institutionnelles sur le marché du travail selon les économistes libéraux, et la faiblesse de la demande effective pour les économistes keynésiens, en raison des politiques d'austérité monétaire et budgétaire.

On considère généralement que la mondialisation engendre des « gagnants » (les travailleurs qualifiés des pays développés) et des « perdants » (les travailleurs peu qualifiés des pays développés), victimes de l'ajustement soit par les quantités (chômage), soit par les prix (baisse des salaires), ou par la dégradation de la qualité de l'emploi et des conditions de travail. Ainsi, le débat sur la mondialisation porterait davantage sur le modèle social souhaité dans la mondialisation, et sur le niveau de redistribution des revenus pour soutenir les travailleurs peu qualifiés et les chômeurs victimes des destructions d'emploi (indemnisation, formation, reconversion). En effet, la théorie économique montre que l'ouverture des économies crée un nouvel équilibre entre les facteurs de production (travail, capital) : le capital devient un facteur très mobile, tandis que le travail l'est beaucoup moins, en raison de barrières culturelles et linguistiques. C'est le travail qui risque alors de devenir la variable d'ajustement (délocalisations, chômage, hausse des impôts, etc.), alors que le capital peut se déplacer partout dans le monde où la rentabilité est la plus élevée.

14.2.3 *Les effets contrastés des politiques protectionnistes*

14.2.3.1 La protection des industries naissantes et vieillissantes

Le protectionnisme désigne l'ensemble des mesures visant à protéger la production d'un pays contre la concurrence étrangère. On distingue les barrières tarifaires, qui sont l'ensemble des taxes (droits de douane) que doivent payer les importateurs du produit soumis à la protection, et les barrières non tarifaires, qui visent à restreindre les flux

d'échange (quotas qui fixent un plafond quantitatif aux importations, obstacles et interdictions administratives, normes techniques visant à limiter les importations). On parle également de protectionnisme « monétaire » lorsqu'un pays agit pour maintenir la sous-évaluation de sa monnaie afin de réduire la compétitivité des importations et de stimuler celle des exportations. Le libre-échange désigne, quant à lui, la politique commerciale qui vise à réduire les barrières tarifaires et non tarifaires afin de stimuler la croissance et l'emploi. Le choix entre libre-échange et protectionnisme (conçus comme deux cas polaires) constitue une question traditionnelle de l'économie politique et renvoie aux débats sur la place à laisser à la régulation marchande : il faut prendre soin de distinguer le protectionnisme de l'autarcie (que personne ne défend), comme d'ailleurs le libre-échange total de l'ouverture internationale, cette dernière étant génératrice de gain de spécialisation, d'accès à des techniques, des capitaux et des débouchés supplémentaires grâce à l'échange. Les auteurs du courant mercantiliste (bullionisme en Espagne, industrialisme colbertiste en France, commercialisme de John Locke en Angleterre et caméralisme porté par les socialistes de la chaire en Allemagne), qui écrivent entre le ^{xvi}^e et le ^{xix}^e siècle, défendent l'idée que la prospérité de la nation et de son souverain dépend de la capacité à freiner les importations et à stimuler les exportations par des mesures de protection de l'économie nationale face à la concurrence étrangère. Si les auteurs classiques comme Adam Smith (1776) rompent avec le mercantilisme au nom du libéralisme économique, toute la palette de l'arsenal protectionniste reste largement utilisée au ^{xix}^e siècle et s'appuie sur un certain nombre de fondements théoriques critiques du libre-échange comme vecteur de croissance et de bien-être.

Dans son *Système national d'économie politique* (1841), l'allemand Friedrich List développe une critique des thèses favorables au libre-échange défendues par l'économiste classique David Ricardo, qu'il accuse de proposer un édifice scientifique (la théorie de l'avantage comparatif) dont le but

politique est de renforcer la suprématie anglaise dans les échanges internationaux. Si List n'est pas fondamentalement hostile au libre-échange et apologétique du protectionnisme, il plaide pour une protection *temporaire* des industries dans l'enfance (ce que l'on appelle le protectionnisme éducateur). Selon lui, le libre-échange ne peut être établi qu'entre pays se trouvant au même stade de développement économique : entre pays se situant à des stades différents, il ne peut que produire une concurrence déloyale qui enferme le pays le moins développé dans son sous-développement (d'autant que toutes les spécialisations ne se valent pas). Les industries naissantes produisent les biens à des coûts supérieurs et ne peuvent générer des économies d'échelle, comme celles des pays les plus avancés : une protection transitoire peut permettre au pays en retard de développement d'améliorer sa compétitivité grâce à des effets d'apprentissage et des gains de productivité.

L'approche de List n'est donc pas contradictoire avec l'insertion dans les échanges internationaux : la protection doit être dégressive et le but poursuivi demeure l'instauration du libre-échange (« le protectionnisme est notre voie, le libre-échange est notre but »). Dans sa *Théorie générale de l'emploi, de l'intérêt et de la monnaie* (1936), John Maynard Keynes propose une réhabilitation au moins partielle de la tradition mercantiliste en relativisant les bienfaits du libre-échange. Dans une économie ouverte, la contrainte extérieure limite l'autonomie des politiques conjoncturelles et paraît condamner les traditionnelles stratégies keynésiennes de relance historiquement situées dans l'espace national.

Le protectionnisme permet alors de renforcer l'efficacité du multiplicateur keynésien en réduisant la propension à importer (économie fermée) : même si les consommateurs sont confrontés à des prix plus élevés et une diversité des biens moins grande, les revenus salariaux peuvent augmenter afin de soutenir la demande intérieure, ce qui n'est guère possible lorsque les entreprises nationales font face à la concurrence des entreprises étrangères et souhaitent

maintenir leur compétitivité. S'amorce alors une boucle vertueuse entre la consommation des ménages, l'investissement des entreprises qui stimule la croissance et l'emploi. L'argumentation en faveur du protectionnisme liée à l'âge de l'industrie s'appuie également sur la défense des industries vieillissantes (ou sénescents) dans les pays développés, lorsque l'industrie a perdu tout avantage par rapport à ses concurrentes. En effet, le déclin d'une industrie a un coût social élevé, en raison du chômage qu'il entraîne et des effets externes négatifs qu'il provoque dans certaines zones du territoire où l'industrie est concentrée (textile, sidérurgie, activités portuaires, construction navale). Des mesures protectionnistes peuvent alors faciliter la réallocation de la main-d'œuvre vers les secteurs plus compétitifs : l'économiste Nicholas Kaldor expliquera, en partant du cas britannique, que le protectionnisme peut s'appliquer aussi aux industries vieillissantes, non pas pour empêcher leur disparition, mais pour réduire le coût économique et social de l'ajustement et faciliter la reconversion de ces activités.

14.2.3.2 Les risques des politiques protectionnistes

Même si les stratégies protectionnistes peuvent conserver un certain attrait dans l'économie mondiale, les économistes mettent en garde les pouvoirs publics contre les effets négatifs de ce type de politique :

- Le protectionnisme peut s'avérer économiquement contre-productif. Dans un système idéal de concurrence pure et parfaite, la protection par des droits de douane réduit le bien-être collectif du pays qui se protège : en effet, le coût pour les consommateurs (qui paient plus cher les biens importés, mais aussi les biens nationaux concurrents, dont les prix s'alignent sur ceux des biens importés) l'emporte sur le gain pour les producteurs du pays et pour l'État (qui perçoit les recettes fiscales liées aux droits de douane). Les mesures protectionnistes entraînent donc un renchérissement du coût de la vie dans les pays qui les adoptent. En théorie, la protection par des

restrictions quantitatives engendre également des pertes, éventuellement supérieures si ce sont les exportateurs étrangers qui bénéficient de la rente due à l'accroissement du prix de vente sur le marché qui applique des protections.

– On considère généralement que le protectionnisme conduit à un affaiblissement de l'incitation à innover dans les pays qui le pratiquent, du fait de la moindre concurrence avec les entreprises étrangères sur le territoire. Le protectionnisme souvent qualifié de temporaire (des industries naissantes ou sénescents), à tort, peut avoir tendance à se transformer en protectionnisme durable et à agir comme une drogue : en raison de la faible incitation à s'améliorer, les firmes nationales, à l'abri de la concurrence, peuvent demander la prolongation de la protection qu'il devient politiquement difficile de lever par la suite.

– Dans des économies largement ouvertes, il est impossible de prendre des mesures protectionnistes au nom de la défense de l'emploi sans s'exposer à des représailles commerciales de la part des partenaires. Ce processus dangereux a déjà été à l'œuvre au cours des années 1930 et les différentes mesures de rétorsion se sont combinées pour produire une forte contraction du commerce mondial, qui a aggravé la crise. En théorie des jeux, les politiques protectionnistes relèvent d'une situation que les économistes appellent le « dilemme du prisonnier » : le protectionnisme au niveau d'un pays n'a de sens que si les autres pays n'agissent pas de manière similaire. Si un pays instaure des mesures de protection à l'encontre d'un partenaire à l'échange, il est alors dans l'intérêt de ce dernier de réagir en élevant lui aussi des barrières. Bien que la situation optimale pour les deux pays soit le libre-échange, si l'un des deux pays s'engage dans le protectionnisme, l'intérêt de l'autre pays est d'agir de même.

– Le renforcement des barrières protectionnistes peut être de nature à freiner le développement des pays émergents, qui ont besoin du commerce international pour soutenir la croissance. La pénétration des investissements directs à l'étranger peut accélérer la croissance de ces pays et enclencher un cercle vertueux bénéfique à toute l'économie (dynamisme de l'emploi, montée en gamme vers les produits à forte valeur ajoutée, réduction des inégalités économiques, etc.). La croissance dégage alors des ressources pour l'État qui peut financer le développement et mettre en place les infrastructures nécessaires au soutien de la croissance à long terme (système éducatif, de santé, infrastructures publiques de transport et de communication, etc.), comme l'explique l'économiste indo-américain Jagdish Bhagwati dans son ouvrage intitulé *Plaidoyer pour la mondialisation* (Odile Jacob, 2010).

– Par ailleurs, le principal problème politique soulevé par la mise en œuvre de mesures protectionnistes est qu'elles ne sont pas appliquées au nom de l'intérêt général, mais servent les intérêts de divers groupes de pression : David Ricardo combattait déjà en son temps les lois sur les blés (*corn laws*) et plaidait pour une ouverture commerciale, afin de desserrer la contrainte sur les profits des industriels dont il était politiquement proche à la Chambre des Communes : l'importation de blés en provenance de pays moins développés constituait un moyen de diminuer le prix du pain et donc les salaires des travailleurs, et permettait de réduire la rente des propriétaires fonciers. Le protectionnisme déclenche en effet des redistributions occultes qui font des gagnants et des perdants dans l'économie : il entraîne la hausse des prix, car des concurrents efficaces sont éliminés du marché, et il se produit alors un transfert de revenu des consommateurs (dont le pouvoir d'achat est réduit) vers les entreprises protégées (dont les profits augmentent sans efforts particuliers des chefs d'entreprise). Dans leur célèbre

manuel d'économie internationale, Paul Krugman et Maurice Obstfeld prennent l'exemple des effets des mesures de protection du marché du sucre aux États-Unis : les producteurs nationaux sont gagnants, de même que les producteurs étrangers (puisque les prix plus élevés sur le marché américain leur assurent une rente), et ce au détriment des consommateurs. Pourtant, si la hausse du prix du sucre pénalise tous les consommateurs, le surcoût ne représente qu'une petite somme supplémentaire pour chacun d'entre eux, et le gain existe surtout pour les producteurs locaux, qui ont su se coaliser pour défendre leurs intérêts. Les groupes de pression seront d'autant plus écoutés par les pouvoirs publics que leur poids politique sera important : l'État peut alors prendre des mesures qui sont contraires à l'intérêt général.

14.3 L'AVENIR DE LA MONDIALISATION ÉCONOMIQUE ET COMMERCIALE

14.3.1 *La mondialisation en perspective historique*

Alors que la tentation du repli sur les espaces nationaux reprend aujourd'hui de la vigueur, l'histoire économique montre que, depuis le XVI^e siècle, le monde a connu différentes phases de mondialisation suivies de phases de fermeture sur les nations : l'ouverture internationale semble déclencher inévitablement des chocs en retour. En effet, si la mondialisation multiplie les flux et les interdépendances, elle entraîne des réactions politiques et des demandes de protection de certaines catégories frappées par la concurrence internationale, comme l'a notamment montré l'historien de l'économie Paul Bairoch.

Une première phase de la mondialisation peut être située à la période des « *Grandes Découvertes* » espagnoles et portugaises du XV^e/XVI^e siècle, avec une réunification de la Terre et la rencontre violente (et tragique) des Amérindiens et

des Européens. Après celle des Croisades, cette époque exprime aussi la fièvre dominatrice et colonisatrice de l'Europe, et la volonté des colonisateurs de trouver de la main-d'œuvre et d'instaurer le travail forcé. Mais l'enjeu de cette toute première phase de la mondialisation est la formation d'une seule humanité agie par les mêmes valeurs et les mêmes croyances, même si des réactions de rejet et de démondialisation se produisent avec des fièvres nationalistes et la formation d'empires. Le libéralisme universaliste s'épanouit toutefois, bouscule les cadres de pensée de l'époque, et conduit aux grandes révolutions politiques de la fin du XVIII^e siècle qui abattent l'absolutisme monarchiste.

La seconde grande période de mondialisation se situe de la fin du XVIII^e siècle à la seconde moitié du XX^e siècle : elle est conduite au XIX^e siècle par l'Angleterre victorienne et ses moteurs sont la finance et l'industrialisation. Ses théoriciens, comme David Ricardo, expriment aussi l'idée centrale du libéralisme, selon laquelle la mondialisation est un jeu à somme positive, tous les pays bénéficiant d'une augmentation de la richesse mondiale, en vertu de la spécialisation et de la division internationale du travail. La multiplication des échanges est alors censée domestiquer les passions nationalistes en vertu du « *doux commerce* » de la doctrine libérale. La révolution industrielle illustre néanmoins les contradictions de l'époque moderne : l'industrie favorise une élévation des niveaux de vie et une amélioration progressive des conditions de vie avec le textile, le chemin de fer ; mais ces progrès sont réalisés avec une certaine brutalité : cycles économiques, chômage, salaires insuffisants, conditions de travail effroyables d'une partie de la main-d'œuvre. Le libéralisme triomphant au XIX^e siècle suscite alors une réaction des masses laborieuses et le creusement des inégalités favorise le succès du courant socialiste, à commencer par celui du socialisme scientifique de Karl Marx (1818-1883). Les années 1850-1914 se traduisent certes par une forte ouverture au monde, avec un flot d'innovations techniques, des flux internationaux d'investissements, de

nouveaux moyens de communication comme le télégraphe... Mais la fin du XIX^e siècle offre déjà le spectacle du déploiement de forces qui poussent à la « *démondialisation* » : politiques protectionnistes, crises diplomatiques, tensions nationales...Jusqu'au cataclysme de 1914. Après la Première Guerre mondiale, et la mondialisation comme jeu à somme nulle (une guerre sanglante de territoires), la quête d'un jeu à somme positive de la mondialisation reprend avec deux grands traités multilatéraux en 1919, et avec la création de la Société des Nations (SDN) et l'Organisation internationale du Travail (OIT). Après le conflit meurtrier de 1914-1918, on voit alors s'ébaucher un embryon de gouvernance mondiale (encadrement de la force, justice internationale, équilibre entre les partenaires sociaux). Après les années 1920 « *rugissantes* » (*roaring twenties*), la Grande Dépression qui ravage les économies américaines et européennes dans les années 1930, après la violente crise boursière de 1929, conduit à une désagrégation de la mondialisation, au retour du protectionnisme commercial des égoïsmes sacrés des nations, et à la fermeture des empires autarciques (Allemagne, Japon), jusqu'à la Seconde Guerre Mondiale, guerre « *totale* ».

Si la Guerre froide (1947-1991) aboutit à la chute de l'URSS, la phase actuelle de mondialisation après la chute de l'URSS a pu laisser penser, à tort, à la « *fin de l'Histoire* » pour reprendre la formule de l'essayiste américain Francis Fukuyama, avec l'universalisme démocratique, l'individualisme et la généralisation de l'économie de marché. Si la nouvelle poussée de mondialisation des économies depuis les années 1990 intègre toutes les activités et les parties du monde avec des interdépendances très fortes entre les nations ouvertes aux flux de biens, de services, de capitaux, elle crée un marché planétaire qui intensifie la concurrence à tous les niveaux, entre États, entreprises, universités, etc. Mais la rapidité de transmission et de propagation des chocs lors des crises économiques légitime

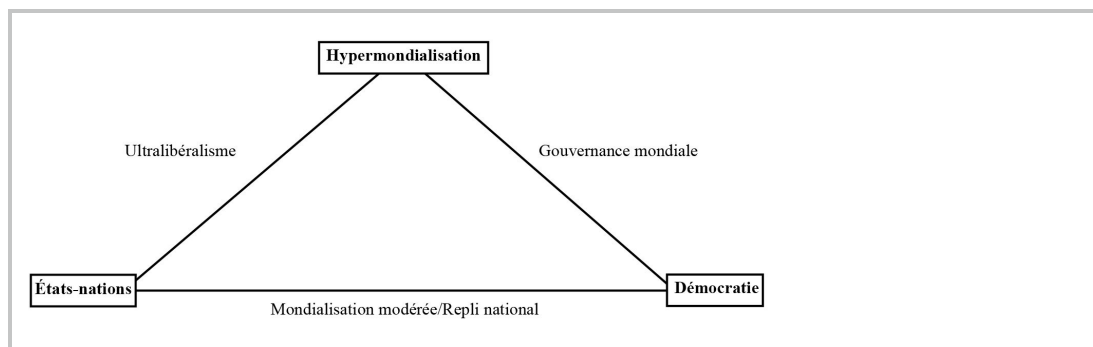
des efforts renouvelés d'édification d'une gouvernance mondiale, comme l'a montré la crise financière mondiale de 2007-2008.

Par ailleurs, la mondialisation permet certes la mobilité des facteurs de production (travail, capital), mais l'individu reste un être de droits qu'en tant que membre d'un État souverain. Et l'État souverain suppose une certaine sédentarisation : l'enracinement et l'identification culturelle à un territoire, à une communauté symbolique, restent vivaces, et elles enclenchent les forces qui peuvent mener à la démondialisation. La question de plus en plus vive des migrations internationales agit comme un révélateur des tensions du droit dans la mondialisation et d'un dilemme de l'État souverain, entre ouverture et fermeture. Dès lors, dans une économie ouverte, le territoire protège tout autant qu'il rend vulnérable. De plus, la mondialisation exacerbe une dynamique de creusement des inégalités économiques avec la compétition généralisée qui vient percuter l'idéal d'égalité sur lequel les sociétés démocratiques, comme la France notamment, sont bâties. La mondialisation génère des gagnants et des perdants, des acteurs économiques insérés dans les échanges internationaux et des exclus confrontés aux insécurités économiques et sociales, et les États sont aujourd'hui, comme par le passé, mis au défi de chercher un équilibre toujours précaire entre insertion dans les échanges et les réseaux internationaux et préservation du contrôle du territoire. Les enjeux de redistribution et de solidarité se construisent avant tout au niveau national des États souverains.

Si la mondialisation multiplie les richesses, les forces qu'elle déclenche bouleversent les équilibres nationaux, les hiérarchies sociales, les cultures, et créent des inégalités au sein de systèmes économiques nationaux placés sous la pression de l'extérieur : la recherche permanente d'avantages concurrentiels peut dès lors renforcer la tentation d'ériger des barrières protectionnistes, avec le risque d'exacerber les tensions nationalistes et populistes.

14.3.2 Vers la fin de la mondialisation ?

En se demandant dès 1997, si « *la mondialisation est allée trop loin* », Dani Rodrik, professeur à Harvard, avait amorcé une critique de la mondialisation actuelle. Il insistait en particulier sur les dangers d'une globalisation financière sans contraintes, parfaitement illustrés, selon lui, par la crise financière récente et les effets domino qui l'ont diffusée à travers le monde. Si cette crise était prévisible, les économistes ont été aveugles selon lui aux pièges pourtant évidents parce qu'ils croyaient à la fable qu'ils avaient eux-mêmes inventée : « *les marchés sont efficaces, l'innovation financière transfère les risques à ceux qui sont les plus à même de les supporter, l'autorégulation marche bien, et l'intervention des gouvernements est inefficace et nocive* ». Rodrik croit toujours dans la capacité de la mondialisation à sortir des millions de gens de la pauvreté et à étendre le bien être, mais seulement si elle est mise en œuvre de manière réfléchie par la prise en compte de ce qu'il nomme « *le paradoxe de la globalisation* ».



GRAPHIQUE 14.3.

Le « trilemme » de Dani Rodrik

Pour que la mondialisation procure ses bénéfices économiques réels et que ceux-ci soient largement partagés dans les sociétés, les démocraties nationales doivent être renforcées et des règles internationales doivent être mises en place pour protéger tous les participants, tout en maintenant de la souplesse pour les entreprises. La mondialisation pose

donc, selon Rodrick, le « *trilemme* » suivant : « *On ne peut simultanément poursuivre la démocratie, la détermination nationale des politiques et la mondialisation économique* ». Il considère en conséquence que l'un des trois objectifs doit être abandonné si les autres veulent être poursuivis. Contre le discours de renoncement qui a suivi la fin des accords de Bretton Woods et le rôle du « *consensus de Washington* » (entre FMI, OMC et Banque mondiale) dans sa diffusion, il affirme que le droit des démocraties à protéger leurs propres arrangements sociaux doit l'emporter sur les exigences de compétitivité dans l'économie globale. Il montre que les progrès de l'Inde et la Chine sont passés par une intégration contrôlée, avec de fortes interventions et des buts clairement définis pour la société. Conscient des difficultés d'une régulation mondiale, il n'en défend pas moins d'abord la nécessité de tenir un autre discours, aussi attentif à la justice et à la durabilité qu'à la croissance : « *Nous avons besoin d'une mondialisation intelligente, pas d'une mondialisation maximale* ».

Les thèses de Dani Rodrik entrent évidemment en forte résonance avec les exigences de régulation mondiale qui sont nées de la crise de 2007/2009, encore débattues aujourd'hui. De cet économiste aujourd'hui parmi les plus écoutés, le nouveau « *triangle d'incompatibilité* » qu'il met en évidence dessine trois voies possibles pour l'économie mondiale: *l'hyper-mondialisation* qui conduit à abandonner la démocratie puisque les choix des nations sont dictés de l'extérieur (agences de notation par exemple), la *gouvernance mondiale* qui mène à la disparition de l'État-nation (à l'instar d'une Europe fédérale) ou la *mondialisation modérée* dans laquelle celle-ci est mise au service des aspirations démocratiques des peuples avant de chercher l'efficacité économique et qui admet la résilience des États nations. Cette troisième partie de l'alternative se rapproche aujourd'hui des positions des tenants en France de la « *démondialisation* ». L'analyse de Rodrik conduit aussi à réexaminer les stratégies de croissance dans les pays en

développement, car subventionner ses biens échangeables ou protéger par des droits de douane les industries nationales peuvent donc être des voies encore valables, à l'instar des exemples chinois ou indien puisque, selon lui, « *marché et États sont complémentaires, pas des substituts* » et « *qu'il n'y a pas une seule voie vers la prospérité* ».

L'ÉCONOMIE OUVERTE : FIRMES MULTINATIONALES ET FRAGMENTATION DU PROCESSUS PRODUCTIF

« Je me range [...] aux côtés de ceux qui voudraient restreindre l'intrication des nations, plutôt que de ceux qui voudraient la voir s'étendre. Les idées, le savoir, la science, l'hospitalité, les voyages – telles sont les choses qui de par leur nature devraient être internationales. Mais faisons en sorte que ce qui peut être fait chez soi le soit autant qu'il est raisonnable et pratique de le faire, et, surtout, faisons en sorte que la finance soit essentiellement nationale. » - John-Maynard Keynes, 1933

« Le problème de l'emploi est bien réel en France et dans d'autres pays développés, mais il ne vient pas tant des délocalisations que du manque de créations d'emplois. » - Suzanne Berger, 2005

SOMMAIRE

[15.1 La mondialisation productive](#)

[15.2 La globalisation des chaînes de valeur](#)

[15.3 Les firmes globales à la recherche de la compétitivité dans l'économie globalisée](#)

[15.4 Les différentes dimensions de la compétitivité](#)

15.1 LA MONDIALISATION PRODUCTIVE

15.1.1 *La montée en puissance des firmes multinationales*

Le mouvement de mondialisation de l'économie s'est appuyé sur la montée en puissance des firmes multinationales (FMN). On peut définir les firmes multinationales comme des entreprises à base nationale qui possèdent au moins un établissement à l'étranger (une filiale) et qui produisent ou commercialisent hors de leur territoire d'origine grâce à leurs filiales. Ainsi, une production à l'étranger (et non pas seulement une distribution) constitue une caractéristique importante d'une firme multinationale : les organismes internationaux retiennent de plus en plus le terme de firme « transnationale » (FTN), puisque le terme « multinationale » pourrait laisser penser que la firme a plusieurs nationalités, alors que l'observation montre que ces firmes ont encore un ancrage national clairement identifié, où est implantée la société mère (les cadres et les dirigeants de la firme étant très majoritairement de la même nationalité).

Les firmes multinationales sont un des principaux agents de la mondialisation contemporaine et de la division internationale du travail (DIT) : elles réalisent aujourd'hui plus du quart du PIB mondial, sont directement responsables d'un tiers du commerce mondial et même des deux tiers, si l'on intègre leurs commandes indirectes et les activités de leurs réseaux de sous-traitance (soit le commerce intra-firme). Si la firme répartit ses activités sur une multiplicité de territoires, les activités les plus créatrices de valeur ajoutée, qui font l'avantage compétitif de la firme (recherche et développement, conception, design, construction de la réputation de la marque, stratégies financières) restent polarisées dans le pays d'origine, soit les pays avancés (Union européenne, États-Unis, Japon), même si l'on

constate une montée en puissance des firmes multinationales des pays émergents (chinoises, notamment). Les activités à plus faible valeur ajoutée, de fabrication, de montage et de distribution, intensives en main-d'œuvre peu qualifiée, sont alors situées dans les pays émergents (souvent chez des sous-traitants).

La mondialisation de la production a entraîné une profonde transformation des stratégies des firmes multinationales, apparues dès le XIX^e siècle (secteur pétrolier), mais surtout après 1945 avec les stratégies des grandes entreprises américaines. Au milieu des années 1960, les firmes multinationales étaient estimées à environ 7 000 dans le monde, contre 85 000 aujourd'hui. En 2017, les 100 plus puissantes d'entre elles dans l'économie mondiale représentaient environ 10% de la capitalisation boursière mondiale, 17% des ventes, 13% de la main-d'œuvre à l'étranger et environ 10% du PIB mondial, principalement dans les secteurs de la haute technologie, celui des mines/pétrole/raffinage, de l'automobile et de l'aéronautique, ou ceux des télécommunications et de l'agroalimentaire. Si la première forme de mondialisation a concerné les échanges de biens et services et la mobilité internationale des ressources nécessaires pour les produire (le travail lorsqu'il s'agit des migrations internationales et le capital lorsque l'on prend en compte les investissements internationaux), on a assisté plus récemment à une mondialisation de l'activité de transformation des ressources en produits, avec la sous-traitance internationale. Les multinationales qui se sont développées à partir des années 1970 avaient pour objectif de rationaliser le processus productif et de viser l'efficacité en divisant le travail de fabrication à l'échelle internationale, mais des firmes multinationales dites *globales*, intégrant des activités commerciales, financières et industrielles, sont apparues par la suite, marquant une nouvelle étape du processus de mondialisation de l'économie dans le cadre de vastes réseaux internationaux de production. L'objectif de ces firmes globales est d'assurer la flexibilité du processus

productif (minimisation des coûts, différenciation des produits), par le biais de réseaux de firmes interdépendantes, souvent réunies par des accords contractuels de partenariat et de coopération.

L'histoire de l'internationalisation des firmes depuis les années 1950 se caractérise par une succession de modèles, de types idéaux, fondés sur le *marché*, puis la *hiérarchie* et enfin la *coopération* :

- Durant les années 1950-1960, c'est le modèle de la firme exportatrice qui s'impose, avec des relations de marché (soit des flux d'exportations entre des entreprises indépendantes) ;
- Durant les années 1970, les firmes multinationales vont s'organiser de manière croissante dans le cadre d'une coordination hiérarchique par l'autorité au sein de groupes contrôlés par la maison mère ;
- À partir des années 1990, la coopération va progresser dans le cadre de réseaux complexes de firmes indépendantes reliées par des accords plus formels (on parle ainsi parfois de « firmes-réseaux »).

15.1.2 La croissance des IDE

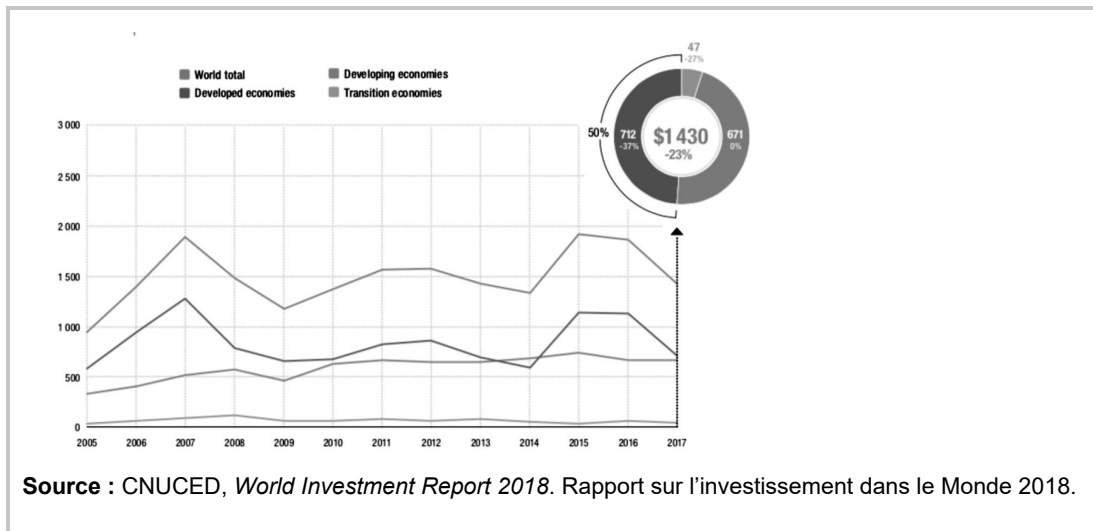
Selon la définition du Fonds monétaire international (FMI), les investissements directs étrangers (IDE), soit en anglais *Foreign Direct Investments* ou FDI, désignent des opérations financières des firmes multinationales destinées à contrôler ou à exercer une influence significative sur la marche et la gestion d'entreprises implantées dans un pays différent de celui des maisons mères. Il existe alors un intérêt durable et une relation à long terme entre l'investisseur direct et la société investie, ainsi que l'exercice d'une influence notable du premier sur la gestion de la seconde. Par convention, on considère qu'il y a intérêt durable et donc investissement direct lorsqu'une entreprise détient au moins 10% du capital ou des droits de vote d'une entreprise résidente d'un pays autre que le sien. Les investissements directs

comprennent non seulement l'opération initiale, qui établit la relation entre les deux unités, mais également toutes les opérations en capital ultérieures entre elles et entre les unités institutionnelles apparentées, qu'elles soient ou non constituées en sociétés.

Il faut soigneusement distinguer un IDE d'un investissement de portefeuille : l'investissement de portefeuille correspond à un ou des investissements réalisés sous forme de titres financiers. L'investissement de portefeuille dans un titre peut représenter un pourcentage substantiel du capital d'une entreprise, mais néanmoins limité, contrairement à un investissement direct, qui a pour objectif la prise de contrôle d'une entreprise. À l'heure actuelle, les firmes multinationales organisent en réseaux leurs sites de production et de commercialisation, qui sont souvent très éloignés dans l'espace. Leur stratégie tend à concevoir le monde comme un marché unifié, et celles-ci développent deux logiques d'investissement direct à l'étranger (IDE).

On peut considérer que l'évolution des IDE dits « entrants » constitue, avec celle des échanges internationaux de biens et services, l'un des principaux indicateurs de l'évolution de la mondialisation de l'économie : on constate en effet que les flux d'IDE sont très dépendants de la conjoncture internationale (accroissement des IDE en période de croissance, ralentissement des flux en période d'incertitudes dans l'économie mondiale voire reflux en période de crise). On a pu d'ailleurs constater une forte croissance des IDE jusqu'à la grande crise financière mondiale en 2007, grâce à la forte croissance des pays émergents, l'accroissement du chiffre d'affaires des firmes multinationales et les opérations de fusions-acquisitions. Après une forte chute des IDE entrants entre 2007 et 2009, ils ont connu une nette reprise avant d'enregistrer une baisse significative en 2017 en raison d'une croissance mondiale atone, mais aussi de facteurs plus structurels, comme la baisse du nombre de fusions-acquisitions dans le monde, et le ralentissement de la

globalisation des chaînes de valeur par les firmes multinationales.



GRAPHIQUE 15.1.

Entrées¹⁸ d'IDE, au niveau mondial et par groupe de pays, 2005-2017 (en milliards de dollars et en pourcentage)

En termes d'accueil des IDE, les pays développés, avec 712 milliards de dollars en 2017, soit 49,8% du total mondial, restent légèrement plus attractifs que les pays émergents (670 milliards en 2016, soit 46,9%). Viennent ensuite l'Europe du Sud Est et la CEI avec 46,7 milliards de dollars et 3,3% du total. En termes de classement des pays les plus attractifs, les États-Unis conservent la tête devant la Chine, tandis que l'Union européenne demeure attractive avec 46,9% du total des pays développés (23,3% du total mondial). En 2017, avec 49,7 milliards de dollars investis sur son sol, la France restait la deuxième destination des IDE en Europe derrière les Pays-Bas.

15.1.3 Les différentes stratégies des firmes multinationales

Ce mouvement de multinationalisation des firmes à partir des années 1980 s'est appuyé sur la déréglementation des flux

de capitaux, qui a favorisé l'essor des flux d'investissements directs étrangers (IDE).

Mais il est également lié aux **diverses stratégies¹⁹ des entreprises** :

- Les stratégies d'approvisionnement et d'accès aux matières premières face à la rareté des ressources ;
- Les stratégies de marché (*market seeking*) à la recherche de nouveaux débouchés, dont le but est de contourner des barrières protectionnistes en produisant directement sur le marché que l'on veut pénétrer, plutôt que d'y exporter (dans les secteurs comme l'automobile, la banque, l'assurance, les agences de publicité, etc.) ;
- Les stratégies de rationalisation à la recherche de moindres coûts des facteurs de production, travail, capital, terre, dans la concurrence internationale (*cost-seeking*), en implantant des filiales-ateliers dans les pays à bas salaires ;
- Enfin, les stratégies de recherche de nouvelles positions de marché, dans le cadre d'une concurrence oligopolistique fondée sur la recherche d'économies d'échelle et d'avantages concurrentiels (*efficiency seeking*).

La théorie traditionnelle de la firme multinationale cherche à traiter trois questions principales : on évoque à ce titre le **paradigme « OLI »**, dont les enseignements partent de deux questions clés. Premièrement, pourquoi, pour certaines firmes, est-il optimal, pour fonctionner, de s'établir dans plus d'un pays, alors que pour d'autres firmes ce n'est pas le cas ? Deuxièmement, qu'est-ce qui détermine le pays dans lequel les installations de production seront situées et celui dans lequel elles ne le seront pas ? Enfin, pourquoi les firmes installent-elles leurs propres installations à l'étranger plutôt que de simplement passer des contrats avec les producteurs et distributeurs locaux ?

Au fil des études, le paradigme « OLI » (Dunning, 1981) est devenu le cadre de référence de la réponse à ces questions

en posant trois conditions au développement à l'étranger :

- Le coût de l'implantation à l'étranger (adaptation à la langue, à la législation, avantage des entreprises autochtones sur les marchés, etc.) oblige d'abord les firmes multinationales à détenir un avantage absolu et spécifique sur leurs concurrentes pour que cette implantation soit profitable : détention d'un brevet (ou d'un secret technique), réputation (type qualité allemande) bien établie, etc. Cet avantage est l'*Ownership Advantage* (O), l'avantage spécifique.
- À ceci doit s'ajouter l'avantage d'une implantation locale (*Location Advantage*, L), qui justifie de ne pas exploiter l'avantage précédent en exportant à partir du pays d'origine. Cela peut venir de la suppression des coûts ou des temps de transports, du contournement des droits de douane, de l'utilisation d'*inputs* moins onéreux que dans le pays d'origine (travail) ou de l'avantage d'une législation moins contraignante (pollution, prévention des risques industriels).
- Enfin, il doit être plus avantageux pour la firme transnationale de faire plutôt que de faire faire parce qu'elle pourrait très bien se contenter d'un contrat avec une firme locale en cédant par exemple une licence d'exploitation. L'incomplétude du contrat, la volonté de maintenir le secret technologique sont alors les raisons qui font de l'intégration au sein de la firme un mode d'organisation supérieur au contrat : c'est l'*Internalization Advantage* (I).

Le choix de la multinationalisation se fera lorsque la firme réunit simultanément les trois avantages O, L et I.

On peut dès lors distinguer les firmes multinationales « classiques » des firmes multinationales « globales » :

- La FMN classique produit sur plusieurs marchés nationaux des biens adaptés à chaque marché, et chaque filiale entretient des relations verticales, hiérarchiques, avec la maison mère, mais non avec les autres filiales ;

- La FMN globale, quant à elle, développe un ensemble de stratégies qui cherchent à unifier la gamme de produits à l'échelle mondiale et à faire de chaque filiale une unité spécialisée dans la fabrication d'un composant particulier du produit fini.

15.2 LA GLOBALISATION DES CHAÎNES DE VALEUR

La libéralisation du commerce international, la réduction des coûts de transport et la convergence des modes de consommation à l'échelle internationale ont permis une nouvelle répartition des activités des firmes multinationales selon une conception mondialisée de leurs opérations et de leurs marchés. Les stratégies des firmes globales sont définies par la maison mère dont la direction cherche à unifier la gamme des produits sur le marché mondial, recourt directement aux marchés des capitaux mondialisés pour son financement, mais décide de la localisation de sa production en fonction des avantages compétitifs de chaque territoire. La globalisation des chaînes de valeur désigne alors la décomposition du processus de production d'un bien en plusieurs opérations prises en charge par des unités de production implantées dans différents pays. En s'appuyant sur la baisse des coûts de transport (avions, bateaux porte-conteneurs...) et de communication, mais aussi sur les entreprises de logistique, cette globalisation de la production a ainsi permis aux firmes multinationales de découper en différentes tranches leurs activités afin de maximiser la valeur ajoutée sur chacune d'entre elles. Depuis une dizaine d'années, les firmes globales ont conclu des contrats de sous-traitance avec des firmes indépendantes situées dans les pays émergents, qui bénéficient d'un réservoir de main-d'œuvre qualifiée à faible niveau de rémunération (comme l'américain Apple avec les entreprises taïwanaises Foxconn et Quanta).

Le but de ce type de stratégie est de limiter le montant des capitaux immobilisés pour les donneurs d'ordres, d'optimiser les coûts par la mise en concurrence des sous-traitants et d'augmenter l'utilisation de composants standardisés et l'échelle de fabrication, ce qui permet des baisses de prix.

Depuis les années 1980, les firmes multinationales ont ainsi mené des stratégies de *délocalisations* : les délocalisations peuvent être définies comme un transfert d'activités économiques du territoire national vers un pays étranger afin de réimporter sur le territoire national l'essentiel des biens produits ou de servir les mêmes marchés, via un investissement direct à l'étranger (IDE). Sont assimilés à des délocalisations (mais sans investissements) les accords de sous-traitance avec une entreprise étrangère (*offshore outsourcing*) ou l'octroi d'une licence à cette entreprise quand cela a eu pour effet de substituer une production étrangère à une production nationale.

Les firmes multinationales franchissent ainsi toutes les étapes pour produire un produit ou un service, et le livrer au client, depuis la conception jusqu'à l'utilisation finale, dans le cadre d'une stratégie d'emblée conçue à l'échelle mondiale. Ces étapes comprennent des activités comme la recherche et la conception, la production, le marketing, la distribution et l'appui au consommateur final. Ce phénomène est étroitement lié au développement des réseaux mondiaux de production : il se traduit donc par une fragmentation physique du processus de production, dont les différentes étapes sont réalisées en des lieux distincts, suivant une logique d'optimisation, les entreprises ayant intérêt à se tourner davantage vers des sources d'approvisionnement internationales pour leurs consommations intermédiaires en fonction des coûts comparés.

Les frontières de la grande firme ont alors tendance à se brouiller, puisqu'elles sont constituées aujourd'hui à la fois de filiales propres (liées à la maison mère par des liens de propriété des actifs) et de fournisseurs indépendants liés par contrat au groupe. Quelques milliers de firmes globales ont

cette capacité de mettre en concurrence les marchés et les avantages comparatifs des différents pays ; elles localisent leurs activités là où la rentabilité est la plus forte, grâce à l'usage des nouvelles technologies et de la production « flexible » (on parle ainsi parfois de « firmes-réseaux »), et elles organisent en réseau leurs filiales étrangères, qui peuvent être elles-mêmes liées avec d'autres sociétés mères multinationales (téléphonie mobile, équipement informatique, automobile, aéronautique civile). La mondialisation des firmes s'appuie ainsi non seulement sur le processus d'externalisation (sous-traitance, *outsourcing*), mais aussi sur les réseaux, les stratégies d'alliances et les accords de coopération entre firmes.

Cette fragmentation des chaînes de valeur globales de plus en plus sophistiquées entraîne une polarisation toujours plus forte des espaces productifs au sein de l'économie mondiale, en raison des effets d'agglomération et de la concentration des activités productives (pour réduire les coûts, gagner en flexibilité, faire circuler l'information plus facilement dans le cadre de districts industriels, et profiter d'un cadre de vie agréable adapté aux cadres très qualifiés des classes moyennes supérieures), et d'un clivage substantiel et grandissant entre les grands centres urbains favorisés et les périphéries des villes.

15.3 LES FIRMES GLOBALES À LA RECHERCHE DE LA COMPÉTITIVITÉ DANS L'ÉCONOMIE GLOBALISÉE

15.3.1 *La mise en concurrence des territoires par les firmes*

Avec la mondialisation économique, l'inscription des firmes industrielles et financières dans un cadre territorial bien défini, l'État-nation, s'estompe au bénéfice d'une mise en

concurrence (un *benchmarking*) des territoires : ceux-ci offrent des dotations diverses en matière de coût du travail, de qualification de la main-d'œuvre et de ressources naturelles. Les stratégies de localisation des firmes globales, « acteurs nomades » dans l'économie mondiale pour reprendre la formule de l'économiste Pierre-Noël Giraud, deviennent plus élaborées et la délocalisation des activités productives s'intensifie en direction notamment des pays à bas salaires et à capacité technologique.

Par ailleurs, la mondialisation financière, en créant un marché unique de l'argent à l'échelle mondiale, soumet les États nations à de nouveaux impératifs, comme celui de créer un environnement favorable pour les firmes multinationales et les opérateurs financiers, en quête d'investissements et de placements profitables.

Avant de s'implanter, la firme effectue un diagnostic précis de son environnement externe en termes d'opportunités et de menaces, en distinguant le microenvironnement et le macro-environnement :

- Le macro-environnement est l'ensemble des grandes tendances subies par l'entreprise et qu'elle ne peut influencer, sur le plan politique (politiques fiscales, stabilité politique, réglementation de la concurrence, droit du travail, etc.), économique (croissance, niveau de revenu, coût du crédit et taux d'intérêt, confiance des consommateurs), social (niveau d'éducation, esprit d'entreprise, conditions de vie) et technologique (investissements publics et privés sur la technologie, découvertes et inventions, etc.) ;
- Le microenvironnement regroupe les cinq forces de la concurrence décrites par M. Porter : la menace de nouveaux concurrents, l'intensité de la concurrence dans le secteur, le pouvoir de négociation des fournisseurs, le pouvoir de négociation des clients et la menace de produits substitués sur le marché.

Ainsi, leurs stratégies et leur puissance financière permettent aux firmes multinationales de modeler, voire de créer la spécialisation productive d'un territoire. Dans cette configuration, les « acteurs nomades », pour reprendre la typologie de l'économiste Pierre-Noël Giraud, comme les firmes globales et les travailleurs très qualifiés, disposent d'un avantage grandissant sur les « acteurs sédentaires », les travailleurs non qualifiés, au sein des pays riches. Dans ce nouveau cadre de mobilité des biens, des capitaux et des services, les États cherchent alors à renforcer leurs avantages compétitifs et à capter les flux d'investissements mondiaux par des politiques d'attractivité de leurs territoires (dynamisme de la demande sur le marché, faible niveau de la fiscalité, qualification élevée de la main-d'œuvre, cadre réglementaire favorable aux affaires, administration efficace, infrastructures publiques adaptées) et des politiques macroéconomiques favorables à la rentabilité des capitaux (taux d'intérêt suffisamment rémunérateurs, stabilité des prix, faible taxation du capital). Les États cherchent à mettre en place des politiques de compétitivité, censées améliorer l'insertion internationale du pays et sa spécialisation productive, notamment dans le cadre des politiques commerciales stratégiques. Avec la mondialisation des systèmes productifs, les économies nationales sont de plus en plus déterritorialisées et les grandes firmes, à la recherche de l'optimisation et de l'efficacité, sont en mesure de comparer les avantages comparatifs des différents territoires (et leurs dotations factorielles). Dans la mesure où les investissements des firmes s'accompagnent de création d'emplois, les États, en particulier ceux qui connaissent des problèmes de chômage importants, sont parfois prêts à faire d'importantes concessions pour les attirer sur leur territoire (baisse de la fiscalité notamment).

15.3.2 *Les déterminants de la compétitivité des firmes*

La compétitivité désigne la capacité d'une entreprise ou d'un pays à affronter avec succès la concurrence internationale et à augmenter ses parts de marché.

- Sur un marché où plusieurs producteurs offrent des biens ou services substituables, les économistes parlent de *compétitivité prix* pour désigner la capacité à proposer des produits à un prix inférieur à celui des concurrents.

Elle résulte :

- Des coûts de production à savoir le coût du capital (les taux d'intérêt), le coût du travail (les salaires), le coût des consommations intermédiaires (énergie, matières premières, composants et équipements achetés) ;
- De la productivité des facteurs de production (efficacité du travail ou du capital) ;
- Des coûts de transport ;
- De l'évolution du taux de change ;

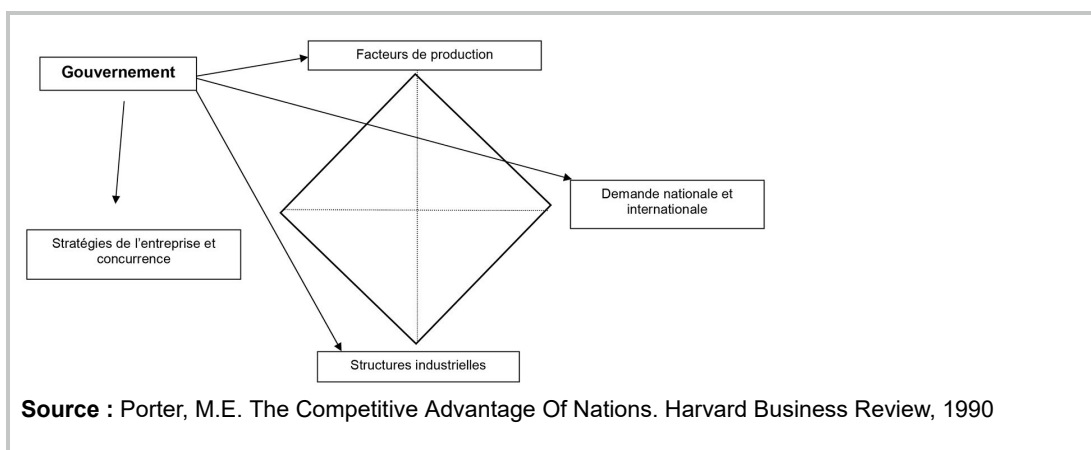
La compétitivité prix est donc principalement fonction de l'évolution des coûts, qui sont liés notamment aux écarts en termes de coûts salariaux unitaires (CSU) entre les économies nationales. Le coût salarial unitaire correspond au coût salarial par unité de valeur ajoutée produite. Il se mesure par le rapport entre le coût total de la main-d'œuvre et la production en volume. Dès lors, lorsque la productivité horaire du travail progresse plus vite que le coût salarial horaire, le CSU baisse. Ce type d'indicateur est très important à mobiliser dans les comparaisons internationales de coût salarial, car il intègre les écarts de productivité par tête. Pourtant, ces indicateurs ne nous donnent qu'une mesure imparfaite de la compétitivité, car les CSU ne sont qu'une partie des coûts de production : il y a aussi les consommations intermédiaires, les marges de production, les marges de distribution, les marges d'exportation.

- La *compétitivité hors prix* renvoie quant à elle aux explications structurelles des performances en termes de qualité des produits, de contenu technologique, de conditionnement, etc. La différenciation des produits en est

généralement à l'origine (les voitures de marques allemandes ont la réputation d'être plus fiables que les voitures françaises, les voitures italiennes d'avoir un design plus élaboré). Une entreprise compétitive présentera une croissance à long terme de sa production supérieure à celle de ses concurrents grâce à un avantage de prix, de qualité, ou de coût.

La compétitivité est donc le produit du croisement de nombreuses variables macroéconomiques, microéconomiques et institutionnelles, ces dernières incluant le rôle du gouvernement, le cadre réglementaire et les institutions de la croissance. Cette relation a été formalisée par Michael Porter, de la *Harvard Business School*, qui propose un graphique en diamant qui reprend quatre éléments de la compétitivité :

- Les conditions des facteurs de production (ressources naturelles, technologie, capital humain) ;
- Les conditions de la demande nationale et internationale ;
- Les structures industrielles et l'efficacité des réseaux ;
- La stratégie de l'entreprise, l'organisation institutionnelle et la concurrence pour investir dans l'innovation.



GRAPHIQUE 15.2.

Facteurs de la compétitivité : schéma en « diamant » de Michael Porter

La cohérence des relations entre ces quatre éléments, conjuguée au rôle du gouvernement, va déterminer la compétitivité internationale globale du pays. Dans la mondialisation, le gouvernement joue donc un rôle important en fournissant un environnement incitatif (stabilité politique et cadre réglementaire). Mais, selon Porter, ce sont bien les entreprises qui doivent *in fine* affronter la concurrence internationale sur des marchés souvent éloignés, volatils et très compétitifs, et ce sont leurs forces internes de concurrence qui font la différence pour gagner des contrats et des parts de marché. Le schéma de Michael Porter met en exergue l'importance des relations dynamiques entre ces facteurs qui se renforcent mutuellement. Cependant, ce schéma demeure limité dans son pouvoir éclairant : il laisse dans l'ombre le rôle de la gouvernance de l'entreprise et n'intègre pas la stratégie de financement de l'entreprise.

Les États nations deviennent aujourd'hui des artisans de la globalisation économique, au sens où la concurrence entre eux et la quête de la compétitivité déterminent leurs politiques d'attractivité (ce ne sont pas les investissements qui sont attractifs mais les pays ou les territoires), avec la mise en place d'incitations fiscales et financières, le recours aux agences de promotion, ou les fournitures de services et d'infrastructures. Les entreprises installées sur un territoire, petites ou grandes, doivent s'adapter à la mondialisation et peuvent alors suivre différentes stratégies de production et de commercialisation, selon qu'elles cherchent à conquérir le marché domestique ou qu'elles ont l'ambition d'occuper une position mondiale. Les firmes globales essaient de mobiliser et de combiner des compétences (conseil, ingénierie, finance, marketing, publicité) disponibles sur ces territoires dans le but de réaliser des projets complexes, en particulier dans le domaine des nouvelles technologies, qui leur permettent de dégager une rente (temporaire) de monopole sur le marché. De fait, elles sont amenées à attirer ou repousser d'autres firmes, créant plus ou moins « d'effets d'agglomération » sur un territoire, c'est-à-dire une concentration d'activités proches

sur un espace géographique donné qui participe à la spécialisation du pays. Cette agglomération des entreprises dans un même espace s'effectue pour bénéficier des externalités positives liées à la proximité géographique : marché du travail local commun, marchés de capitaux, infrastructures publiques, réseaux de communications, retombées technologiques, gains d'information, etc.

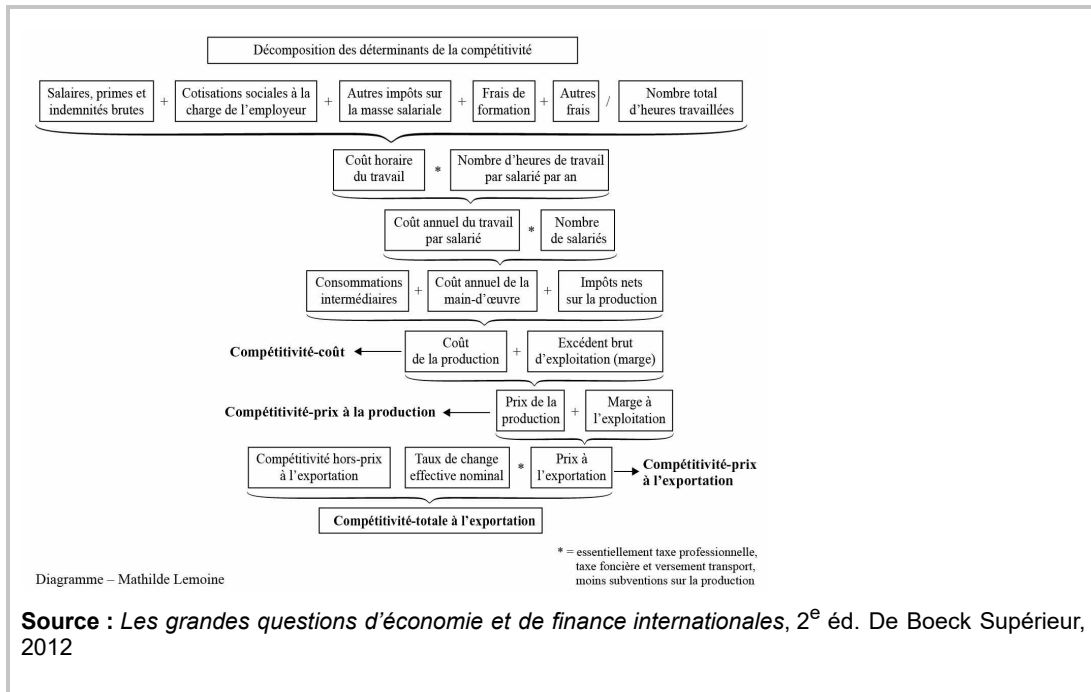
La notion de compétitivité englobe une pluralité de variables économiques, et il est difficile de l'évaluer par un indicateur unique. Cependant, de nombreux travaux tentent de la mesurer. Le *World Economic Forum* publie ainsi chaque année un rapport intitulé *Global Competitiveness Report* dans lequel il dresse un classement des pays les plus compétitifs. Ce rapport annuel se base sur plus de cent indicateurs organisés en douze « *piliers* » : les institutions, les infrastructures, l'environnement macroéconomique, la santé et l'enseignement primaire, la taille du marché, le développement technologique, l'innovation, l'efficacité du marché des biens, l'efficacité du marché du travail, l'enseignement supérieur et la formation professionnelle, l'évolution des marchés financiers, la sophistication des activités commerciales et l'innovation.

La notion de compétitivité fait également l'objet d'un débat macroéconomique qui est abordé dans le chapitre 17.

15.4 LES DIFFÉRENTES DIMENSIONS DE LA COMPÉTITIVITÉ

La notion de compétitivité est un enjeu capital dans la mondialisation. Pourtant, il n'est pas simple de la cerner sans verser dans le simplisme et donc dans l'erreur. Dans *Les grandes questions d'économie et de finance internationales*, De Boeck Supérieur 2012, Mathilde Lemoine, Philippe Madiès et Thierry Madiès précisent ainsi que « la compétitivité est une notion difficile à définir ».

15.4.1 La compétitivité relève d'abord d'une approche microéconomique.



GRAPHIQUE 15.3.

Les multiples facteurs de la compétitivité des firmes A redessiner.

En effet, dans le même ouvrage, les auteurs précisent la distinction entre compétitivité prix et compétitivité coût. La compétitivité prix à l'exportation est fonction du taux de change effectif nominal, du taux de change effectif réel, des prix relatifs des exportations, des écarts aux taux de change d'équilibre, en parité de pouvoir d'achat (PPA). La compétitivité coût est fonction de l'évolution des CSU (coûts salariaux unitaires) entre les économies. Pourtant, ces indicateurs ne nous donnent qu'une mesure imparfaite de la compétitivité, car les CSU ne sont qu'une partie des coûts de production. Il y a aussi les consommations intermédiaires, les marges de production, les marges de distribution, les marges d'exportation. La compétitivité hors prix dépend de l'innovation technologique, du design, du marketing... Donc, au niveau microéconomique, nous observons que le concept

de compétitivité recouvre déjà un nombre important de dimensions !

15.4.2 *L'approche macroéconomique de la compétitivité*

L'approche macroéconomique de la compétitivité est à l'origine d'ambiguïtés et de confusions encore plus nombreuses. La définition de la compétitivité au niveau macroéconomique pose un vrai problème aux économistes. Dans un rapport du Conseil d'Analyse économique (CAE) intitulé « Compétitivité », 2003, Michel Debonneuil et Lionel Fontagné pointaient dès la première phrase les ambiguïtés du terme compétitivité dès lors qu'on l'applique à une économie nationale. Ils ajoutent que pour Paul Krugman (prix Nobel en 2008 pour ses travaux en économie internationale), « *la compétitivité est un mot vide de sens lorsqu'il est appliqué aux économies nationales* » et « *qu'un problème de compétitivité d'une économie nationale pouvait être purement et simplement un problème intérieur de productivité.* » Pourtant, cette approche macroéconomique, globale, de la compétitivité marque les travaux et les décisions politiques internationales, en particulier en Europe. Ainsi en était-il de la Stratégie de Lisbonne adoptée en 2000 qui prévoyait qu'en dix ans, l'économie européenne puisse « devenir l'économie de la connaissance la plus compétitive et la plus dynamique du monde, capable d'une croissance économique durable accompagnée d'une amélioration quantitative et qualitative de l'emploi et d'une plus grande cohésion sociale », et « dans le respect de l'environnement » (ajouté en 2001). Enfin, le Conseil européen du 14 janvier 2003 explique que « la compétitivité d'une nation est la capacité à améliorer durablement le niveau de vie de ses habitants et à leur procurer un haut niveau d'emploi et de cohésion sociale. »

Une place centrale est donc attribuée à cette forme élargie de la compétitivité pour favoriser le développement de nos économies nationales et de l'économie européenne

considérée comme un tout. D'une certaine manière, et il suffirait d'avoir pour objectif cette compétitivité globale pour avoir la croissance soutenue et durable, le plein emploi, la cohésion sociale... ! Mais, dans ces conditions, conformément au principe « qui trop embrasse, mal étreint » : cette approche trop globale de la compétitivité n'est ni pertinente ni propice pour atteindre les multiples objectifs visés. Un seul concept (et indicateur) associé pour expliquer la croissance économique, d'emploi, de cohésion sociale, de politique d'éducation et de recherche...

La compétitivité d'une économie ne peut donc être mesurée seulement à l'aune des déséquilibres extérieurs. Paul Krugman, récompensé par le prix Nobel d'économie en 2008 pour ses travaux en économie internationale, insiste depuis longtemps sur les dangers du concept de compétitivité appliqué à une économie nationale. Dans son manuel *Macroéconomie* (De Boeck Supérieur, 2008), il nous rappelle qu'un pays ayant un déficit commercial n'est pas nécessairement en mauvaise santé et qu'un déficit commercial (ou courant) n'est pas forcément le signe de dysfonctionnement de l'économie. « *Les déficits et excédents commerciaux sont des phénomènes macroéconomiques. Ils sont donc le résultat de situations dans lesquelles le tout est différent de la somme des parties. Il serait tentant de penser que les pays ayant des travailleurs très productifs ou des produits et services largement demandés devraient être en excédent commercial, et que les pays ayant des travailleurs improductifs ou des produits et services de faible qualité devraient être en déficit commercial. Mais la réalité est qu'il n'y a pas de relation simple entre la réussite d'une économie et sa situation d'excédent ou de déficit commercial* ».

Nous vous donnons également rendez-vous au chapitre 17 du livre sur cette question des déséquilibres macroéconomiques. En résumé, prétendre que la compétitivité d'une économie s'est dégradée parce que son compte courant (ou seulement sa balance commerciale) devient de plus en plus déficitaire est au moins partiellement

erroné ! Pourtant l'analyse des soldes de la balance commerciale et de la balance des transactions courantes est souvent menée exclusivement en termes de compétitivité globale !

15.4.3 Productivité, compétitivité et attractivité

Il convient encore de ne pas confondre les notions proches de productivité, compétitivité et attractivité. Selon Matthieu PLANE, économiste de l'OFCE, dans l'ouvrage *L'Économie française 2013*, « la compétitivité d'un pays est la capacité du secteur productif à répondre à la demande intérieure et étrangère tout en offrant aux résidents un niveau de vie qui, à la fois, s'élève et puisse être préservé à long terme. L'appréciation de la compétitivité d'un pays fait appel à deux approches complémentaires : les indicateurs de performance, d'un côté, et leurs déterminants, de l'autre. La première approche axée sur les résultats est une approche ex post de la compétitivité. L'indicateur le plus pertinent est les parts de marché des exportations dans le commerce mondial et le PIB par tête en parité de pouvoir d'achat. La seconde approche (compétitivité ex ante), qui s'intéresse aux déterminants internes de la croissance, regroupe des éléments touchant à l'attractivité... ». (...) « Selon la Direction de la prévision, l'attractivité peut être définie comme la capacité à attirer des capacités nouvelles et des facteurs de productions mobiles – capitaux, travailleurs qualifiés – sur le territoire afin d'améliorer la compétitivité et d'augmenter le niveau de vie de ses habitants. »

Enfin, malgré ces quelques éclaircissements sur la nature multidimensionnelle de la compétitivité, il est difficile d'ignorer quelques paradoxes inhérents à une approche de la croissance économique par la compétitivité ! Il est en effet paradoxal de se focaliser sur la compétitivité relative des économies ouvertes dans la mesure où théoriquement et empiriquement, nous apprenons dans tous les manuels d'économie que l'échange international est surtout un jeu à somme positive (gagnant-gagnant) alors que la quête de

compétitivité (des uns par rapport aux autres) nous conduit à un jeu à somme nul, où ce que gagnent les uns est perdu par les autres. Puisque le Monde est une économie fermée, les excédents extérieurs des uns s'accompagnent des déficits extérieurs des autres. La compétitivité des uns a donc pour contrepartie le manque de compétitivité des autres.

Nous disposons d'indicateurs de compétitivité. Mais, faisons au moins l'effort de les utiliser avec pertinence et cohérence. Pour le dire autrement, voici une nouvelle preuve qu'en économie, il ne suffit pas d'avoir des outils (indicateurs), encore faut-il savoir s'en servir !

18 Les entrées d'IDE constituent des importations de capitaux. Les sorties d'IDE correspondent à des exportations de capitaux. Les stocks sont constitués de la somme des flux passés. Les flux correspondent à des variations de stocks.

19 Dans son énumération des stratégies, Michalet mentionne les trois premières : approvisionnement, marché, stratégie productive (dont la recherche de la main d'œuvre bon marché constitue l'une des possibilités), mais il peut également s'agir de la recherche d'une main d'œuvre qualifiée, d'un environnement productif favorable, infrastructures, etc...

En revanche, il ne cite pas la 4^e (recherches de nouvelles positions de marché) qui ressemble à un sous-ensemble de la seconde.

Pour la stratégie de marché, contourner les barrières peut constituer l'une des motivations, mais pas la seule. Cela peut être également s'implanter dans un pays où connaître la clientèle (goût, culture...) ou bien pour se faire connaître. Autrement, il peut y avoir une stratégie de marché, même sans barrières douanières

L'ÉCONOMIE OUVERTE : L'INTÉGRATION ÉCONOMIQUE EUROPÉENNE

« Nous ne pouvons avoir simultanément 1) le libre-échange, 2) la mobilité du capital 3) l'indépendance de la politique monétaire domestique, et 4) des changes fixes. Le cercle ne peut être bouclé : un des éléments doit être abandonné pour éviter l'incohérence. » - Tommaso Padoa Schioppa, 1982

« De fait, le gouvernement de l'Europe ressemble plus souvent à un gouvernement par des règles qu'à un gouvernement par des choix. » - Jean-Paul Fitoussi, 2002

« La monnaie unique ne peut survivre sans un véritable fédéralisme économique qui suppose à la fois une solidarité financière, une responsabilité budgétaire collective, une harmonisation fiscale et une coordination étroite des instruments de politique économique compatibles avec le retour de la croissance dans la région et la réindustrialisation du continent. » - Patrick Artus, 2011

SOMMAIRE

[16.1 La dynamique de l'intégration économique européenne](#)

[16.2 La marche vers l'Union économique et monétaire \(UEM\)](#)

[16.3 Le système de politique économique de la zone euro](#)

[16.4 Les évolutions institutionnelles de la zone euro](#)

16.1 LA DYNAMIQUE DE L'INTÉGRATION ÉCONOMIQUE EUROPÉENNE

16.1.1 *Les degrés d'intégration économique*

L'intégration économique désigne le processus par lequel plusieurs nations s'accordent pour supprimer les barrières aux échanges entre elles, afin d'unifier progressivement leurs marchés, et d'en tirer des avantages mutuels. Il existe différents degrés d'intégration, qui conduisent à la formation d'espaces économiques constitués de plusieurs nations, liées par un traité attribuant aux membres de la zone des droits qui ne sont pas accordés au reste du monde (comme la suppression des tarifs douaniers). Ces espaces sont appelés unions économiques régionales, ou zones d'intégration régionales.

Dès l'origine, dans les années 1950, la construction de l'Europe fait le choix de l'intégration par le marché, en vertu de la formule de Robert Schuman selon lequel la construction européenne, garante de la Paix sur le continent, devait avant tout progresser par « *des réalisations concrètes basées sur des solidarités de fait* », c'est-à-dire des mesures concrètes dans le domaine économique notamment. L'achèvement de l'Union économique et monétaire (UEM) et la création de l'euro peuvent donc apparaître comme un prolongement naturel du processus d'intégration commerciale qu'a connu l'Union européenne dès l'origine. Pourtant, l'Union européenne demeure une construction politique originale, qui mêle les politiques communautaires et les décisions intergouvernementales : à mi-chemin entre la fédération et la confédération, la construction européenne a établi un système de politique économique qui combine les institutions supranationales (comme la Banque Centrale européenne en charge de la politique monétaire européenne) et les États

membres qui conservent une souveraineté encadrée par des normes (comme dans le domaine de la politique budgétaire).

On distingue différents degrés dans le processus d'intégration économique, selon l'importance de l'unification des marchés, et selon la nature des accords entre les pays de la zone. Selon la « *Théorie de l'intégration économique* » (1961) de l'économiste B. Balassa, il existe cinq degrés d'intégration.

- La zone de libre-échange se caractérise par une diminution ou une suppression des barrières douanières à l'intérieur de la région. Ce type d'accord laisse libre chaque membre de sa politique commerciale envers les pays extérieurs à la zone : le libre-échange reste intra-régional, ce qui préserve l'autonomie des États dans leurs politiques commerciales nationales. L'ALENA, zone de libre-échange entre les États-Unis, le Mexique et le Canada créée en 1992, en fournit un exemple. Par nature, la création d'une zone de libre-échange entraîne un risque de contournement par les importateurs des tarifs douaniers : pour importer dans le pays B où le tarif douanier extérieur à la zone est élevé, il suffit de faire transiter les marchandises par un pays A membre de la zone où le tarif douanier est plus faible. Afin d'éviter ce contournement de la politique commerciale, et les déséquilibres qu'il génère dans leur balance commerciale, les pays membres instituent généralement des règles d'origine : c'est la nationalité d'origine d'un produit qui fixe son niveau de taxation, pas le pays par lequel il a transité avant de pénétrer sur le territoire de la nation importatrice.

- L'union douanière est une zone de libre-échange dont les membres décident d'adopter une politique commerciale unique vis-à-vis du reste du monde, en fixant des tarifs douaniers extérieurs communs. Généralement, le passage à une union douanière nécessite de fixer des règles de partage des recettes douanières. Le Mercosur est organisé sur ce mode depuis sa création en 1991. On trouve des exemples d'unions douanières dès le XIX^e siècle, comme le Zollverein, créé entre les États allemands en 1834.

– Dans un marché commun, les pays membres ajoutent à l'union douanière la libre circulation des facteurs de production (capital, travail, brevets...). Cela suppose une harmonisation poussée des réglementations nationales, par exemple la fixation de règles communes concernant les diplômes et l'accès à des professions protégées par leur statut. Après le Traité de Rome de 1957, l'usage était de parler de « marché commun » pour désigner ce qui constitue à cette époque une union douanière, au sens de B. Balassa. En revanche, le « marché unique européen » entré en vigueur en 1993 constitue bien un marché commun, en raison de la libre circulation des marchandises et des facteurs de production.

– L'union économique peut se définir par l'adoption d'objectifs de politique économique communs, ce qui conduit à une harmonisation progressive des politiques économiques dans la zone. Pour B. Balassa, cela inclut une politique monétaire commune. Certains économistes distinguent l'union économique, qui entraîne uniquement l'harmonisation des politiques économiques, et l'union monétaire qui implique en outre une politique monétaire commune. Ainsi, l'Union européenne s'est engagée sur la voie de l'union économique et monétaire, définie en 1992 par le traité de Maastricht.

– La construction d'une véritable fédération d'États constitue le degré ultime d'intégration économique, ce que B. Balassa appelle « l'intégration totale » : à l'union économique et monétaire s'ajoute la constitution d'un pouvoir politique fédéral, avec l'harmonisation des politiques fiscales et sociales.

Toutefois, l'intégration européenne ne répond pas uniquement à des objectifs économiques (favoriser la croissance au niveau régional), mais aussi à des objectifs politiques (garantir une paix durable en Europe par l'intégration des nations, avec des institutions communes). L'intégration européenne n'est pas un processus uniforme et linéaire, qui conduirait l'Europe à suivre dans l'ordre les différentes étapes

de l'intégration économique distinguées par B. Balassa. C'est au contraire un processus complexe, où les différentes étapes s'enchevêtrent. Par exemple, l'intégration monétaire débute très précocement, avec l'Union européenne de paiements qui, entre 1950 et 1958, dans un contexte de pénurie de devises, organise une coopération monétaire en Europe par un mécanisme de compensation des créances et des dettes. De même, il existe des éléments de politique économique commune dès la signature du Traité de Rome en 1957, comme la politique agricole, alors que la libre circulation des marchandises et des facteurs de production est pleinement réalisée avec la création du « marché unique européen » en 1993.

L'intégration régionale (IR) ne doit donc pas être confondue avec la tendance naturelle des échanges à s'intensifier autour de pôles régionaux. Ce phénomène, qu'on appelle régionalisation, est une fonction inverse de deux paramètres essentiels : la distance économique et la distance géographique (cf. modèles de gravité de Paul Krugman, 1991).

La mondialisation, en exacerbant les exigences de compétitivité, de crédibilité et de pouvoir de négociation, incite les pays à se regrouper en zone régionale. Et en retour, l'IR, loin de freiner le processus de mondialisation, apparaît comme une étape nécessaire vers l'intégration globale.

16.1.2 Le processus d'intégration économique de l'Europe

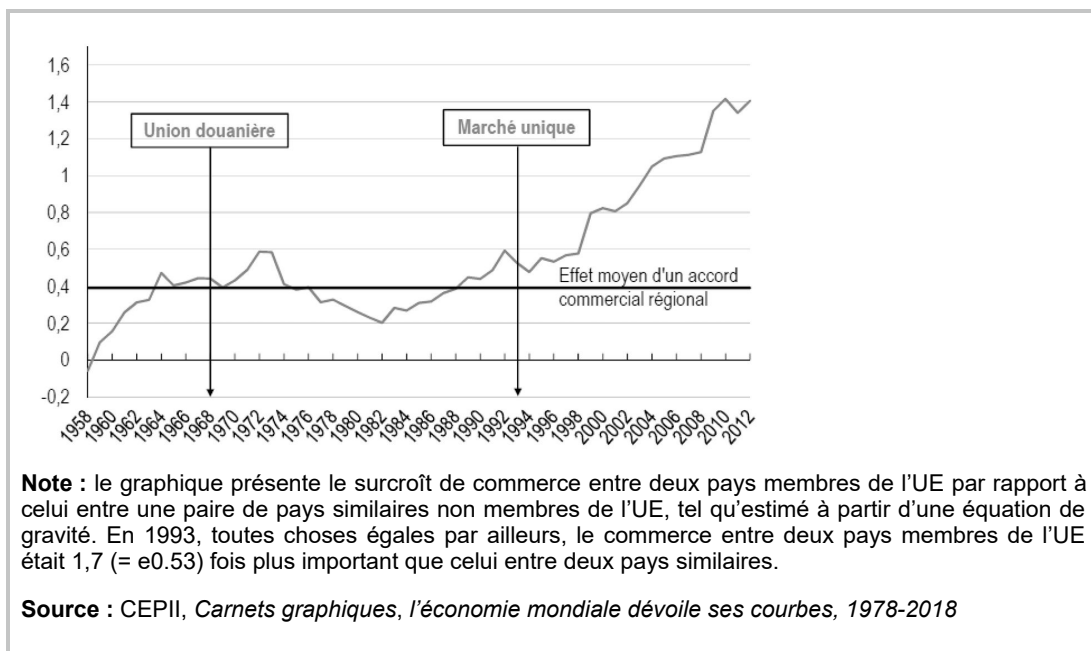
L'unification de l'Europe s'inscrit dans un large processus d'intégration économique et commerciale, résultant d'accords entre les nations en vue de constituer un ensemble aux réglementations, aux institutions, voire aux mécanismes économiques définis sur une base commune. Dans l'esprit des pères fondateurs de l'Europe, le progrès économique, dans le cadre d'une économie de marché fondée sur la concurrence libre et non faussée, est le moteur du progrès social. En effet, l'intégration européenne s'appuie d'emblée sur le mécanisme de croissance décrit jadis par A. Smith dès le XVIII^e siècle : l'élargissement de la taille du marché et l'intensification de la concurrence entre les entreprises doivent stimuler la recherche des gains de productivité, gages d'une croissance des niveaux de vie moyens. La croissance de la production qui en résulte enclenche alors un cercle

vertueux générateur de distribution de revenus, d'économies d'échelle, de baisses de prix et de gains de pouvoir d'achat pour les consommateurs. La concurrence et la spécialisation internationale sur le marché unique européen sont censées transformer la construction européenne en jeu à somme positive, puisque le gain mutuel à l'échange crée une dynamique dont tous les États membres profitent.

Décrits dans le rapport Cecchini dès 1988, ces bienfaits attendus du marché unique justifient ainsi l'élargissement progressif de l'Union européenne, accentuant encore la dynamique de compétitivité et d'efficacité des appareils productifs. Les revenus réels des ménages augmentent, tandis que les entreprises sont plus compétitives sur les marchés extérieurs, ce qui permet une amélioration de la balance commerciale. La croissance du PIB permet d'augmenter les recettes fiscales et les créations d'emploi. La conception allemande de l'économie sociale de marché soutient cette logique fondamentalement libérale : la liberté des échanges et des prix sur les marchés est le meilleur moyen d'atteindre l'optimum économique, et l'État peut contenir le creusement des inégalités les plus criantes qui en résultent par de modestes mécanismes sociaux.

Mais l'intégration économique européenne est aussi passée par la création de dispositifs institutionnels permettant de garantir une certaine stabilité monétaire et échapper aux turbulences monétaires liées à la fin du système de « Bretton Woods » fondé sur la convertibilité or/dollar. Au début des années 1970, les relations monétaires internationales entrent dans une phase d'instabilité : la suspension de la convertibilité-or du dollar (pour les banques centrales étrangères) en 1971 fait entrer le système monétaire international dans une période de désordres, contre lesquels les pays européens vont essayer de lutter. Ils décident d'instaurer le Serpent monétaire le 21 mars 1972. Il s'agit de diminuer de moitié les marges de fluctuations entre monnaies européennes (+/- 2,25%), en contenant leurs fluctuations à l'intérieur des bandes de fluctuations prévues vis-à-vis du

dollar (ce que l'on a appelé le « serpent dans le tunnel »). La décision est complétée par la création du FECOM, un fonds de coopération monétaire pour donner aux membres les moyens de leur politique de change. En effet, comme les taux de change sont définis par rapport au dollar, c'est seulement au pays dont la monnaie sort du tunnel qu'il incombe d'intervenir, ce qui réduit la dimension coopérative du serpent. Le serpent monétaire européen va toutefois échouer et le flottement du dollar après 1973 fait disparaître le « tunnel » tandis que les pays à monnaie faible renoncent à payer le prix de la défense de leur taux de change (taux d'intérêt élevés, perte des réserves de change). La livre quitte le serpent dès 1972, la lire en 1973, le franc quant à lui fera un « aller-retour » pour sortir définitivement en 1976. Seule une petite zone mark subsiste alors, mais elle disparaîtra en 1978. Malgré cet échec, l'organisation des taux de change européens apparaît toujours nécessaire. Cette nécessité est exprimée par Valéry Giscard d'Estaing et Helmut Schmidt lors du sommet franco-allemand de mars 1978. Les sommets de Copenhague et de Brême qui le suivent fixent les grandes lignes du nouveau système qui sera acté en décembre au cours du Sommet de Bruxelles.



GRAPHIQUE 16.1.

Effet de l'intégration européenne sur le commerce intra-européen

Le système monétaire européen (SME) entre en vigueur le 13 mars 1979 et son fonctionnement repose sur l'écu et la définition de cours pivots dont va dépendre le mécanisme de change. L'écu, qui évoque une ancienne monnaie française, mais qui tire son nom de l'acronyme de *European Currency Unit*, est constitué à partir d'un panier des monnaies de la communauté. La pondération des monnaies dans le panier dépend de la taille du pays (PIB) et du volume de ses échanges commerciaux. La fixité du change dépend de la définition d'un cours central, le « cours pivot », entre les monnaies des pays membres prises deux à deux. Le franc français a ainsi un cours pivot en marks, en liras, en francs belges, etc. Une marge de fluctuation de +/- 2,25% autour de ce cours pivot est tolérée. Elle a même été étendue jusqu'en 1990 à +/- 6% pour la lire ainsi que pour la peseta espagnole (1989), la livre sterling (qui n'entre dans le mécanisme de change qu'en 1990) et l'escudo portugais (1992). Les vagues spéculatives de 1993 conduiront à un élargissement général des marges à +/- 15%. Le SME sera marqué par un fonctionnement asymétrique à l'avantage de l'Allemagne, dont la monnaie, le Mark, connaîtra une appréciation tendancielle grâce à une stratégie macroéconomique de taux de change fort et stable, obligeant les pays partenaires (dont la France) à maintenir des taux d'intérêt plus élevés pour défendre le taux de change et rassurer les marchés financiers.

Par ailleurs, si l'intégration économique a sans nul doute favorisé une certaine prospérité sur le Vieux Continent et le rattrapage du niveau de vie des États-Unis, le blocage de la croissance et le ralentissement du progrès technique sont à l'origine d'un décrochage de l'Europe par rapport au PIB/habitant américain (de l'ordre de 30%) à partir de la seconde moitié des années 1990. Les explications sont diverses et variées et suscitent des débats, mais les écarts

persistent de part et d'autre de l'Atlantique : une durée annuelle de travail moyenne plus faible en Europe (fondée sur l'hypothèse d'une préférence pour le loisir, néanmoins discutée par les macro-économistes), des taux d'emploi inférieurs, des politiques économiques inadaptées, un investissement insuffisant dans les nouvelles technologies, ou encore un rythme du progrès technique trop lent figurent parmi les facteurs explicatifs les plus souvent cités. Pourtant, si les avancées en matière d'intégration politique et d'Europe sociale restent modestes, l'intégration économique se poursuit à partir des années 1990 avec l'union monétaire de l'Europe, conçue comme une nouvelle étape indispensable pour compléter le marché unique et faciliter les échanges.

Les 28 pays membres de l'Union européenne (États membres) : l'Allemagne, l'Autriche, la Belgique, la Bulgarie, Chypre, la Croatie, le Danemark, l'Espagne, l'Estonie, la Finlande, la France, la Grèce, la Hongrie, l'Irlande, l'Italie, la Lituanie, la Lettonie, le Luxembourg, Malte, les Pays-Bas, la Pologne, le Portugal, la République tchèque, la Roumanie, le *Royaume-Uni*, la Slovaquie, la Slovénie et la Suède. *Suite au référendum du 23 juin 2016 en faveur du Brexit, le Royaume-Uni est engagé dans un processus de séparation de l'UE, qui doit se terminer en mars 2019.*

Les 19 pays membres de la zone euro selon la date d'adhésion de ces pays à la monnaie unique sont les suivants : Allemagne : 1999, Autriche : 1999, Belgique : 1999, Chypre : 2008, Espagne : 1999, Estonie : 2011, Finlande : 1999, France : 1999, Grèce : 2001, Irlande : 1999

Autres États utilisant l'euro :

Quatre micro-États, hors zone euro sont autorisés à utiliser la monnaie unique (Andorre, Monaco, Saint-Marin et le Vatican) et certains pays européens, comme le Monténégro ou le Kosovo, l'utilisent de facto.

Les 9 pays de l'UE hors zone euro sont : Bulgarie, Croatie, Danemark, Hongrie, Pologne, République tchèque, Roumanie, Royaume-Uni, Suède

16.2 LA MARCHE VERS L'UNION ÉCONOMIQUE ET MONÉTAIRE (UEM)

16.2.1 *Les efforts de convergence avant l'adoption de l'euro*

Ratifié en 1992, le Traité de l'Union européenne prévoit la création d'une monnaie unique et un calendrier pour y parvenir. Parmi les dispositions prévues par ce traité dit de Maastricht, figuraient en particulier les objectifs et les missions des pays membres de l'Union européenne sur le

plan macroéconomique. Selon l'article 2 du traité : « *La Communauté a pour mission, par l'établissement d'un marché commun, d'une union économique et monétaire et par la mise en œuvre de politiques et d'actions communes, de promouvoir un développement harmonieux et équilibré des activités économiques dans l'ensemble de la Communauté, une croissance durable et non inflationniste [...], un haut degré de convergence des performances économiques, un niveau d'emploi et de protection sociale élevé* ».

L'article 3A du traité donnait clairement la priorité à la lutte contre l'inflation : « *L'action des États membres comporte la fixation irrévocable des taux de change conduisant à l'instauration d'une monnaie unique [...] ainsi que la définition et la conduite d'une politique monétaire et d'une politique de change uniques dont l'objectif principal est de maintenir la stabilité des prix et, sans préjudice de cet objectif, de soutenir les politiques économiques générales dans la Communauté, conformément au principe d'une économie de marché ouverte où la concurrence est libre* ».

De la même manière, les critères de convergence inscrits dans le traité de Maastricht donnaient la priorité à la stabilité monétaire et budgétaire : en effet, le processus d'adoption de la monnaie unique a consisté en une période de convergence des politiques économiques au cours des années 1990, afin de faciliter le rapprochement des situations macroéconomiques des pays candidats.

Ainsi, les critères de convergence nominale suivants ont été choisis pour réaliser l'union monétaire, et inscrits dans le Traité de Maastricht :

- Un degré élevé de stabilité des prix apprécié par un taux d'inflation proche (pas plus de 1,5 point de plus) de celui des trois États membres (au plus) qui ont les meilleurs résultats en la matière ;
- Une capacité d'emprunt : le taux d'intérêt des emprunts à long terme (10 ans) ne doit pas dépasser de plus de 2 points la moyenne des taux des trois pays où l'inflation

est la plus faible. Ce second critère est presque un corollaire du premier (le taux d'intérêt dépend du taux d'inflation anticipé) et accorde une réelle importance au jugement des marchés financiers sur la qualité de la gestion de l'économie ;

- Une stabilité des taux de change : elle suppose la participation de la monnaie nationale au mécanisme de change européen, l'absence de dévaluation du taux officiel déclaré dans ce cadre et le maintien depuis deux ans du taux effectif du marché dans les marges normales de fluctuation (fixées autrefois à $\pm 2,25\%$ par rapport au taux officiel et élargies en août 1993 à $\pm 15\%$) ;

- Une situation des finances publiques soutenable (absence de déficit public excessif) : les valeurs de référence sont 3% pour le rapport entre déficit prévu ou effectif et le PIB et 60% pour le rapport entre la dette publique et le PIB. Le Traité laisse cependant une marge d'appréciation en cas de dépassement faible et exceptionnel ou de dépassement faible, mais décroissant par rapport aux seuils fixés.

Ainsi, chaque État devait, avant son entrée dans l'euro, mettre en œuvre les mesures nécessaires à un assainissement de ses finances publiques (réduction des dépenses publiques, hausse des impôts) et garantir la stabilité monétaire comme objectif prioritaire (stratégies de rigueur monétaire), notamment en réformant les statuts de la banque centrale, pour en assurer l'indépendance face au pouvoir politique.

16.2.2 Les coûts et les bénéfices de l'Union monétaire

La monnaie unique a été conçue comme complément du marché unique : on sait que l'usage d'une même monnaie permet de faciliter les échanges au sein d'une zone-commerciale.

Les avantages théoriques de la monnaie unique ont été décrits par M. Emerson, l'un des directeurs de la Commission européenne, dans un rapport publié en 1990. La monnaie unique est censée apporter de nombreux bénéfices microéconomiques :

- L'Union économique et monétaire (UEM) permet de réduire l'incertitude liée aux fluctuations des taux de change au sein de la zone, puisque les dévaluations des monnaies nationales peuvent perturber le commerce extérieur (un exportateur a des difficultés à anticiper ses recettes) et le calcul de la rentabilité anticipée des investissements étrangers ;
- L'UEM entraîne la suppression des coûts de transaction et favorise le développement des échanges intracommunautaires, puisque les conversions entre les monnaies nationales généraient des coûts sous forme de commission de change (pour les touristes notamment) ;
- L'UEM favorise la transparence en termes de prix et de coûts, ce qui améliore les conditions de la concurrence et la circulation de l'épargne au sein de la zone (allocation du capital plus efficiente), et une baisse des taux d'intérêt favorable à l'investissement productif.

Le projet de monnaie unique visait également des objectifs macroéconomiques :

- La mise en place de l'euro devait permettre de partager le pouvoir monétaire au sein de l'UEM. Au cours des années 1980, le fonctionnement du système monétaire européen (SME) est progressivement devenu asymétrique. Le choix d'une banque centrale indépendante du pouvoir politique et les statuts retenus en matière de politique monétaire devaient permettre de maintenir une inflation faible sur longue période, gage d'un fonctionnement plus efficace des marchés.
- L'UEM devait faire de l'euro une monnaie internationale largement utilisée dans les transactions internationales, et

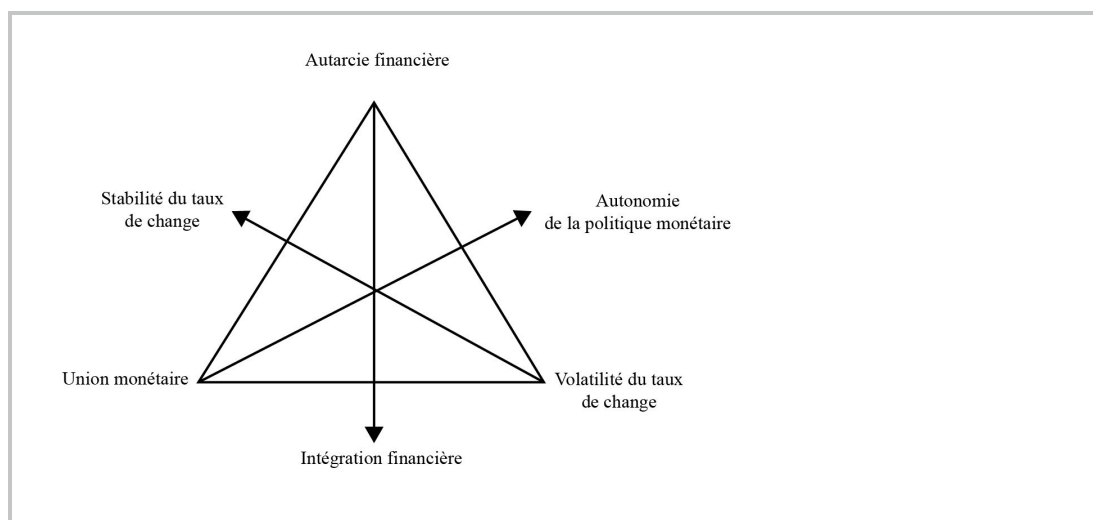
une monnaie de réserve pour les banques centrales internationales.

Pour autant, le passage à l'intégration monétaire ne s'est pas effectué sans coûts :

– L'Union économique et monétaire nécessite d'accepter de perdre de l'autonomie dans la politique monétaire nationale. Dans un contexte de globalisation financière et de mobilité des capitaux, la fixité des taux de change dans l'UEM fait peser une forte contrainte sur les pays à monnaie faible, condamnés à maintenir des taux d'intérêt élevés pour empêcher les sorties de capitaux. Cette contrainte a été représentée dans le « triangle des incompatibilités » exposé par R. Mundell (prix Nobel d'économie en 1999), selon lequel il est impossible de combiner trois facteurs : les taux de changes fixes (comme dans le SME), la mobilité parfaite des capitaux et l'autonomie des politiques monétaires. En effet, un pays peut choisir d'adapter deux de ces trois points, mais devra nécessairement renoncer au troisième. Par exemple, un pays engagé dans un régime de changes fixes et qui souhaite conserver l'autonomie de sa politique monétaire doit conserver une réglementation restreignant la mobilité des capitaux : si ce n'était pas le cas, toute politique monétaire faisant varier le taux d'intérêt déclencherait des mouvements de capitaux qui affecteraient le taux de change fixé sur le marché. Le pays serait alors contraint d'utiliser la politique monétaire pour ramener le taux de change à sa parité. De la même manière, si le pays souhaite privilégier la mobilité des capitaux et conserver l'autonomie de sa politique monétaire, il ne doit pas avoir à défendre la parité de change et il doit donc adopter un régime de changes flexibles.

Les pays européens ont longtemps cherché à maintenir un régime de changes fixes entre leurs membres dans le cadre du SME : mais dans un contexte de mobilité des capitaux et de globalisation financière à partir des années 1990, les États devaient consacrer leur politique monétaire

au maintien de la parité externe, au détriment de la recherche de l'équilibre interne. Mais les vagues spéculatives qui ont frappé les différentes monnaies européennes au cours des années 1990 engagées dans le SME ont rendu cet objectif toujours plus difficile à atteindre : l'Union économique et monétaire (UEM) a pu apparaître comme une réponse à ce dilemme puisque les pays ont renoncé à leur politique monétaire nationale, et l'ont transférée à une institution supranationale, la Banque Centrale européenne (BCE). Dès lors, la politique monétaire unique peut être employée pour atteindre un objectif interne à la zone euro (la maîtrise de l'inflation), car les changes flexibles par rapport au reste du monde lui permettent de retrouver une certaine autonomie.



GRAPHIQUE 16.2.
Le triangle des incompatibilités

Les modalités de la mise en place de la monnaie unique ont conduit à une perte de souveraineté des politiques économiques des États membres. Ainsi, le transfert de la souveraineté monétaire à la Banque Centrale européenne (BCE) implique que chaque État perd sa capacité à stimuler la croissance et l'emploi à court terme et à lutter contre les chocs asymétriques, soit des chocs exogènes qui frappent un État membre de la zone en particulier.

- Par ailleurs, la politique budgétaire des États membres est encadrée par des règles de discipline communes et se trouve subordonnée à la politique monétaire (dont l'objectif prioritaire est la stabilité des prix selon les statuts de la Banque Centrale européenne), ce qui limite également sa capacité à lutter contre les chocs asymétriques. Or, c'est en théorie l'autonomie de la politique budgétaire qui est censée les amortir dans une zone où la politique monétaire est unique.
- Enfin, l'Union économique et monétaire implique une politique de change unique vis-à-vis des monnaies extérieures à la zone euro (zone yen et zone dollar, notamment) : cela signifie qu'un pays de la zone euro ne peut procéder à un ajustement monétaire pour restaurer la compétitivité de ses exportations sur les marchés internationaux.

16.3 LE SYSTÈME DE POLITIQUE ÉCONOMIQUE DE LA ZONE EURO

16.3.1 *La zone euro, une zone monétaire (non) optimale*

L'instauration d'une monnaie unique dans une économie où les marchés des biens, des services et des facteurs de production sont intégrés, comme sur le marché européen, suppose de respecter certaines conditions. La théorie des zones monétaires optimales (ZMO) examine justement ce lien entre intégration économique et intégration monétaire : selon cette théorie, les pertes dues à l'abandon de l'instrument du taux de change dans une union monétaire sont d'autant plus faibles qu'un certain nombre de critères sont respectés. Comme nous l'avons vu, la perte principale due à la politique monétaire unique est l'impossibilité de répondre à un choc asymétrique : le pays en question ne peut plus mettre en

œuvre de relance monétaire ni opérer une dévaluation en cas de crise, s'il a perdu sa souveraineté monétaire et l'autonomie de sa politique de change. Selon R. Mundell, une zone monétaire est dite optimale si « *lors des chocs externes asymétriques affectant la demande de produits ou l'offre de facteurs d'une au moins des régions membres, des procédures automatiques d'ajustement sont mises en œuvre, évitant ainsi des modifications des taux de change nominaux entre les régions* » (Mundell, 1961).

La théorie des ZMO montre donc que la création d'une union monétaire comme la zone euro est pertinente si et seulement *si* les avantages de la disparition des coûts de transaction et du risque de change *dépassent* les inconvénients de la perte de la politique monétaire pour faire face aux chocs asymétriques qui frappent un ou plusieurs États membres de la zone, et que l'on distingue des chocs symétriques qui exercent des effets macroéconomiques pour toute la zone.

Un certain nombre de conditions doivent être réunies pour qu'un groupe de pays ait intérêt à rejoindre une union monétaire comme la zone euro. Les principaux critères des zones monétaires optimales sont les suivants :

- R. Mundell (“A Theory of Optimum Currency Area”, *American Economic Review*, 1961) insiste sur la mobilité du facteur travail : lorsqu'un pays de la zone monétaire est frappé par une crise, le déplacement des travailleurs vers les pays qui créent des emplois doit permettre de réduire le chômage dans le pays concerné. L'ajustement par la mobilité du travail peut compenser l'abandon du taux de change comme variable d'ajustement. De plus, la flexibilité des salaires doit permettre de remédier aux chocs de demande asymétriques : la hausse du chômage dans le pays en crise doit baisser les salaires et restaurer la compétitivité des produits.
- Selon R. McKinnon (“Optimum Currency Areas”, *American Economic Review*, 1963), une union monétaire est pertinente si le degré d'ouverture des économies qui la

composent est relativement proche. Une union monétaire doit inclure les principaux partenaires commerciaux d'une région qui n'ont plus à craindre la volatilité des taux de change bilatéraux : si une économie est très ouverte, son taux d'inflation sera peu différent de celui de ses partenaires, en raison de la concurrence liée à l'intégration commerciale. L'arme du taux de change serait peu efficace, puisqu'une dévaluation se traduirait par une hausse du prix des biens importés et par une perte de compétitivité par rapport aux partenaires commerciaux.

– P. Kenen (*The theory of optimum currency areas: an eclectic view*, Chicago University Press, 1969) fait valoir qu'une économie très diversifiée dans sa production et dans sa consommation est moins sensible aux chocs économiques exogènes : si l'économie est spécialisée dans un secteur particulier, un choc frappant celui-ci peut déstabiliser toute l'économie nationale. Une économie très diversifiée est moins sensible aux chocs asymétriques et peut intégrer une union monétaire.

– D'autres travaux ont ajouté un certain nombre de conditions préalables, parmi lesquelles le fédéralisme budgétaire et fiscal (un budget commun au sein de l'union monétaire pourrait permettre d'opérer des transferts budgétaires pour compenser les différences d'activité économique entre les pays membres), ou la nécessité de l'intégration financière et de la mobilité du capital, celle de la convergence des taux d'inflation entre les pays et de la convergence des institutions sociales et politiques.

– En parallèle avec les avancées empiriques, a émergé dans la recherche en économie une « nouvelle » théorie des zones monétaires optimales. Cette théorie insiste sur le caractère endogène de la satisfaction des critères d'optimalité au sens de R. Mundell. Cela signifie que les critères de la convergence nominale au sein d'une union monétaire sont sous-tendus par la théorie de la crédibilité et la prise en compte de la cohérence temporelle des décisions de politique économique. Cette nouvelle théorie

des ZMO s'appuie notamment sur les hypothèses de la nouvelle école classique (NEC) qui introduit la réaction des agents privés (ménages, entreprises) aux annonces des décideurs publics en matière de politique économique. L'hypothèse d'anticipations rationnelles de la part des agents privés implique alors que l'efficacité de la politique monétaire dépend largement de la crédibilité des autorités monétaires. L'adhésion à une union monétaire peut avoir pour effet de renforcer la réputation des autorités dans la conduite de la politique économique.

La zone euro se rapproche certes d'une ZMO par certains aspects :

- Le processus de convergence macroéconomique et le respect des critères de Maastricht avaient pour but de rapprocher les pays candidats à la monnaie unique d'une ZMO, et l'euro a été conçu comme un prolongement de l'intégration économique. Avant l'euro, la création du marché unique supposait une certaine harmonisation des politiques économiques et la garantie des quatre libertés concernant la circulation des biens et des services, du travail et du capital à l'intérieur de l'unité économique ainsi créée, afin que les conditions d'accès aux marchés des pays membres soient les mêmes pour toutes les entreprises et tous les consommateurs, quelle que soit leur origine.
- Le taux d'ouverture des économies européennes est relativement élevé et une grande partie des échanges commerciaux a lieu avec des partenaires européens aux structures productives et au niveau de vie moyen relativement proches. De plus, les différentes phases de libéralisation des échanges ont permis de réduire les écarts de prix au sein de la zone euro (par la convergence des taux d'inflation).
- En matière de diversification productive, un certain nombre de travaux montrent que les pays membres de la zone euro ont une spécialisation productive moins affirmée

que les États fédérés des États-Unis. Les pays européens ont donc une production diversifiée et consomment de manière globalement similaire : ils sont donc moins sujets aux chocs asymétriques, puisqu'ils sont plus homogènes que les États américains.

Pourtant la zone euro reste encore éloignée d'une ZMO dans de nombreux domaines :

- Même si les différents marchés nationaux au sein de la zone euro sont mis en relation directe, ils restent pour ainsi dire juxtaposés, et certaines règles du jeu économique sont fixées par chaque État, ce qui constitue un puissant obstacle aux échanges à l'intérieur du marché unique. Par exemple, malgré les progrès de l'intégration économique, les échanges commerciaux restent soumis aux barrières non tarifaires (comme des réglementations différentes dans les différents États), et les économies européennes restent très diverses en matière d'institutions et de systèmes sociaux.

- La mobilité des facteurs de production, comme d'ailleurs aux États-Unis, reste imparfaite : l'hétérogénéité des systèmes de protection sociale et la persistance des barrières culturelles et linguistiques peuvent être des obstacles importants à la mobilité des travailleurs au sein de la zone euro, et limiter les déplacements de main-d'œuvre en cas de crise, d'autant qu'un certain nombre de rigidités institutionnelles réduisent la flexibilité des salaires sur les différents marchés du travail. Par exemple, une montée du chômage en Espagne ne conduit pas forcément à des déplacements de main-d'œuvre vers les pays du nord de la zone euro.

- La concurrence sur le marché unique a accentué l'hétérogénéité de la zone euro en matière de spécialisation productive : la libéralisation des échanges a favorisé une organisation territoriale des activités productives polarisée entre les pays du nord et les pays du sud de l'Union européenne. Le Nord a renforcé ses

activités industrielles tournées vers les nouvelles technologies et créatrices d'emplois qualifiés, et a développé des économies d'agglomération : en situation d'incertitude, les entreprises vont en effet profiter de la faible distance qui les sépare les unes des autres pour bénéficier des fournisseurs et des clients des autres, de la diversité de l'offre productive, de l'importance du marché du travail et du partage des infrastructures ; tandis que le Sud a entamé un processus de désindustrialisation et une spécialisation tournée vers les services intensifs en main-d'œuvre peu qualifiée.

Déséquilibres courants en zone euro : un problème de taille du secteur exportateur

– Enfin, dans la zone euro, les États ont conservé leur souveraineté budgétaire (même si des règles de discipline budgétaire limitent les déficits publics et la dette publique) et cette absence de réel fédéralisme budgétaire et fiscal l'éloigne d'une véritable zone monétaire optimale (ZMO). La taille du budget communautaire reste par ailleurs trop faible (environ 1% du PIB de l'Union européenne) pour réaliser des transferts budgétaires en cas de chocs asymétriques (si un État est en crise, un budget commun peut prélever des ressources dans les États en croissance et les redistribuer vers les zones en difficulté).

16.3.2 *L'organisation du “policy-mix” dans la zone euro*

L'Union économique et monétaire (UEM) est un accord entre les pays membres de l'Union européenne visant à créer une monnaie unique (l'euro) avant le premier janvier 2002 et à transférer leur souveraineté monétaire à une banque centrale unique : la Banque Centrale européenne (BCE). L'union monétaire de l'Europe s'inscrit dans un vaste processus d'intégration économique et commerciale historique qui caractérise la construction européenne. La zone euro, qui regroupe les pays de l'Union européenne (UE) ayant adopté

l'euro, compte aujourd'hui 19 membres. Cette UEM demeure une construction institutionnelle originale qui rend délicate la mise en œuvre de la politique conjoncturelle, soit la combinaison de la politique monétaire et de la politique budgétaire afin de réguler la demande globale et stimuler la croissance sans accélération de l'inflation. Dans l'UEM, cette politique conjoncturelle est caractérisée par une politique monétaire unique, fédérale et indépendante des États membres, tandis que les politiques budgétaires restent décidées à l'échelon national. Les Traités européens ont encadré la politique conjoncturelle par des normes strictes, à la fois sur le plan monétaire et sur le plan budgétaire pour limiter les effets négatifs dus à l'interdépendance de la politique monétaire unique et des 19 politiques budgétaires dans l'UEM. Sur le plan monétaire, la Banque Centrale européenne (BCE) a un objectif de stabilité des prix prioritaire dans le cadre de ses statuts, tandis que son indépendance implique que les États membres de l'UEM ne peuvent pas lui donner d'ordre (notamment créer de la monnaie pour financer les déficits publics).

16.3.2.1 La politique monétaire unique de la BCE

Dans la zone euro, la politique monétaire est confiée au Système européen de banques centrales (SEBC), composé de la BCE et des banques centrales nationales de l'Union.

L'**objectif principal** de la Banque Centrale européenne (BCE) est de **maintenir la stabilité des prix** dans la zone euro, en vertu de l'article 2 du texte fixant ses statuts. Le traité sur le fonctionnement de l'Union européenne précise dans son article 105 que « *l'objectif principal du SEBC est de maintenir la stabilité des prix. Sans préjudice de l'objectif de stabilité des prix, le SEBC apporte son soutien aux politiques économiques générales dans la Communauté, en vue de contribuer à la réalisation des objectifs de la Communauté, tels que définis à l'article 2* ».

Les tâches principales du Système européen de banques centrales sont les suivantes :

- La mise en œuvre de la politique monétaire dans la zone euro (article 105 (2) du traité sur l'Union européenne et article 3 des statuts du SEBC) ;
- La conduite des opérations de change ; la détention et la gestion des réserves de change des États membres ;
- Le développement des systèmes de paiement dans la zone.

La BCE et les banques centrales nationales sont indépendantes des gouvernements nationaux et de la Commission européenne, et elles ne peuvent accepter d'instructions de leur part (en vertu de l'article 108 du Traité sur l'Union européenne). En outre, les banques centrales ne sont plus autorisées à « *accorder des découverts ou tout autre type de crédits aux autorités publiques* » ; « *l'acquisition directe, auprès d'elles, par la BCE ou les banques centrales nationales, des instruments de leur dette est également interdite* » (en vertu de l'article 104 du traité). L'Eurosystème, institution qui regroupe la Banque Centrale européenne et les banques centrales nationales des États membres de l'Union européenne qui ont adopté l'euro, est l'autorité monétaire de la zone euro. Son objectif principal est de « *maintenir la stabilité des prix pour le bien commun* » et il agit également « *en tant qu'autorité financière de premier plan* », et se donne pour mission de « *sauvegarder la stabilité financière et de promouvoir l'intégration financière européenne* ».

Le SEBC est dirigé par trois organes de décision de la Banque Centrale européenne : le directoire, le conseil des gouverneurs et le conseil général. Dans les faits, les deux organes de décision principaux sont le conseil des gouverneurs et le directoire. Avec l'Union économique et monétaire (UEM), la politique monétaire devient unique, les banques centrales des pays de la zone euro doivent appliquer les mesures décidées par la BCE et sont ainsi devenues avant tout des courroies de transmission. En effet, la crédibilité de la BCE repose notamment sur sa capacité à

préciser son indépendance vis-à-vis des gouvernements nationaux.

L'objectif prioritaire de la Banque Centrale européenne demeure la lutte contre l'inflation : maintenir la stabilité des prix dans la zone euro, à des niveaux proches d'une cible de taux d'inflation fixée à 2%.

En matière de politique de change, le traité (article. 219) prévoit un partage des compétences entre le Conseil et le Système européen des banques centrales (SEBC), lequel conduit, au jour le jour, la politique de change en tenant compte de la souveraineté des États et de son objectif de stabilité des prix.

Même s'ils font l'objet de controverses dans le champ de la recherche en économie, la Banque Centrale européenne escompte tirer un certain nombre de bénéfices de son objectif principal de stabilité des prix dans la zone euro :

- La stabilité des prix facilite la transparence des prix relatifs : dans un environnement où la volatilité des prix est faible, il est plus aisé de comparer le prix des différents biens.
- La stabilité des prix réduit les primes de risques incorporées dans les taux d'intérêt de long terme et favorise l'investissement productif par une baisse des coûts de l'endettement.
- La stabilité des prix réduit les coûts nécessaires pour se protéger contre l'inflation : elle limite les risques de fuite devant la monnaie et l'achat d'instruments de couverture par les entreprises, indispensables lorsque l'inflation est forte.
- La stabilité des prix limite les distorsions fiscales : ainsi la hausse de prix peut gonfler les revenus des agents et les faire passer dans une tranche d'impôts supérieure et désinciter au travail et à la production et à l'épargne.
- La stabilité des prix favorise la cohésion sociale et politique : de nombreux épisodes de troubles politiques dans l'histoire des nations européennes ont été favorisés

par des périodes d'hyperinflation (comme en Allemagne en 1923). La stabilité des prix est enfin un moyen de préserver le pouvoir d'achat des ménages et de maintenir la compétitivité-prix des produits européens sur les marchés internationaux.

Les avantages attendus et les avantages effectivement constatés peuvent diverger.

16.3.2.2 Les politiques budgétaires nationales encadrées

Les politiques budgétaires des États-membres de la zone euro sont encadrées par le Pacte de Stabilité et de Croissance (PSC) a été adopté au Conseil européen d'Amsterdam en juin 1997. Ce traité d'Amsterdam instaure le PSC qui limite le déficit budgétaire des États à 3% du PIB et la dette publique à 60% du PIB. L'objectif est d'éviter les dérives liées aux politiques budgétaires expansionnistes et à leurs effets comme l'accélération de l'inflation, et le gonflement de la dette publique. En effet, les règles budgétaires doivent permettre d'éviter que les excès des uns ne soient pas subis, payés, par les autres. Il s'agit d'éviter les comportements de passager clandestin qui consistent à faire payer à tous la hausse des taux d'intérêt induite par la hausse des dépenses publiques de quelques-uns. En effet, en zone monétaire, il existe des effets de débordements (sur les partenaires) positifs, ou négatifs. Mais le PSC se traduit dans la pratique par une gestion procyclique du budget et donc par la fin du jeu des stabilisateurs automatiques.

Le PSC comporte deux types de dispositions :

- La surveillance multilatérale est une disposition préventive qui oblige les États de la zone euro à présenter leurs objectifs budgétaires à moyen terme dans le cadre d'un programme de stabilité.

En 2005, la notion de solde proche de l'équilibre a été remplacée par la proximité de l'équilibre structurel par ailleurs avec des soldes structurels différents selon les pays (de -0,5 à + 0,5% du PIB). Un système d'alerte rapide permet au Conseil ECOFIN, réunissant les ministres de l'Économie et

des Finances de l'Union, d'adresser une recommandation à un État en cas de dérapage budgétaire.

– La procédure des déficits excessifs, disposition dissuasive, est enclenchée dès qu'un État dépasse le critère de déficit public fixé à 3% du PIB, sauf circonstances exceptionnelles. Le Conseil ECOFIN adresse alors des recommandations pour que l'État mette fin à cette situation. Si tel n'est pas le cas, le Conseil peut prendre des sanctions et obliger l'État à faire un dépôt auprès de la Banque Centrale européenne qui peut devenir une amende (de 0,2 à 0,5% du PIB de l'État en question) si le déficit excessif n'est pas comblé.

16.3.3 *Le dilemme de la régulation conjoncturelle dans la zone euro*

16.3.3.1 Les contradictions de la gouvernance de la politique économique dans la zone euro

La crise des dettes souveraines à partir de 2009 a démontré les contradictions de la gouvernance de la politique économique dans la zone euro. Les États membres sont en effet soumis à l'heure actuelle à un dilemme :

- En phase de récession, si les pays fragilisés par la crise des dettes publiques s'engagent dans des plans d'austérité budgétaire trop précoces qui asphyxient la demande et freinent la croissance potentielle, la défiance des marchés financiers risque de grandir quant à leur solvabilité future ;
- S'ils laissent s'aggraver les déficits publics pour éviter la spirale de la récession, ils s'exposent à une remontée des taux d'intérêt et à une hausse du coût de leurs emprunts sur les marchés de capitaux.

Dès lors, les écarts ne peuvent que se creuser au sein de la zone euro. Avec le choix structurel de se priver d'une politique monétaire autonome et, en pratique, en l'absence d'une mobilisation de la politique de change, la faiblesse

actuelle du budget européen (et son nécessaire équilibre) empêche une réelle redistribution entre les États sous la forme de transferts financiers suffisamment puissants (ce que l'on appelle le « fédéralisme budgétaire »).

Enfin, la zone euro reste un espace où la tentation est forte de recourir à la concurrence fiscale pour attirer les investisseurs étrangers et abaisser le coût du travail, puisque les leviers de la politique conjoncturelle sont contraints : dans une économie globalisée caractérisée par une forte mobilité du capital, la réduction des impôts sur le capital peut devenir un moyen de restaurer la compétitivité, au risque d'un niveau sous-optimal de services collectifs et de protection sociale (puisque l'État voit ses recettes fiscales se tarir).

De fait, en l'absence de fédéralisme, les déficits de balance des transactions courantes sont très difficiles à réduire sans payer le coût d'une forte récession et d'une forte hausse à long terme des dettes extérieures (l'unification monétaire sans fédéralisme impose l'équilibre extérieur à chaque État membre). Aux États-Unis, les États fédérés peuvent accumuler des déficits extérieurs importants qui sont compensés par des transferts à l'échelon fédéral.

16.3.3.2 Le renforcement de la discipline budgétaire après la crise des dettes publiques

Les défaillances de la gouvernance économique de la zone euro qui se sont révélées au moment de la crise des dettes publiques à partir de 2009 sont largement liées aux incohérences institutionnelles du projet originel : la zone euro reste un modèle de gouvernance très particulier qui combine une politique monétaire unique et dix-neuf politiques budgétaires, certes souveraines, mais encadrées par des règles. Ce choix d'un gouvernement par les règles et les normes budgétaires semble compliquer considérablement la mise en œuvre du "policy-mix", soit la combinaison de la politique monétaire et de la politique budgétaire, qui demeure sous-optimale dans la zone euro puisque la politique monétaire unique est confiée à une institution fédérale qui

dispose d'une indépendance forte, la Banque Centrale européenne, tandis que la politique budgétaire demeure nationale et soumise au respect de normes inscrites dans les traités européens.

C'est ce qui explique que la gestion macroéconomique de la zone euro depuis 1999 n'a pas pu être optimale, car il n'existait qu'une seule et même politique monétaire pour des pays membres qui se trouvaient dans des situations économiques très différentes, du fait de l'hétérogénéité structurelle originelle et croissante entre économies. En effet, la politique monétaire unique se traduit par la conjonction d'un taux d'intérêt nominal unique et de taux d'inflation différents : cela entraîne des écarts importants de taux d'intérêt *réels* qui modifient le coût relatif de la dette, les incitations et les arbitrages en matière d'endettement et d'investissement.

On a pu observer une coordination des politiques macroéconomiques de nature intergouvernementale. Un seul instrument restait à disposition des États, légitime d'un point de vue démocratique, dans le cadre du Pacte de stabilité et de croissance (PSC) : la politique budgétaire. Mais cette organisation du "policy-mix" a pu favoriser les attitudes de type « passager clandestin » ou « cavalier seul » (*free rider*), au détriment de la coopération et de l'intérêt collectif : si par exemple un État membre met en œuvre une politique de relance budgétaire, les conséquences négatives en termes de freinage de la demande globale sont socialisées dans toute la zone euro à partir du moment où la BCE relève ses taux d'intérêt pour prévenir le risque d'inflation.

16.4 LES ÉVOLUTIONS INSTITUTIONNELLES DE LA ZONE EURO

16.4.1 *La zone euro dans les crises et les réponses institutionnelles*

En raison de la crise financière puis de la transmission de ses effets à l'économie « *réelle* » (ce que certains économistes

ont appelé la « *Grande Récession* ») les gouvernements des pays de la zone euro ont été confrontés à un creusement des déficits publics et un gonflement de leur dette publique, bien au-delà du seuil des normes inscrites dans les textes européens. Après l'introduction de la Monnaie unique, une bulle de crédit s'est formée, puis a explosé avec la crise financière ; le risque de défaut de remboursement a brutalement augmenté avec la chute des prix de l'immobilier, la récession et la stagnation des revenus, contraignant les pays du sud de la zone à de difficiles plans de rigueur budgétaire, à l'instar de la Grèce (et de l'Espagne).

La crise des dettes souveraines (voir chapitre 20) a souligné le manque de coordination des politiques économiques en Europe et l'insuffisance du fédéralisme budgétaire (ce que certains appellent le « *coût de la non-Europe* »). Sous l'effet de la crise, diverses modifications institutionnelles ont été adoptées (création d'un Mécanisme européen de stabilité, adoption du Traité sur la stabilité, la coordination et la gouvernance en Europe, notamment). Cependant, un débat subsiste aujourd'hui entre deux perspectives : celle, soutenue par l'Allemagne, qui met l'accent sur le respect des règles budgétaires, dans la tradition de l'ordo-libéralisme, et celle, soutenue notamment par la France, qui défend un renforcement de la solidarité et un renforcement des politiques communes en matière d'intégration, de croissance et de cohésion sociale.

Dans la note de France Stratégie intitulée « Quelle architecture pour la zone euro ? », publiée en décembre 2016, les auteurs Christophe Gouardo et Vincent Aussilloux précisent qu'« aux États-Unis, les États fédérés sont chacun responsables de leur endettement public. En contrepartie, et bien que les États s'en soient eux-mêmes dotés, l'État fédéral n'impose pas de règles budgétaires aux échelons infranationaux. En revanche, en temps de crise, les transferts du budget fédéral (notamment au titre de l'assurance chômage) permettent d'assurer une forme de stabilisation macroéconomique ». Selon eux, face aux défaillances de la

coordination des politiques budgétaires en zone euro, « cette architecture pourrait être déclinée à l'échelle de la zone euro afin de combiner, d'une part, la volonté de renforcer la souveraineté budgétaire au niveau national – non contrainte par des règles européennes – et la nécessité, d'autre part, de créer une fonction de stabilisation à l'échelle de la zone ».

16.4.2 *Les débats autour des réformes de la gouvernance de la zone euro*

D'un point de vue plus institutionnel, les États et la BCE ont mis en place en 2010 des mécanismes de solidarité budgétaire entre États à la suite de la crise des dettes souveraines dans la zone euro, le Fonds européen de stabilité financière (FESF). C'est une structure de financement *ad hoc* pouvant mobiliser jusqu'à 440 milliards d'euros. Les ressources du FESF sont des titres de dette émis sur les marchés qui sont garantis par les États, au prorata de leur participation au capital de la BCE. Le FESF a été pérennisé en devenant le MES (Mécanisme européen de stabilité) en 2013. Avec le *Pacte Europlus*, le but est de renforcer la gouvernance de l'Union économique et monétaire, en déplaçant le curseur d'une discipline des pairs (surveillance mutuelle) vers une discipline de marché et des règles institutionnelles.

Par ailleurs, le Traité sur la stabilité, la coordination et la gouvernance (TSCG) au sein de l'Union économique et monétaire, dit « Pacte budgétaire européen », a été ratifié par la plupart des États membres de l'Union européenne pour imposer une discipline financière renforcée en réponse à la crise de la zone euro, et le traité est entré en vigueur au 1^{er} janvier 2013. Le Royaume-Uni et la République tchèque n'ont pas signé le traité et sont donc dispensés de cette mise en œuvre (l'article 15 du TSCG leur laisse la possibilité d'y adhérer *a posteriori* s'ils changent d'avis). Le TSCG comprend trois parties : le « pacte budgétaire » proprement

dit (avec sa règle budgétaire), la coordination des politiques économiques et leur convergence, la gouvernance de la zone euro. Ainsi, l'article 3 du TSCG énonce clairement une règle d'équilibre : « *a) la situation budgétaire des administrations publiques d'une partie contractante est en équilibre ou en excédent ; b) la règle énoncée au point a) est considérée comme respectée si le solde structurel annuel des administrations publiques correspond à l'objectif à moyen terme spécifique à chaque pays, tel que défini dans le pacte de stabilité et de croissance révisé, avec une limite inférieure de déficit structurel de 0,5% du produit intérieur brut aux prix du marché.* » Le TSCG prévoit que les États membres qui dépassent la norme de dette publique de 60% du PIB, qu'il réaffirme, réduisent leur niveau d'endettement public d'un vingtième par an, de manière concertée. Par exemple, si le PIB d'un pays s'élève à 1.000 milliards d'euros et que sa dette atteint 700 milliards (70% de son PIB), l'écart entre la dette effective et les 60% de dette (600 milliards d'euros) est de 100 milliards d'euros ($700 - 600 = 100$). La première année, ce pays devra donc réduire cet écart de 5 milliards d'euros ($1/20 \times 100 \text{ milliards} = 5$).

Le TSCG constitue un engagement solennel des États signataires à viser la stabilité budgétaire en adoptant des mesures appropriées afin de restaurer la confiance et réduire le poids de leur endettement, nécessaire au retour de la croissance, sous peine d'éventuelles sanctions financières. Pour autant, le TSCG est un traité original d'une certaine souplesse aux États. Le Traité sur la stabilité, la coordination et la gouvernance (TSCG) constitue un engagement solennel des États qui l'ont signé à viser la stabilité budgétaire en adoptant des mesures appropriées afin de restaurer la confiance et réduire le poids de leur endettement, nécessaire au retour de la croissance. C'est un traité « intergouvernemental » original, signé en dehors des procédures habituelles de l'Union européenne tout en reprenant les 5 règlements et une directive (le « six pack » de 2011) et en utilisant les institutions de l'Union (Commission et

Cour de justice). Cependant, le respect des engagements est confié au Conseil de l'Union européenne sous l'autorité du Conseil européen ce qui enlève la décision d'application à la Commission européenne qui, elle, reste gardienne des engagements pris en soumettant son rapport au Conseil. Les Parlements nationaux, par le biais de leurs commissions, sont associés à celle du Parlement européen dans le but de « *débattre des politiques budgétaires et d'autres questions régies par le traité* ». Les juges européens n'auront pas le droit de sanctionner les dérapages budgétaires, ce que voulait l'Allemagne, mais seulement le non-respect du traité dans le droit national d'un pays. En cas d'infraction avérée, la Cour peut alors prononcer des sanctions financières dont le produit est ensuite versé au Mécanisme européen de stabilité (MES).

Même s'il impose une discipline budgétaire stricte, le « Pacte budgétaire » laisse des marges d'interprétation non négligeables : prenant acte du fait que le solde structurel de beaucoup d'États membres n'est pas équilibré, il fait du respect de la « règle budgétaire » un « objectif de moyen terme » et le retour à l'équilibre est inscrit dans un « calendrier de convergence » élaboré par les États membres. Les 0,5% maximum à obtenir sont portés à 1% pour les pays plus rigoureux, ayant une dette publique inférieure à 60%. Les États membres qui sont responsables de la mise en œuvre des mesures correctives ont le choix et la responsabilité de les adapter au contexte. La notion de « *solde structurel* » exclut « *les variations conjoncturelles* » et les « *mesures ponctuelles exceptionnelles* » et le Traité stipule que la règle peut être suspendue « en cas de circonstances exceptionnelles » (dont « *les graves récessions économiques* » et les « *faits inhabituels indépendants de la volonté des États* »).

À l'heure actuelle, de nombreux macro-économistes (dont Agnès Bénassy-Quéré, Jean Pisani-Ferry, Patrick Artus, Michel Aglietta, etc.) insistent sur les défaillances du système de politique économique de la zone euro. Lorsqu'ils

examinent ce qui n'a pas fonctionné depuis 1999, ils soulignent le fait que la conduite de la politique économique est restée polarisée sur la surveillance budgétaire sans prêter suffisamment attention aux divergences structurelles (différences de salaire, de compétitivité entre les pays membres), et aux déséquilibres financiers nourris par la dette privée. Certains, comme Michel Aglietta, plaident pour une gestion d'ensemble de la demande en zone euro et des investissements publics et des dettes émises en commun (eurobonds). Ainsi, dans le cadre de la stratégie « *Europe 2020* », de nombreux experts insistent sur la nécessité de renforcer la croissance potentielle de l'Union européenne, mais aussi de repenser les conditions d'entrée dans la zone euro (à la lumière de la crise grecque et de l'expérience des douze premières années de l'euro) et d'avancer sur le dossier de la coordination des politiques fiscales et de l'union budgétaire. En l'absence d'un véritable fédéralisme budgétaire, sans doute hors d'atteinte à l'heure actuelle, ils préconisent d'éviter les ajustements budgétaires trop brutaux pour les pays en crise, tandis que certains plaident également pour l'établissement d'un système européen d'assurance chômage ou un fonds commun d'investissement qui pourraient jouer un rôle stabilisateur contracyclique (pour amortir les chocs en période de crise).

Dans une note du Conseil d'Analyse économique (CAE) de février 2016, A. Bénassy-Quéré, X. Ragot et G. Wolff tirent les enseignements de la crise de l'euro à l'été 2015 et recommandent de progresser rapidement sur la voie de l'union budgétaire dans la zone euro. En l'absence d'un véritable fédéralisme budgétaire. Ils préconisent d'éviter les ajustements budgétaires trop brutaux pour les pays en crise, de laisser le futur Conseil budgétaire européen coordonner la politique budgétaire en relâchant les contraintes du Pacte de stabilité et de croissance (PSC) dans les périodes de crise, et plaident également pour l'établissement d'un système européen d'assurance chômage qui pourrait jouer un rôle stabilisateur contracyclique. Selon les auteurs, au-delà des

réformes et priorités de court terme, il convient de garder à l'esprit les enjeux de long terme quand on réfléchit au fonctionnement de la zone euro. Dans le cadre de la stratégie « Europe 2020 », il est impératif de renforcer la croissance potentielle de l'Union européenne, mais aussi de repenser les conditions d'entrée dans la zone euro (à la lumière de la crise grecque et de l'expérience des douze premières années de l'euro) et d'avancer sur le dossier de la coordination des politiques fiscales et du fédéralisme budgétaire.

Néanmoins, si l'austérité budgétaire est aujourd'hui critiquée pour ses effets délétères sur la croissance, une majorité d'économistes mettent clairement en garde contre la tentation de démanteler la zone euro, à cause des coûts exorbitants d'une telle solution, et ce pour trois raisons, selon un économiste comme Patrick Artus (in *Euro. Par ici la sortie*, avec Marie-Paule Virard, Fayard, 2017) : 1) La ruine des épargnants et des emprunteurs qui s'ensuivrait à cause des distorsions de change qui alourdiraient les dettes extérieures en cas de retour aux monnaies nationales ; 2) La fragmentation désormais très poussée des chaînes de valeur mondiales des firmes multinationales qui rend le retour au protectionnisme et aux dévaluations compétitives impraticable ; 3) Une polarisation industrielle accrue entre le Nord et le Sud qui n'arrangerait pas les affaires, loin de là, des pays du sud de la zone euro, dont l'appareil productif serait alors plombé par des monnaies dépréciées, et des taux d'intérêt exorbitants imposés par les marchés financiers en cas de sortie de l'euro.

Malgré les blocages actuels, Patrick Artus et Marie-Paule Virard estiment que la zone euro doit sortir de la crise « par le haut » : ils plaident, avec d'autres macro-économistes, pour un renforcement du fédéralisme pour mieux coordonner les politiques économiques, et pour faire face en commun aux chocs macroéconomiques et stabiliser l'euro. Ils préconisent en particulier d'instaurer un budget commun de la zone euro, de créer un fonds d'investissement suffisamment doté, et de

développer un système d'assurance chômage, ou un revenu minimum européen.

Depuis l'élection d'E. Macron en 2017, la France plaide pour un budget commun spécifique dans la zone euro, et de taille assez conséquent (représentant plusieurs points de PIB), à la fois pour investir, mais aussi pour une fonction de stabilisation en cas de choc macroéconomique. La Commission européenne a ainsi proposé en mai 2018 un embryon de budget, via deux instruments distincts : 25 milliards d'euros pour soutenir les réformes jugées prioritaires (retraites, marché du travail...) et jusqu'à 30 milliards d'euros de prêts pour les pays qui subiraient un choc économique important (comme la survenue d'une catastrophe naturelle). Mais des divergences politiques avec l'Allemagne persistent : cette dernière se dit toutefois favorable à un budget « d'investissement » de plusieurs dizaines de milliards d'euros et à des prêts à court terme pour les pays confrontés à des crises d'origine extérieure, qui seraient accordés par un futur « Fonds monétaire européen ». Les États européens ont en effet acté au printemps 2018 la transformation du Mécanisme européen de stabilité (MES), chargé de superviser les prêts aux États membres en difficulté (comme la Grèce), en un « Fonds monétaire européen » (FME), qui pourrait venir en aide aux pays en crise, mais avec un droit de regard sur les stratégies budgétaires menées par les pays bénéficiaires. Ce FME pourrait endosser le rôle de « prêteur en dernier recours » (ou "backstop") pour les banques en détresse, dans le cas où les mécanismes de sauvegarde déjà prévus s'avéraient inefficaces. Pourtant, la création d'un budget commun de la zone euro (ou même celle d'un poste de ministre des Finances de la zone euro) rencontrent pour l'heure de profondes résistances politiques, et si en juin 2018 l'Allemagne et la France semblent s'être mis d'accord sur le principe d'un budget commun de la zone euro, les États peinent à se mettre d'accord sur sa taille et les modalités de fonctionnement, et au-delà, sur la voie d'une plus grande intégration politique au sein de l'euro.

L'IMPOSSIBILITÉ D'UNE UNION MONÉTAIRE SANS FÉDÉRALISME

En l'absence de fédéralisme, c'est-à-dire de transferts publics entre les pays, les déficits de la balance courante sont impossibles, car ils conduisent à une hausse insupportable à long terme des dettes extérieures.

Il existe donc une contradiction impossible à résoudre entre :

- l'unification monétaire sans fédéralisme qui impose l'équilibre extérieur à chaque État membre ;
- la spécialisation productive normale des pays qui génère des déficits extérieurs.

Le cas des États-Unis est instructif. Le fédéralisme permet aux États d'avoir des déséquilibres extérieurs considérables puisqu'ils sont compensés par les transferts fédéraux. Notamment les dépenses de santé (Medicare, Medicaid), les salaires des fonctionnaires sont financés par les impôts fédéraux.

Source : Patrick Artus, *La crise de l'euro*, Armand Colin, 2012

LES DÉSÉQUILIBRES MACROÉCONOMIQUES ET FINANCIERS INTERNATIONAUX

« Le contrôle de la liquidité mondiale et son maintien à un taux de croissance raisonnable supposeraient donc une coordination internationale forte des politiques monétaires et de change. Il faudrait que, d'une part, les pays déficitaires (Etats-Unis) mènent des politiques d'accroissement de leurs taux d'épargne, d'autre part, que les pays excédentaires (Asie, exportateurs de pétrole) essaient de réduire leurs taux d'épargne, passent à des régimes de change plus flexibles, qui conduiraient à une moindre accumulation de réserves de change. » - Patrick Artus, 2008

« Les économies de marché émergentes dégagant un excédent de leur compte de transactions courantes ont-elles donc généralement acquis des actifs sûrs et liquides, principalement aux États-Unis ? La réponse est positive. » - Ben Bernanke, 2011

« Comprendre le lien entre les déséquilibres internationaux et la stabilité financière est essentiel pour bien saisir non seulement ce qui s'est produit à l'occasion de la crise, mais aussi ce qui pourrait arriver à l'avenir.

Les déséquilibres et vulnérabilités qui affectaient l'économie mondiale avant la crise ont commencé à réapparaître. Une demande mondiale insuffisante et des taux d'intérêt exceptionnellement bas dans de nombreux pays créent un contexte propice à des comportements extrêmement risqués dans le secteur financier. Ne pas prendre de mesures pour

corriger les déséquilibres mondiaux ne peut en définitive que miner la reprise économique et la stabilité financière à l'échelle internationale. » - Marque Carney, Banque de France, 2011

SOMMAIRE

[17.1 Les déséquilibres macroéconomiques et financiers en « économie fermée »](#)

[17.2 Les déséquilibres macroéconomiques et financiers en économie ouverte](#)

[17.3 Les déséquilibres macroéconomiques et financiers avec distinction agents privés/agents publics](#)

Le financement des économies ouvertes, c'est-à-dire le financement des opérations et des agents économiques en économie ouverte, nous conduit à nous interroger sur les déséquilibres économiques et financiers internationaux. L'analyse de la relation entre épargne et investissement qui en découle, au niveau macroéconomique, est fondamentale pour comprendre le fonctionnement d'une économie. Alors qu'en économie fermée, l'épargne domestique détermine l'investissement, ce n'est plus vrai en économie ouverte. En effet, les échanges de capitaux permettent aux agents résidents d'être, au moins en partie, financés par les agents non résidents et réciproquement. Avec le processus de globalisation financière, les possibilités de financement se sont accrues de manière exponentielle.

Agents résidents et agents non résidents. Pour retracer l'ensemble des opérations des agents économiques résidents d'un pays avec l'étranger, la comptabilité nationale crée une catégorie « reste du monde » qui regroupe en fait les ménages, les entreprises, les administrations et les institutions financières non résidents qui effectuent des opérations avec des agents résidents. Un agent est considéré comme **résident** s'il exerce une activité sur le territoire national pendant au moins un an. Ainsi, un touriste anglais de passage pour une semaine à Paris est **non résident**, mais une entreprise allemande ou un travailleur immigré installé en France sont des agents résidents. Inversement, un Français travaillant à l'étranger est un non-résident (du point de vue de l'économie française).

17.1 LES DÉSÉQUILIBRES MACROÉCONOMIQUES ET FINANCIERS EN « ÉCONOMIE FERMÉE »

17.1.1 *De l'équilibre Ressources-Emplois en biens et services...*

Une économie fermée est une économie dans laquelle les agents résidents n'échangent pas avec les agents non résidents, ceux qui résident dans les autres économies nationales (pays). Il n'y a donc ni importation ni exportation. Autrement dit, en l'absence d'échanges entre l'économie nationale et le reste du monde (RDM), entre les agents résidents et les agents non-résidents, **l'Équilibre**

Ressources-Emplois en biens et services est donné par l'équation suivante :

$Y = C + I$ (1) Approche par la demande

Cette relation est une identité. Elle est toujours vérifiée, du fait de sa construction, et elle signifie que les biens et services à notre disposition ont une seule origine possible : notre propre production ($Y = \text{PIB}$). Ils sont utilisés pour la consommation (C) et l'investissement (I). Dans ce modèle, très simplifié, on ne fait pas la distinction entre les agents publics et les agents privés.

Étant donné que $Y = C + S$ (2) Approche par le revenu

En combinant (1) et (2), on obtient $C + I = C + S$.

17.1.2 ... à l'équation de l'équilibre entre épargne et investissement

Il vient :

$S = I$ (3) Équilibre Épargne = Investissement

Clairement, la satisfaction des besoins de financement des uns est limitée par la capacité de financement des autres. Si $S < I$, il ne sera possible de financer I qu'au prix d'une baisse de C , une hausse de S .

En économie ouverte, cette contrainte de l'épargne va être, en partie, levée.

17.2 LES DÉSÉQUILIBRES MACROÉCONOMIQUES ET FINANCIERS EN ÉCONOMIE OUVERTE

Par souci de simplification, nous avons fait l'hypothèse, et nous la ferons ci-dessous pour les économies ouvertes, qu'il n'existe pas de stocks, autrement dit tout ce qui est produit est vendu rapidement.

17.2.1 Les déséquilibres sans distinction entre agents publics et agents privés

Une économie ouverte est une économie dans laquelle les agents résidents échangent avec les agents non résidents, ceux qui résident dans les autres économies nationales (pays). Il existe des importations et des exportations. Les résidents exportent pour pouvoir importer. Les importations (M) sont des achats. Les exportations (X) sont des ventes. Or, pour pouvoir acheter, il faut vendre, sinon il faut s'endetter.

En économie ouverte, quand **on ignore la distinction entre les agents publics et les agents privés**, l'Équilibre Ressources-Emplois en biens et services devient :

$Y + M = C + I + X$ (1) Approche par la demande

Cette relation est une identité. Elle est toujours vérifiée, du fait de sa construction, et elle signifie que les biens et services à notre disposition ont deux origines possibles : notre propre production ($Y = \text{PIB}$) et nos importations (M). Ils sont utilisés pour la consommation (C), l'investissement (I) et les exportations (X). Dans ce modèle, très simplifié, on ne fait pas la distinction entre les agents publics et les agents privés.

Étant donné que $Y = C + S$ (2) Approche par le revenu

En combinant (1) et (2), on obtient : $C + I + X - M = C + S$

$(X - M) = (S - I)$ (3)

Cette identité est l'équation des déséquilibres en économie ouverte.

17.2.2 Les enseignements de l'équation des déséquilibres en économie ouverte

Il existe une relation entre les flux financiers internationaux (S-I) et les flux réels internationaux (X-M). En effet, à partir de cette simple identité, on peut comprendre pourquoi les déséquilibres extérieurs, observés en utilisant les balances des paiements courants (transactions courantes) des différentes économies, ont pour contrepartie des déséquilibres dans les flux de capitaux qui les financent. Autrement dit, les déséquilibres des paiements courants

s'expliquent par les déséquilibres au niveau domestique (interne) entre épargne et investissement.

Les flux financiers internationaux, destinés au financement de l'investissement et l'accumulation de capital, et les flux réels (de biens et services), sont comme les deux faces d'une même réalité.

Deux cas de figure sont le plus souvent observés (l'équilibre extérieur est davantage l'exception que la règle) : Excédent extérieur et Déficit extérieur.

Considérons les deux situations observables (excédents ou déficits) pour donner du sens au financement des économies ouvertes, donc des agents et des opérations économiques en économie ouverte.

17.2.3 *Excédent extérieur et excédent d'épargne*

$$\boxed{(X - M) = (S - I)}$$
$$> 0 > 0$$

Un excédent extérieur $(X - M) > 0$ aura nécessairement pour **contrepartie** une **épargne nette positive** $(S - I) > 0$.

Ceteris paribus, dans un pays où on observera un **excédent extérieur**, nous aurons également un **excédent de l'épargne** sur l'investissement. C'est donc un pays dans lequel les agents ont globalement une capacité de financement. Les agents résidents peuvent financer (prêts) des dépenses d'agents non résidents. Cette capacité de financement supplémentaire, relativement aux investissements domestiques, ira financer des besoins de financement dans des économies où les capacités de financement seront insuffisantes par rapport aux besoins. On observera, dans cette économie en excédent d'épargne (et en excédent courant), des **sorties nettes de capitaux**. *In fine*, les agents à capacité de financement du pays créateur, les créanciers-prêteurs, deviennent propriétaires d'une partie du rendement futur du capital du pays débiteur.

Un pays est en excédent extérieur $(X - M) > 0$ quand il vend plus (X) qu'il n'achète (M) au reste du monde. C'est un pays dans lequel on épargne (S) plus qu'on n'investit (I) : $(S - I) > 0$, et donc on ne consomme (C) pas assez. Notez que l'excédent extérieur s'explique aussi par un excédent d'épargne sur l'investissement, et non pas seulement par une grande compétitivité. **L'excédent extérieur n'est donc que le symptôme d'un déséquilibre interne, l'excédent d'épargne sur l'investissement** (l'insuffisance de consommation).

17.2.4 Déficit extérieur et déficit d'épargne

$$\boxed{(X - M) = (S - I)}$$
$$< 0 < 0$$

Un déficit extérieur $(X - M) < 0$ aura nécessairement pour contrepartie une épargne nette négative $(S - I) < 0$. *Ceteris paribus*, dans un pays où on observera un déficit extérieur, nous aurons également un déficit d'épargne par rapport à l'investissement. C'est donc un pays dans lequel les agents ont globalement un besoin de financement. Ce besoin de financement supplémentaire, relativement à l'épargne domestique, sera financé par des capacités de financement excédentaires provenant des économies où les besoins de financement seront insuffisants par rapport aux capacités. Les agents résidents doivent se faire financer par des agents non résidents. Concrètement, des résidents empruntent à des résidents, mais comme cela ne suffit pas, des résidents empruntent aussi à des non-résidents, car l'épargne des résidents est insuffisante.

On observera, dans cette économie en déficit d'épargne (et en déficit courant), des entrées nettes de capitaux. *In fine*, les agents à besoin de financement du pays débiteur, les débiteurs-emprunteurs, cèdent une partie de la propriété du rendement futur du capital de leur économie. Un pays est en déficit extérieur $(X - M) < 0$ quand il achète (M) plus qu'il ne vend (X) au reste du monde. C'est aussi un pays qui épargne

(S) moins qu'il n'investit (I) et donc qui consomme (C) trop. Notez que le déficit extérieur s'explique par un déficit d'épargne sur l'investissement et non pas seulement par une faible compétitivité. Le déficit extérieur n'est que le symptôme d'un déséquilibre interne, le déficit d'épargne par rapport à l'investissement, lié à un excès de consommation.

Un déficit extérieur ne peut exister que s'il est financé. En situation de déficit extérieur $(X - M) < 0$, les recettes issues des exportations ne permettent pas de payer les importations. Sans financement externe, les importations doivent baisser. C'est seulement l'existence d'un potentiel financement, provenant des non-résidents, qui permet de financer l'achat des produits importés, donc l'existence d'un déficit extérieur.

17.2.5 *Les déséquilibres entre économies ouvertes au sein d'une économie mondiale fermée*

Monde (économie fermée)	
Économie nationale (ouverte)	Reste du monde (autres économies ouvertes)
$(X - M) > 0$ et $(S - I) > 0$ => Prêts aux non-résidents : sorties de capitaux Flux de capitaux sortant de l'économie nationale vers le reste du Monde	$(X - M) < 0$ et $(S - I) < 0$ => Emprunts aux résidents de l'économie nationale : entrées de capitaux Flux de capitaux provenant de l'économie nationale
$(X - M) < 0$ et $(S - I) < 0$ => Emprunts aux non-résidents : entrées de capitaux Flux de capitaux provenant du reste du Monde et entrant dans l'économie nationale	$(X - M) > 0$ et $(S - I) > 0$ => Prêts aux résidents de l'économie nationale : sorties de capitaux Flux de capitaux sortants, vers l'économie nationale

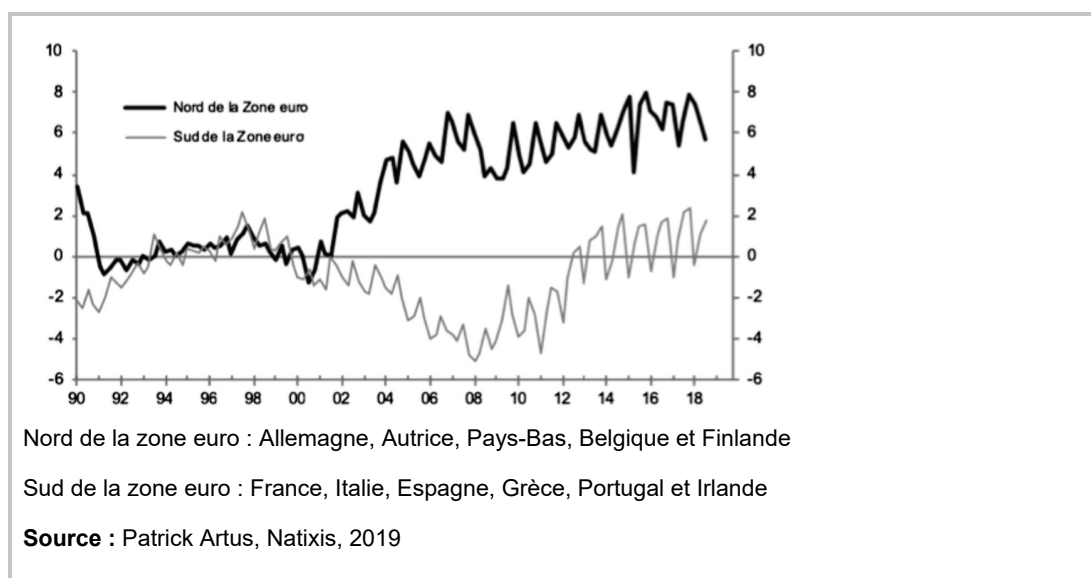
En résumé, dans les économies ouvertes, nous avons des agents résidents d'un pays qui ont une capacité de financement et qui vont prêter aux agents d'autres pays qui ont un besoin de financement. Voilà comment les excédents

extérieurs des uns financent les déficits extérieurs des autres. Étant donné que le monde est divisé en « deux pays » (l'économie nationale et le reste du monde) et que, de surcroît, le monde est une économie fermée (aux échanges avec les extraterrestres), si l'un des pays est en excédent extérieur, l'autre sera nécessairement en déficit extérieur.

17.2.6 *La zone euro et les déséquilibres macroéconomiques et financiers*

On peut observer cela dans le cas de la zone euro, au cours des années 2000. La zone euro étant une grande économie, et résultant des logiques de la construction économique européenne, elle est relativement plus fermée que la moyenne des économies beaucoup plus petites.

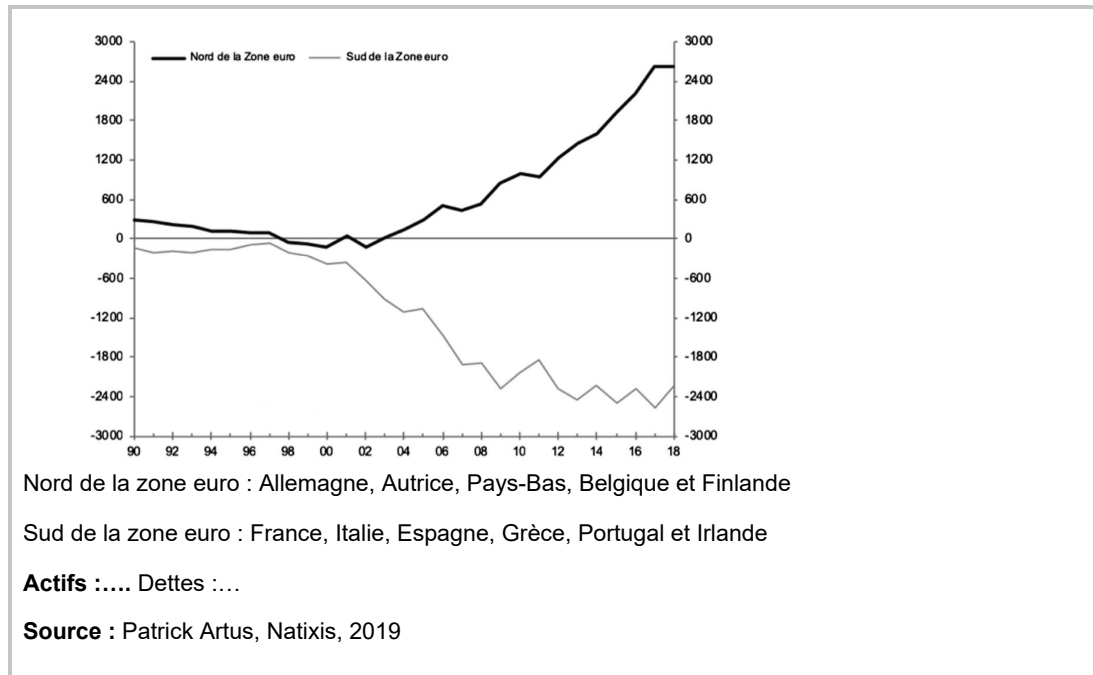
En effet, dans la zone euro, les capacités de financement des pays dits du Nord qui sont en excédents extérieurs [$(X - M) > 0$] notamment (Allemagne, Pays-Bas...) couvrent les besoins de financement des pays du Sud qui sont en déficit extérieur [$(X - M) < 0$] (Grèce, Portugal, Espagne).



GRAPHIQUE 17.1.

Zone euro : Balance courante en % du PIB en volume, 1990-2018

On observe clairement, dans le graphique ci-dessus, que les excédents courants des uns s'accompagnent, dans le graphique ci-dessous, de sorties de capitaux, et ont pour contreparties les déficits courants des autres, qui s'accompagnent d'entrées de capitaux.



GRAPHIQUE 17.2.

Actifs ou dettes nets extérieurs, 1990-2018 (en Mds d'€)

17.2.7 Les déséquilibres globaux représentent un enjeu majeur pour l'économie mondiale

Au cours des années 2000, on a observé le développement au sein de l'économie mondiale de déséquilibres qualifiés de « globaux ». Les déséquilibres globaux désignent des déséquilibres durables des échanges internationaux entre les économies ouvertes du monde. Il y a des déséquilibres dans les échanges internationaux de biens et services, mais aussi dans les échanges de capitaux entre les économies, c'est-à-dire entre les agents économiques résidents et non résidents. Cependant, derrière les déséquilibres financiers, il existe des déséquilibres au sein des économies « réelles » elles-mêmes

(inégaie répartition des revenus, polarisation de l'emploi, désindustrialisation, spécialisations internationales). Les déséquilibres extérieurs sont toujours la traduction de déséquilibres internes.

Plus largement, les déséquilibres globaux sont d'une part le reflet de la mondialisation des échanges commerciaux, financiers et monétaires, et d'autre part l'expression de multiples interdépendances.

Il y a consensus pour dire que ces déséquilibres globaux et leurs conséquences, en termes d'abondance de liquidité en particulier, ont permis les excès comportementaux des agents économiques qui ont conduit à la crise financière qui s'est déclarée en 2007 et qui s'est propagée dans le monde depuis 2008 (voir chapitre 20).

De surcroît, si, médiatiquement, l'attention a été attirée sur les défaillances des comportements au niveau microéconomique, en revanche, ces déséquilibres globaux (d'ordre macroéconomique) et leurs conséquences furent largement sous-estimés, voire ignorés.

Pourtant, on ne peut espérer comprendre l'enchaînement de ces crises financières économiques et des finances publiques sans avoir une connaissance minimum des flux économiques, monétaires et financiers, de leur traduction et conséquences illustrées par ces déséquilibres globaux. La compréhension des ressorts microéconomiques de la crise qui a débuté en 2007 doit absolument s'accompagner d'une meilleure compréhension des causes macroéconomiques et globales de celle-ci, si nous voulons vraiment, d'une part, résoudre les problèmes posés par les crises en cours et, d'autre part, éviter de créer, en « soignant » le présent, les conditions d'occurrence d'une crise encore plus grave dans l'avenir.

Déséquilibres globaux, interdépendances et coopération

En effet, le premier intérêt des problématiques relatives aux déséquilibres globaux est de mettre en évidence de nombreuses interdépendances et la nécessaire coopération qui devra en découler pour pouvoir résoudre les crises

présentes et les crises de l'avenir. Nous pouvons par exemple retenir les interdépendances suivantes : entre économies nationales, entre opérations réelles, monétaires et financières, c'est-à-dire entre flux internationaux de capitaux, de biens et de services, de devises, entre agents économiques privés et agents publics, entre politiques macroéconomiques et structurelles des différents pays, entre politiques monétaires, entre politiques budgétaires, entre politiques budgétaires et politiques monétaires, entre formes de liquidité...

Analyser les problématiques relatives aux déséquilibres globaux permet de comprendre les causes macroéconomiques de la crise financière de 2007-2008, bien au-delà des causes microéconomiques et des questions liées à l'architecture financière internationale. On peut alors prendre conscience du fait que les réformes structurelles de nature microéconomique doivent être complétées par des réformes au niveau macroéconomique et par davantage de coopération entre les nations. Il n'y aura pas de retour à une croissance soutenue et durable, soutenable, sans résolution simultanée des problèmes microéconomiques et des déséquilibres macroéconomiques.

Cela étant, pour parvenir à comprendre la nature et les origines de ces déséquilibres globaux, nous reprendrons l'identité fondamentale de la macroéconomie des économies ouvertes déjà présentée plus haut.

17.3 LES DÉSÉQUILIBRES MACROÉCONOMIQUES ET FINANCIERS AVEC DISTINCTION AGENTS PRIVÉS/AGENTS PUBLICS

17.3.1 Équation des déséquilibres dans des économies ouvertes avec distinction entre agents privés et agents publics

En économie ouverte, quand **on fait la distinction entre les agents publics et les agents privés**, l'Équilibre Ressources-Emplois en biens et services devient :

$$Y + M = C + I + X + G$$

$$Y = C + I + G + X - M \quad (1)$$

Cette **identité** signifie que les biens et services à notre disposition (ressources) ont deux origines possibles : notre propre production ($Y = \text{PIB}$) et nos importations (M).

Ils sont utilisés (emplois) pour la consommation (C), l'investissement (I), les dépenses publiques (G) et les exportations (X)

$$Y = C + I + G + (X - M) \quad (2)$$

Le PIB, noté ici Y , a donc pour contrepartie les différentes composantes de la demande globale avec :

$(C + I + G)$, qui correspond à la demande intérieure ou absorption

$(X - M)$, qui correspond à la demande extérieure nette

Cette présentation va nous permettre de mettre en lumière les grands équilibres/déséquilibres macroéconomiques en faisant le lien avec le modèle offre globale (ou agrégée) et demande globale (ou agrégée). Ce modèle offre globale (OG)/demande globale (DG) est le modèle économique de base approprié pour l'analyse macroéconomique des grands équilibres et déséquilibres de l'économie, considérée dans son ensemble.

Ressources en biens et services = Emplois de ces biens et services

$$Y + M = C + I + G + X \quad (3)$$

$$\text{OG} = \text{DG}$$

17.3.2 Équilibre ressources-emplois dans une économie ouverte et interdépendances

On peut simplement mettre en évidence les interdépendances multidimensionnelles (sur les échanges de

biens et services, sur les échanges de capitaux et sur les échanges de devises) en économie ouverte.

Quand on fait la distinction entre les agents privés et les agents publics, les équations changent un peu.

L'Équilibre Ressources = Emplois en biens et services pour une économie ouverte devient :

$$Y + M = C + I + X + G \Rightarrow Y = C + I + G + X - M \quad (1)$$

Cette identité signifie que les biens et services à notre disposition ont deux origines possibles : notre propre production ($Y = \text{PIB}$) et nos importations (M). Ils sont utilisés pour la consommation (C), l'investissement (I), les dépenses publiques (G) et les exportations (X).

Selon une autre approche, par le revenu et son utilisation, on a aussi :

$$Y = C + S + T \quad (2) \text{ avec } T, \text{ les recettes fiscales}$$

En combinant (1) et (2), on obtient

$$Y = C + I + G + X - M \text{ et } Y = C + S + T$$

donc

$$C + I + G + X - M = C + S + T$$

et

$$(X - M) = (S - I) + (T - G)$$

Cette identité est une **égalité comptable** qui est toujours réalisée, du fait de sa construction.

À partir de cette simple identité, on peut comprendre pourquoi les déséquilibres extérieurs, observés en utilisant les balances des paiements courants (transactions courantes) des différentes économies, ont pour contrepartie des déséquilibres dans les flux de capitaux qui les financent.

Autrement dit, les déséquilibres des paiements courants s'expliquent d'un côté par les déséquilibres au niveau domestique (interne) entre épargne et investissement des agents économiques privés et d'un autre côté par les déséquilibres au niveau du budget de l'État (des administrations publiques au sens large).

TABLEAU 17.1.

Déséquilibres économiques et financiers internationaux

Cas 1 : Excédent extérieur et excédent d'épargne	Cas 2 : Déficit extérieur et déficit d'épargne
$(X - M) = (S - I) + (T - G)$ $> 0 > 0 > 0$ Ou $(X - M) = (S - I) + (T - G)$ $> 0 > 0 < 0$ Ou $(X - M) = (S - I) + (T - G)$ $> 0 < 0 > 0$ L'Excédent extérieur $(X - M) > 0$ aura pour contrepartie une Épargne nette positive du secteur privé $(S - I) > 0$ et/ou Épargne nette positive du secteur public $(T - G) > 0$. On a donc ici des Sorties nettes de capitaux. <i>Ceteris paribus</i>, pour un solde (épargne – investissement) équilibré pour le secteur privé, dans un pays où on observe un excédent extérieur, nous avons toutes les chances de constater un excédent budgétaire au niveau de l'État.	$(X - M) = (S - I) + (T - G)$ $< 0 < 0 < 0$ Ou $(X - M) = (S - I) + (T - G)$ $< 0 < 0 > 0$ Ou $(X - M) = (S - I) + (T - G)$ $< 0 > 0 < 0$ Déficit extérieur $(X - M) < 0$ aura pour contrepartie une Épargne nette négative du secteur privé $(S - I) < 0$ et/ou Épargne nette négative du secteur public $(T - G) < 0$. On a donc là des Entrées nettes de capitaux. <i>Ceteris paribus</i>, pour un solde (épargne – investissement) équilibré pour le secteur privé, dans un pays où on observe un déficit extérieur, nous avons toutes les chances de constater un déficit budgétaire au niveau de l'État.

L'équilibre extérieur étant davantage l'exception que la règle, on peut donc rencontrer les deux types de situations suivant :

- l'excédent extérieur avec excédent d'épargne
- le déficit extérieur avec déficit d'épargne

Cette fois-ci, le solde extérieur $(X - M)$ a pour contrepartie la somme des épargnes nettes des agents privés et des agents publics.

Par exemple :

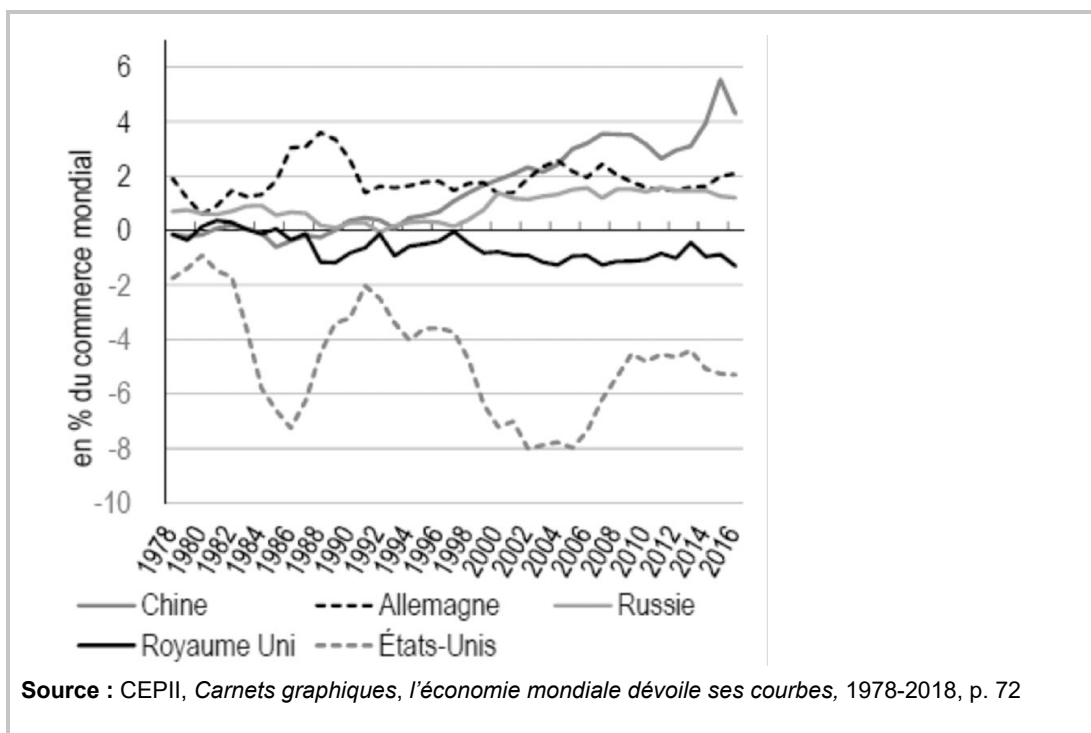
- Un excédent extérieur $(X - M) > 0$ aura pour contrepartie une épargne nette positive du secteur privé $(S - I) > 0$ et/ou une épargne nette positive du secteur public $(T - G) > 0$.
- Un déficit extérieur $(X - M) < 0$ aura pour contrepartie une épargne nette négative du secteur privé $(S - I) < 0$ et/ou

une épargne nette négative du secteur public ($T + G$) < 0.

17.3.3 Les déséquilibres globaux au cours des années 2000

Les déséquilibres commerciaux entre les États-Unis et la Chine

Dans l'économie mondiale des années 2000, nous pouvons nous intéresser aux échanges entre la Chine et les États-Unis et observer que la Chine ayant un excédent courant ($X - M$) > 0, elle vend (X) plus qu'elle n'achète (M), car elle dépense peu en consommation (C), mais épargne (S) beaucoup par rapport à son investissement.



GRAPHIQUE 17.3.

Déséquilibres commerciaux entre 1978 et 2016

La Chine a donc une capacité de financement. Les résidents chinois sont « obligés » d'épargner car il n'y a pas de protection sociale... Dans ces conditions, une épargne de précaution se révèle indispensable. De plus, les résidents

chinois peuvent préfèrent prêter à des non-résidents américains, d'autant plus que le système financier chinois étant encore peu développé, il offre peu d'opportunités de placements financiers. Au contraire, le marché financier américain est beaucoup plus profond, mieux structuré, organisé et *in fine*, bien plus liquide, donc plus sécurisant pour les investisseurs.

A *contrario*, les États-Unis ont un déficit courant $(X - M) < 0$. Les résidents américains achètent (M) plus qu'ils ne vendent (X) et donc épargnent (S) peu par rapport à leur investissement. En effet, ils consomment plus que leurs revenus, donc ils doivent emprunter.

Dans cette logique, les Chinois, qui consomment peu, peuvent prêter aux Américains, qui consomment beaucoup, et finalement leur prêter de quoi acheter les produits chinois qu'ils ne peuvent acheter eux-mêmes. Résultat, les Américains sont très endettés vis-à-vis des Chinois ! Il existe donc de grandes interdépendances économiques et financières entre la Chine et les États-Unis.

Dans l'exemple États-Unis/Chine des années 2000, nous avons donc :

$$\begin{aligned}\text{En Chine } (X - M) &= (S - I) + (T + G) \\ &> 0 > 0 > 0\end{aligned}$$

Aux États-Unis, on aurait :

$$\begin{aligned}(X - M) &= (S - I) + (T - G) \\ &< 0 < 0 < 0\end{aligned}$$

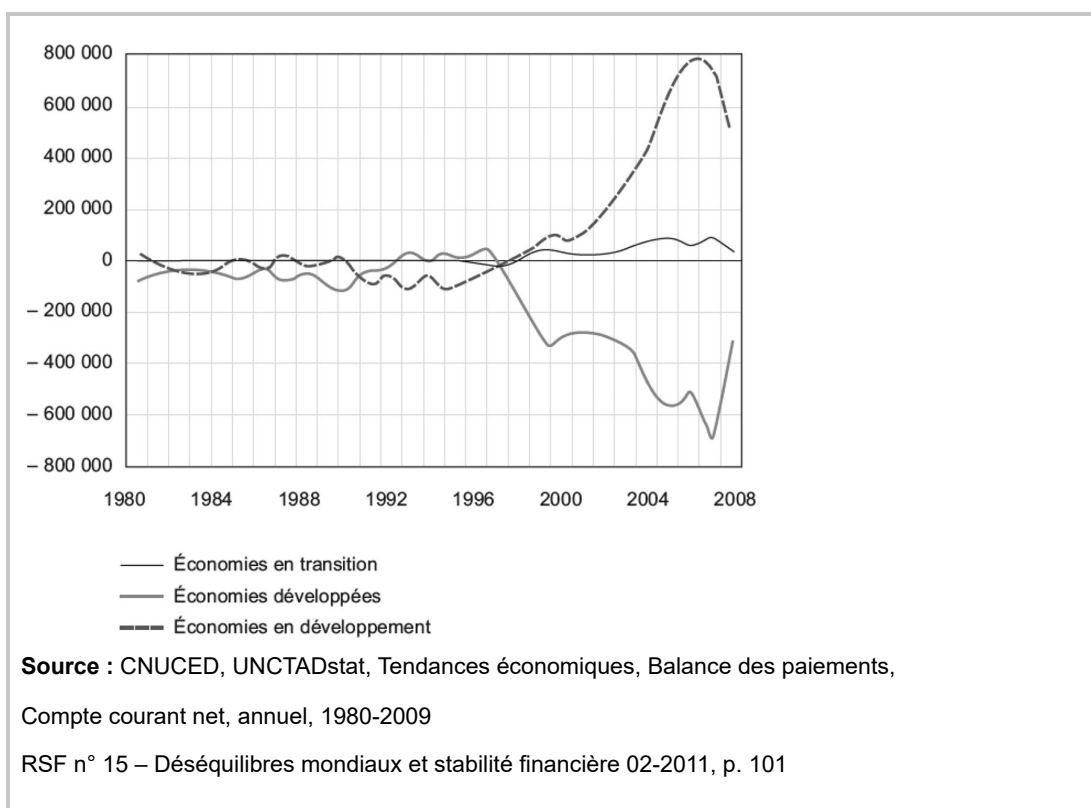
On en déduit que les agents privés et publics chinois vont « prêter » (transférer des capitaux) aux agents privés et publics américains. Ainsi, les Chinois deviennent doublement dépendants des Américains, par leurs exportations et leurs prêts.

Il existe donc des liens très étroits entre déficits extérieurs (endettement vis-à-vis du reste du monde), d'une part, et déficits budgétaires (endettement public), excès de

consommation et insuffisance d'épargne (endettement privé), d'autre part.

Les déséquilibres courants entre les économies développées et les économies en développement

On peut observer que la croissance de l'excédent courant des économies en développement et, dans une moindre mesure, celle des économies en transition, a pour contrepartie le creusement du déficit courant des économies développées. Ainsi, depuis la fin des années 90, et paradoxalement, ce sont les pays les résidents des pays les moins développés qui « prêtent » aux résidents des pays développés pour qu'ils puissent importer plus qu'ils n'exportent. Autrement dit, le manque d'investissement (relativement à l'épargne) des premiers permet de financer l'excès d'investissement (et de consommation) des seconds.



GRAPHIQUE 17.4.

Soldes des transactions courantes des économies développées et économies en développement entre 1980 et 2008 (en milliards de dollars)

17.3.4 *La compétitivité au regard des déséquilibres macroéconomiques et financiers*

Un déficit extérieur n'est pas un problème en soi et un excédent extérieur n'est pas, en soi, un signe de compétitivité. C'est pour cela qu'il convient de regarder quelles en sont les explications en matière d'épargne et d'investissement.

Un déficit extérieur pourrait être le signe d'un développement économique rapide de pays en voie de développement ayant des besoins de financement d'investissements très importants. En effet, un déficit extérieur a pour contrepartie un endettement extérieur qui devra être remboursé par les générations futures à partir de la croissance économique future. Cette identité des déséquilibres en économie ouverte nous donne des informations cruciales, explicitement (directement) et implicitement (indirectement), sur les déséquilibres globaux, leur nature et leurs origines profondes.

Selon Paul Krugman (Prix Nobel en 2008 pour ses travaux en économie internationale), « *la compétitivité est un mot vide de sens lorsqu'il est appliqué aux économies nationales* » et « *un problème de compétitivité d'une économie nationale peut être purement et simplement un problème intérieur de productivité* ».

Pourtant, cette approche macroéconomique, globale, de la compétitivité marque les travaux et les décisions politiques internationales, en particulier en Europe. Un pays ayant un déficit commercial n'est pas nécessairement en mauvaise santé. Un déficit commercial (ou courant) n'est pas forcément le signe de dysfonctionnement de l'économie et d'un manque de compétitivité. « *Les déficits et excédents commerciaux sont des phénomènes macroéconomiques [...] la réalité est qu'il n'y a pas de relation simple entre la réussite d'une économie et sa situation d'excédent ou de déficit commercial.* » (Selon Paul Krugman).

De même, un excédent commercial (ou courant) n'est pas forcément le signe d'un fonctionnement optimal de l'économie et d'une solide compétitivité.

L'analyse microéconomique nous explique pourquoi dans les pays les agents économiques résidents se spécialisent et échangent (avantages comparatifs, dotations factorielles...) et non pourquoi les économies nationales sont en situation d'excédent ou de déficit commercial. Les déséquilibres extérieurs (commercial, courant) dépendent donc des décisions d'épargne et d'investissement des agents économiques (privés et publics) des pays qui échangent.

L'excédent extérieur (qui semble être un signe de compétitivité et de bonne santé économique) est lié à un excédent d'épargne, c'est-à-dire un manque d'investissement (ce qui n'est pas bon signe à long terme) et une insuffisance de consommation (ce qui n'est pas une bonne nouvelle à court terme). Le déficit extérieur (qui semble être un signe de faible compétitivité et de mauvaise santé économique) est lié à un excédent d'investissement (ce qui n'est pas mauvais signe à long terme) et une consommation élevée (ce qui n'est pas forcément une mauvaise nouvelle à court terme).

De surcroît, une économie qui possède un excédent d'épargne est une économie dans laquelle les agents résidents disposent d'une capacité de financement qui servira à financer les besoins de financement des agents économiques des économies qui sont en déficit extérieur, en manque d'épargne et de capacité de financement.

À cet aune, la question de la compétitivité extérieure mesurée par le solde extérieur prend un tout autre sens.

17.3.5 Politiques économiques et soldes extérieurs courants : enseignements fondamentaux

L'identité des déséquilibres en économie ouverte peut nous permettre d'estimer l'impact des politiques économiques sur

la balance des paiements courants en présentant leurs conséquences sur l'épargne et l'investissement.

Les politiques économiques qui contribuent à faire croître l'investissement et à diminuer l'épargne entraînent un déficit extérieur, alors que les politiques économiques qui contribuent à faire baisser l'investissement et à augmenter l'épargne entraînent un excédent extérieur !

Les comportements d'endettement, de consommation, d'épargne, de placement de cette épargne et d'investissement des agents économiques privés sont influencés par les politiques monétaires menées par les banques centrales, lesquelles, en faisant varier les taux d'intérêt de court terme (et les taux d'intérêt à long terme via les anticipations réalisées par les agents économiques), vont agir sur l'endettement et donc la consommation, l'épargne, l'investissement.

Ces comportements sont aussi influencés par les politiques budgétaires (dépenses publiques et fiscalité), domestiques et étrangères, qui vont agir sur la dépense privée et les taux d'intérêt à long terme.

La combinaison de politiques monétaires excessivement expansionnistes dans les pays de l'OCDE et de stratégies de stabilité du taux de change des autorités monétaires dans les économies émergentes et exportatrices de pétrole se traduit par une accumulation de réserves de changes des banques centrales de ces pays... lesquelles ont contribué à faire durablement croître la base monétaire mondiale.

Cet excès de liquidité mondiale, associé à des taux d'intérêt durablement très bas, a permis d'alimenter l'endettement privé avant 2007, ce qui a permis de doper la croissance économique, d'alimenter des bulles financières et de faciliter l'endettement public. L'interdépendance qui existe entre les politiques monétaires et budgétaires a donc contribué à alimenter la liquidité mondiale, l'endettement privé et public, avant la crise de 2008.

Manque de coordination des politiques macroéconomiques et déséquilibres globaux

Les « policy mix » (combinaison des politiques budgétaires et monétaires) non coordonnés permettent d'expliquer, au moins en partie, les déséquilibres $(S - I) + (T - G)$ et donc $(X - M)$.

Le manque de coopération en présence d'interdépendances multiples est très néfaste. Les déséquilibres internes (consommation, investissement, épargne, dépenses publiques, fiscalité) permettent d'expliquer les déséquilibres externes. La combinaison de ces deux problèmes se traduit par un excès d'endettement et une forte croissance de la liquidité mondiale, qui a elle-même contribué aux bulles financières qui ont déclenché la crise de 2008.

Cette crise, qui est fondamentalement une crise de la dette privée, a provoqué quelques années plus tard une crise des dettes publiques. Ces crises de dettes publiques contribuent à maintenir des déséquilibres globaux et elles en sont indirectement et partiellement à l'origine.

Cette identité nous montre enfin qu'il n'est pas possible de résoudre durablement les crises des dettes publiques sans réduire les déséquilibres $(X - M)$ et $(S - I)$.

Donc, sans coopération, la sortie de crise peut se révéler compliquée !

A cet aune, on peut commencer à savoir comment traiter les causes profondes des crises, de différentes natures, mais interdépendantes, qui se cumulent depuis juillet 2007.

Dans les économies ouvertes, nous voyons que des agents privés ou publics d'un pays en excédent extérieur et d'épargne ont une capacité de financement et qu'ils vont prêter... aux agents privés ou publics d'autres pays, qui ont des besoins de financement. L'interdépendance entre flux réels et flux financiers apparaît déjà clairement. Comme, en outre, les échanges réels et financiers se font en monnaie et devises, il ne faut pas être étonné de voir que les déséquilibres de paiements courants (et donc d'épargne-investissement) se traduisent par des mouvements des taux

de change, lesquels en retour vont rétroagir sur ces flux réels et financiers. Les phénomènes réels et financiers sont donc interdépendants des phénomènes monétaires.

17.3.6 *Les politiques économiques face aux équilibres et déséquilibres macroéconomiques*

Face à ces situations de déséquilibres permanents, les politiques économiques entrent en action afin de faciliter le retour vers l'équilibre de l'économie : stabiliser le chômage et l'inflation à des niveaux bas, maintenir la croissance à un niveau soutenu sans tensions inflationnistes et sans déséquilibres extérieurs trop importants. La main visible des pouvoirs publics va tenter d'accompagner la main invisible des marchés pour qu'*in fine* il y ait convergence entre les intérêts individuels d'une part et les intérêts individuels et collectifs de l'autre.

Retenons d'abord que ces grands équilibres/déséquilibres macroéconomiques se retrouvent dans les principaux objectifs des politiques économiques conjoncturelles, dits du « carré magique ».

Le « carré magique » a été inventé par Nicolas Kaldor en 1971. C'est un graphique qui représente sur quatre axes les quatre principaux objectifs des politiques macroéconomiques conjoncturelles. Ces objectifs sont une croissance économique soutenue et durable, le plein emploi, la stabilité des prix et l'équilibre extérieur. Les objectifs sont interdépendants. Le carré des objectifs de la politique macroéconomique conjoncturelle est dit magique, car il serait magique d'atteindre les quatre objectifs simultanément. Ils constituent les principaux indicateurs du tableau de bord de l'économie globale.

Ces différents déséquilibres macroéconomiques peuvent être soit la cause, soit la conséquence de déséquilibres de la demande globale, de l'offre globale ou de leurs composantes. Ces trois déséquilibres sont liés à la demande globale. En

effet, l'inflation pénalise la consommation finale, l'investissement et le commerce. Le chômage diminue la consommation finale et l'investissement, et les déséquilibres extérieurs pénalisent la production domestique et accroissent le chômage, ce qui provoque une diminution du PIB. Quand un pays est en excédent extérieur, il y aura un ou plusieurs autres pays en déficit extérieur. Par exemple, quand l'Allemagne est en excédent courant par rapport au reste de l'Europe, d'autres pays sont logiquement en déficit courant. Les composantes de l'offre et de la demande globale des économies varient en sens contraire. Ces quatre grands déséquilibres doivent donc être combattus, en particulier par les politiques macroéconomiques conjoncturelles, qui vont agir essentiellement sur les différentes composantes de la demande globale, mais également par les politiques structurelles, qui vont agir spécifiquement sur les différentes composantes de l'offre globale.

PRINCIPES DE POLITIQUE ÉCONOMIQUE

« Laisser faire ne doit pas signifier ne rien faire, mais laisser agir la libre concurrence. Là donc où la libre concurrence ne peut agir, il y a lieu, pour l'État, d'intervenir afin d'y suppléer ; et là où la libre concurrence peut agir, il y a lieu, pour l'État, à intervenir afin de l'organiser, d'assurer les conditions et le milieu de son fonctionnement. » - Léon Walras, 1880

« Nous vivons aujourd'hui la plus grande catastrophe économique du monde moderne - la plus grande catastrophe due presque entièrement à des causes économiques. On me dit qu'à Moscou, l'on considère que ceci est la dernière et la suprême crise du capitalisme, et que l'ordre social existant n'y survivra pas. On croit aisément ce qu'on désire. Mais si cette crise se trouve plus tard analysée par l'historien économique, il y a, c'est vrai, une possibilité - je n'irai pas plus loin - qu'elle soit interprétée comme un tournant majeur [...] Pas plus qu'une possibilité, cependant. Car je crois profondément que notre destin est encore entre nos mains et que nous pouvons sortir de cette crise si nous effectuons les choix adéquats - ou plutôt si ceux qui détiennent l'autorité de par le monde le font. » - John Maynard Keynes, 1931

« Nous devrions en savoir assez long pour éviter de détruire notre civilisation en étouffant le processus spontané de l'interaction des individus, en chargeant une quelconque autorité de la diriger. Mais pour ne pas tomber dans cette faute, nous devons rejeter l'illusion d'être capables de délibérément "créer l'avenir de l'humanité". » - Friedrich Hayek, 1979

« Bref, une politique budgétaire expansionniste constitue l'un des rares moyens de sortir d'une récession. » - Joseph Stiglitz, 2002

« C'est la fonction même de la politique macroéconomique : fournir aux habitants du pays une assurance collective d'activité. » - Jean-Paul Fitoussi, 2005

SOMMAIRE

[18.1 Les fondements de la politique économique](#)

[18.2 Les objectifs de la politique économique](#)

[18.3 Les instruments de la politique économique](#)

[18.4 Les contraintes qui limitent les effets de la politique budgétaire](#)

[18.5 Le « policy-mix » et le pilotage conjoncturel](#)

[18.6 Les politiques économiques structurelles](#)

18.1 LES FONDEMENTS DE LA POLITIQUE ÉCONOMIQUE

L'instabilité de la croissance économique, ici dans la zone euro, et ses nombreuses conséquences sont inhérentes aux économies capitalistes de marché. Le retour à davantage de stabilité ne peut se faire en fonction des seuls mécanismes de marché. L'intervention des pouvoirs publics est indispensable.

18.1.1 *La justification de la politique économique*

La **politique économique** désigne l'étude des processus par lesquels les pouvoirs publics orientent, par leurs décisions, les choix et l'activité des agents économiques privés (ménages et entreprises), dans une direction jugée souhaitable pour l'ensemble d'un territoire (région, nation, ou union plurinationale comme la zone euro). La politique économique utilise principalement des instruments de nature financière, et son action transite principalement par la demande globale (c'est-à-dire la consommation et l'investissement). Au XVIII^e siècle, les premiers penseurs de l'économie ont surtout insisté sur les vertus du marché et du libéralisme économique pour organiser les échanges. Dans ce cadre d'analyse, le jeu spontané de l'intérêt individuel et les bienfaits de la concurrence suffisent à assurer le bien-être collectif : c'est aussi le sens de la célèbre métaphore de la « main invisible » d'Adam Smith, dans son ouvrage intitulé *Recherches sur la nature et les causes de la Richesse des nations* publié en 1776, qui défend cette harmonie naturelle des intérêts, où l'État se désengage de la sphère économique. Le meilleur gouvernement est alors celui qui ne perturbe pas l'ordre économique spontané des marchés.

Mais avec la survenue de crises industrielles récurrentes au XIX^e siècle, et leur cortège de chômage, de pauvreté, et de déséquilibres monétaires (inflation ou déflation), l'idée d'une intervention correctrice de l'État s'est progressivement imposée. Au XX^e siècle, c'est en particulier à l'occasion de la grande crise des années 1930 (la « Grande Dépression ») que l'idée d'une politique économique stabilisatrice a montré son utilité, tandis que les analyses de John Maynard Keynes (1883-1946), à la même époque, ont donné aux politiques de régulation conjoncturelle et de relance une assise scientifique. Aux fonctions régaliennes de l'État (police, justice, défense nationale), se sont ainsi ajoutées, au XX^e siècle, des politiques de régulation conjoncturelle ; des politiques structurelles avec des objectifs de long terme (qui peuvent tout à la fois concerner la compétitivité, le tissu industriel, la recherche-développement, l'innovation, la cohésion sociale) ; des politiques microéconomiques centrées sur des secteurs particuliers de l'économie (comme l'aide au secteur automobile ou à l'aéronautique, par exemple) ; des politiques sociales de protection et de redistribution ; et enfin d'éventuelles fonctions de production à travers les nationalisations de certaines industries ou de certains services publics (voir chapitre 7).

Pour envisager la théorie de la politique économique, on peut puiser, dans une première approche, trois concepts structurants dans la boîte à outils des économistes : l'optimalité, l'efficacité et la crédibilité.

- L'**optimalité** se base sur l'idée selon laquelle, à l'image du consommateur ou du producteur, les choix de l'État en tant qu'acteur de la politique économique peuvent être optimisés par une **maximisation sous contrainte**, ce qui revient à chercher à affecter des ressources de la manière la plus productive possible.

- L'**efficacité**, préoccupation essentielle de tous les décideurs de politique économique, repose quant à elle sur la capacité d'obtention d'un résultat recherché par l'État. On retient généralement **deux règles** de politique

économique : la règle de **cohérence** de Jan Tinbergen établie en 1963 indique que, dans une politique économique à objectifs fixés, le nombre d'instruments indépendants utilisés doit être égal au nombre d'objectifs à atteindre (il faut mobiliser autant d'instruments que d'objectifs). La règle d'**efficience** de Robert Mundell (énoncée en 1960) implique que les instruments indépendants de la politique économique doivent être spécialisés selon les objectifs poursuivis, en fonction de leur efficacité comparative (il faut affecter à un objectif l'instrument relativement le plus efficace).

– Enfin, dans une économie mondialisée, où les biens et les capitaux circulent librement, la contrainte de **crédibilité** prend une importance considérable pour les pouvoirs publics, en particulier pour **ancrer les anticipations** des agents économiques privés dont les décisions sont influencées par les choix de politique économique.

18.1.2 *Les acteurs de la politique économique*

Les **marchés** ne peuvent pas fonctionner parfaitement : cela ne signifie pas forcément que l'État doit se substituer au marché, mais qu'une intervention publique est nécessaire pour corriger ces imperfections et assurer le fonctionnement concurrentiel des marchés. La conduite de la politique économique est censée permettre un fonctionnement plus harmonieux de l'économie de marché.

L'intervention de l'État dans l'économie s'effectue au titre de **trois fonctions principales** : l'**allocation** des ressources, la **stabilisation** du cycle économique et la **redistribution** des richesses pour corriger les inégalités. Comme les mesures de politique économique doivent être cohérentes et coordonnées entre elles, il est nécessaire qu'une seule autorité décide : c'est donc le gouvernement qui, avec l'aide des ministères et plus particulièrement du ministère de l'Économie, prend les mesures de politique économique. Dans certains pays comme les États-Unis ou dans la zone euro, certaines décisions importantes sont prises par des institutions

indépendantes des États, comme la banque centrale (la Banque Centrale européenne dans la zone euro).

L'**État** est donc le principal acteur de la politique économique : tous les systèmes économiques que nous connaissons aujourd'hui sont des systèmes mixtes combinant le mécanisme du marché et l'intervention de la puissance publique.

La notion de politique économique est toutefois trompeuse. Elle suggère une impulsion unique, celle de l'État ou plus encore le gouvernement qui en incarne la volonté. Or, si l'État est bien un centre d'impulsion, il n'est ni omnipotent ni omniscient : la politique économique résulte d'un **compromis politique** entre des acteurs aux intérêts divergents (par exemple, elle découle souvent de négociations dans le cadre de la démocratie sociale). Il faut rappeler quels sont les différents acteurs qui participent à l'élaboration de ce compromis, tandis que la politique économique est soumise à un certain nombre de contraintes (règles supranationales, contrainte extérieure dans l'économie globalisée, etc.).

Les **groupes de pression** interviennent tout au long du processus de décision des acteurs publics pour les influencer par l'information qu'ils leur fourniront, les débats ou forums qu'ils susciteront, et les manifestations qu'ils organiseront. Il existe de nombreux groupes de pression qui défendent des intérêts catégoriels (consommateurs, travailleurs, fonctionnaires, contribuables, etc.). Si on peut voir dans l'existence des groupes de pression une source de pluralisme, ou le moyen de faire entendre les moins bien placés, plusieurs critiques continuent de peser fortement sur leur rôle. Certains insistent sur l'exacerbation des corporatismes qui rendrait les choix plus difficiles et probablement moins soucieux d'un intérêt collectif. D'autres observateurs considèrent que les groupes de pression se préoccupent plus de partager le revenu national que de l'augmenter, critique d'autant plus forte que la situation économique est défavorable. Enfin, la multiplication des groupes de pression serait une source de conservatisme et

de statu quo plutôt que de progrès. Lorsque l'on mène des réformes, il est toujours plus facile de faire un front de ceux qui sont mécontents que de ceux qui pourraient être satisfaits dès lors qu'il s'agit de modifier des intérêts acquis au profit d'intérêts potentiels et aléatoires.

Dans les démocraties pluralistes, ce sont a priori **les préférences des électeurs** qui décident des politiques économiques. Mais la réalité est plus complexe : lorsqu'ils sont appelés à voter, les électeurs doivent trancher sur des programmes dont le contenu est constitué d'un paquet de mesures, où les choix de politique économique sont souvent insérés dans des propositions qui dépassent les seuls enjeux économiques. De plus, les électeurs se prononcent à intervalles fixes, et sauf cas exceptionnels, ce sont donc les gouvernements qui, durant leur mandat, prennent des décisions courantes concernant la politique économique. Par ailleurs, les décisions prises durant l'exercice du pouvoir peuvent différer des programmes qui ont permis l'élection, au gré des circonstances, de la conjoncture et des choix politiques. Toutefois, les gouvernements ont intérêt à respecter de très près les objectifs de ceux qui les ont choisis et/ou élus s'ils entendent eux-mêmes viser une reconduction et/ou réélection.

Enfin, les **conseillers économiques** et les centres de recherche en économie jouent également un rôle en livrant une expertise aux décideurs de la politique économique (Insee, CAE, OFCE, CEPIL, COE, BDF²⁰).

18.2 LES OBJECTIFS DE LA POLITIQUE ÉCONOMIQUE

18.2.1 *Les différents types de politiques économiques*

L'analyse économique permet de connaître les effets des décisions de l'État. Le travail de l'économiste est alors d'explicitier aux décideurs les implications de tel ou tel choix

en termes d'efficacité (la capacité à atteindre l'objectif souhaité), mais aussi d'équité (veiller aux effets de redistribution des revenus entre les groupes sociaux). On peut prendre l'exemple d'un gouvernement qui veut réduire le taux de chômage en supposant que deux politiques conduisent à ce résultat, mais au prix pour l'une d'une moindre progression des salaires, et de l'autre d'une plus forte augmentation des inégalités. Dans cet exemple, pour arbitrer entre ces deux solutions, il faut pouvoir comparer leur coût social respectifs et leur bénéfice social spécifiques. Les décideurs publics, souvent conseillés par les économistes, doivent donc arbitrer, compte tenu des différents acteurs en présence (qui ont des intérêts différents), en faveur de la mesure qui présente le meilleur rapport coût budgétaire/efficacité, tout en étant cohérente avec les options du gouvernement et les annonces déjà faites ou promises.

Dans ces conditions, il faut distinguer les différents objectifs de la politique économique. Aux fonctions incompressibles de l'État, que l'on appelle les fonctions régaliennes (police, justice, armée) se sont historiquement ajoutées :

- des politiques **structurelles** avec des objectifs de long terme, qui peuvent concerner la compétitivité, l'élévation de la croissance potentielle, le tissu industriel, la recherche-développement, l'innovation, la cohésion sociale ;
- des politiques **microéconomiques** centrées sur des secteurs particuliers, notamment pour aider les secteurs stratégiques qui ont des effets d'entraînement sur les autres secteurs de l'économie (l'automobile ou l'aéronautique, par exemple) ;
- des politiques **sociales** de protection et de redistribution ;
- d'éventuelles fonctions de **production** à travers les nationalisations de certaines industries ou de certains services publics ;

– des politiques **conjoncturelles** visant à amortir l'ampleur des cycles économiques.

18.2.2 *Une pluralité d'objectifs pour la politique économique conjoncturelle*

On retient généralement **quatre objectifs** économiques des politiques conjoncturelles que nous allons privilégier ici : le plein emploi, le soutien de l'activité économique et la croissance, l'équilibre extérieur et la stabilité des prix. L'idéal serait d'atteindre simultanément les quatre objectifs, lesquels sont interdépendants. La politique conjoncturelle vise les grands équilibres macroéconomiques à court terme, qui sont résumés dans le « carré magique » de l'économiste Nicholas Kaldor (1939).

La croissance économique

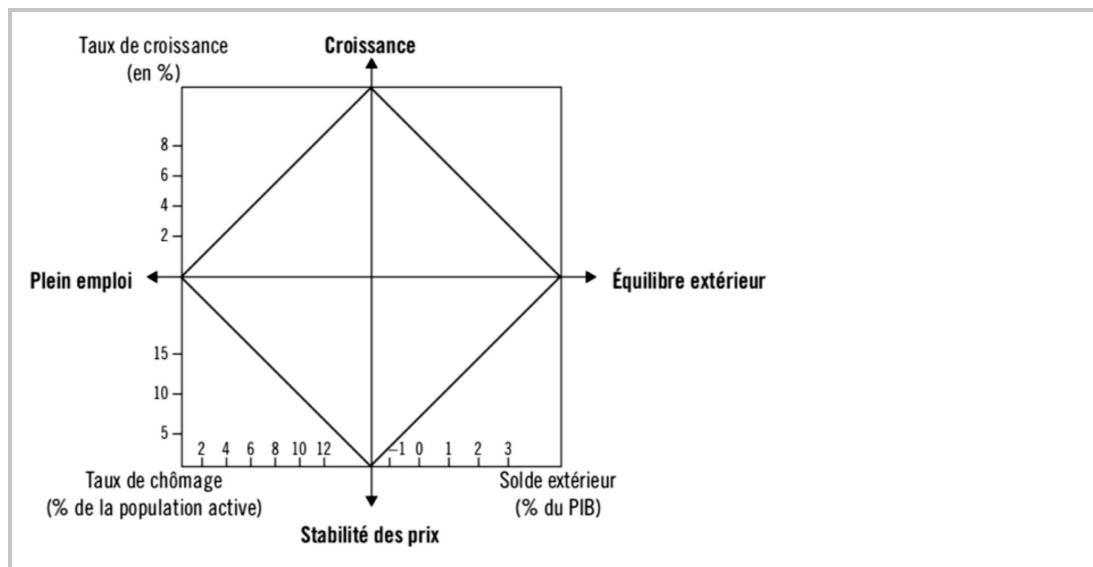
Le taux de croissance économique est mesuré par le pourcentage de variation annuelle du produit intérieur brut (PIB) en volume. Le revenu vient de la production : la croissance du PIB améliore donc le revenu national, constitue le moteur principal du développement économique et permet de couvrir les besoins de la population. Une croissance du PIB supérieure à celle de la population (croissance du PIB/habitant) améliore le revenu réel moyen de la population et la capacité des individus à satisfaire leurs besoins. La croissance du PIB suppose l'utilisation plus intensive des facteurs de production et contribue donc au plein emploi. La croissance permet de préparer l'avenir : le développement rapide des revenus permet de dégager une épargne qui assure à son tour le financement des investissements productifs ; or, ces investissements productifs assurent la production et l'emploi futurs. Quand elle développe le pouvoir d'achat de la majorité des individus, la croissance atténue les conflits entre les différents groupes sociaux pour le partage du revenu national. La croissance permet aussi de financer la protection sociale et des services collectifs qui constituent une part essentielle du bien-être de la population : justice, éducation, santé, infrastructures routières et portuaires, etc. L'objectif de croissance consiste donc à maintenir le PIB effectif proche du PIB potentiel, défini comme le PIB qui assure le plein emploi des facteurs de production sans accélération de l'inflation. Si le PIB se développe au-delà du PIB potentiel, l'inflation s'accélère ; si le PIB reste en deçà du potentiel, le chômage se développe. Si la croissance reste l'objectif principal de la politique économique, la mesure du taux de croissance peut poser des difficultés et la croissance peut générer des coûts économiques et sociaux (inflation, externalités négatives de pollution, inégale répartition des fruits de la croissance).

Le plein emploi

Le plein emploi est réalisé si l'ensemble des facteurs de production disponibles dans l'économie sont utilisés de la façon la plus efficace possible, c'est-à-dire dans l'usage qui leur apporte la plus forte productivité. Le problème du plein emploi concerne donc en théorie autant le capital que le travail, même si, dans les préoccupations politiques, la question est le plus souvent ramenée au seul facteur travail. Le chômage constitue avant tout un gaspillage des facteurs de production : une utilisation maximale de la population active disponible permettrait une production et donc un revenu national plus élevés (allocation optimale des ressources). Le coût du sous-emploi ne se limite pas à la perte de production et de revenu des agents directement concernés. À ces coûts privés s'ajoutent des coûts sociaux : le chômage, par exemple, entraîne une indemnisation des chômeurs qui est financée, d'une manière ou d'une autre, par des prélèvements sur les revenus des autres agents (travailleurs employés et entreprises). Quand le chômage représente une fraction importante de la population active, il implique une faible consommation d'une partie encore plus importante de la population. Une demande plus faible pour les biens de consommation réduit les revenus des entreprises installées dans ce secteur. Le caractère inégalitaire du chômage peut entraîner des coûts sociaux importants, liés en particulier au chômage de longue durée, qui fragilise le lien social.

L'équilibre extérieur

Le déficit des échanges extérieurs (la différence entre les exportations et les importations), s'il est durable, débouche sur un appauvrissement de la nation. Ce déficit souligne généralement une perte de compétitivité des entreprises dans la concurrence mondiale. La dépréciation de la devise nationale renchérit les importations. Le déficit commercial appelle tôt ou tard des mesures de rigueur monétaire ou budgétaire pour freiner la demande intérieure. À long terme, on ne peut durablement tolérer une détérioration trop importante des échanges extérieurs, dans un contexte de surveillance des politiques économiques par les marchés financiers, qui jugent de la crédibilité des différents États : s'ils interprètent cette détérioration du commerce extérieur comme un signe d'affaiblissement de la compétitivité, ils jugeront le risque plus grand et prêteront des capitaux au pays à des taux d'intérêt plus élevés.

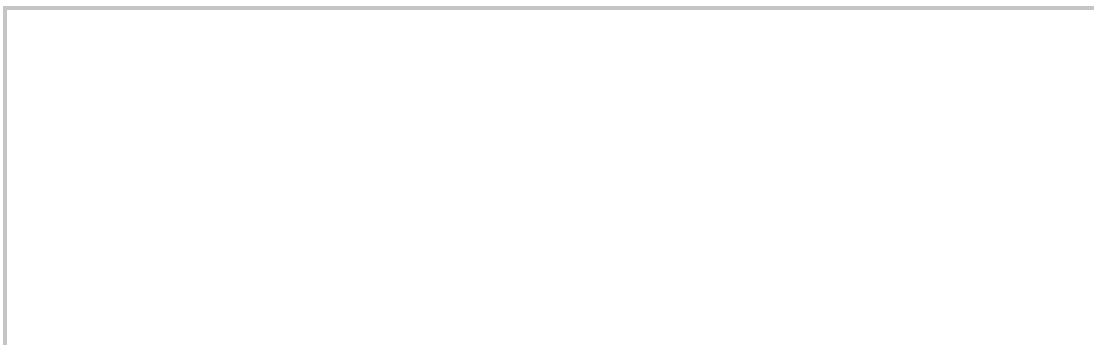


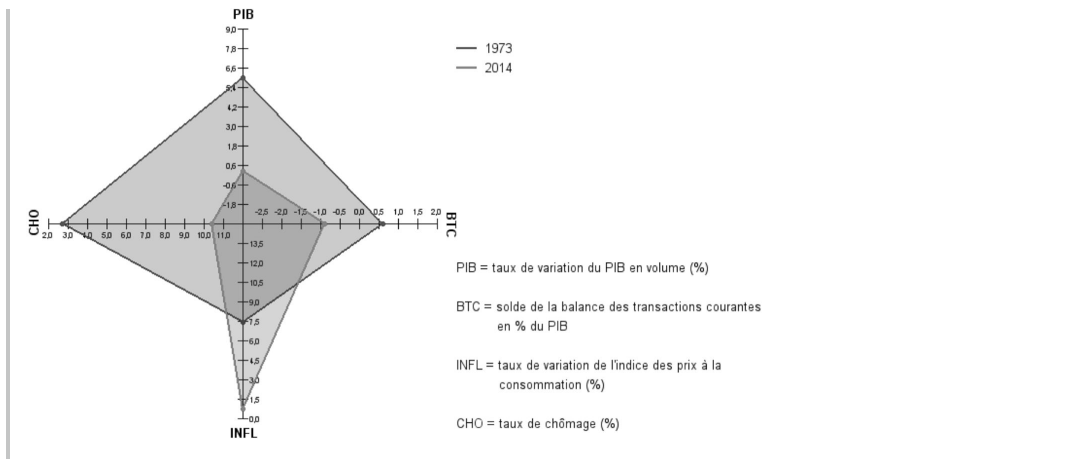
GRAPHIQUE 18.1.

Le carré magique

La stabilité des prix

Les grandes banques centrales visent généralement un taux d'inflation faible. L'inflation engendre en effet un certain nombre de coûts : elle constitue une dépréciation de toutes les valeurs monétaires : elle dégrade le pouvoir d'achat des agents dotés de revenus fixes non indexés sur le niveau des prix ; elle diminue la valeur réelle des créances détenues par les agents qui prêtent leurs capitaux à d'autres agents. L'inflation opère des redistributions occultes des revenus dans l'économie, au détriment de ceux dont le patrimoine, ou les rémunérations, sont plus mal protégés contre la dépréciation de la monnaie. Elle trouble le message donné aux agents par les variations de prix : en période d'inflation, la hausse des prix dans un secteur peut refléter aussi bien un mouvement général de dépréciation de la monnaie qu'un déplacement en faveur de ce secteur. L'inflation peut gonfler les profits nominaux alors que le marché est en régression et créer un biais statistique. Dans l'économie de marché, les signaux donnés par les prix deviennent chargés d'une information fausse et ceci perturbe l'allocation des facteurs de production. La contrainte extérieure se trouve renforcée par l'inflation : elle dégrade la compétitivité-prix des produits nationaux et alimente la spéculation contre la monnaie nationale sur les marchés financiers, en raison de la perte de crédibilité de la politique économique. Le risque le plus grave associé à l'inflation est la tendance de l'inflation à s'auto-développer et à s'accélérer, voire à dégénérer en hyperinflation (boucle prix-salaires, inflation importée).





GRAPHIQUE 18.2.

Le carré magique de la France en 1973 et en 2014

L'observation de la situation d'une économie dans le temps nous permet de voir qu'il est difficile, et même impossible, d'atteindre les quatre objectifs simultanément.

18.3 LES INSTRUMENTS DE LA POLITIQUE ÉCONOMIQUE

18.3.1 *La politique macroéconomique conjoncturelle*

L'État se donne pour objectif d'assurer une croissance économique soutenue, durable et équilibrée, c'est-à-dire de faire en sorte d'assurer la progression régulière du niveau de vie moyen de la population et le plein emploi sans accélération de l'inflation et sans déficit extérieur. L'agrégation des comportements des agents économiques peut entraîner une instabilité des marchés, tandis que le rôle des anticipations peut conduire à des excès d'optimisme, suivis d'excès de pessimisme dans une économie : l'État a alors une **fonction de régulation (ou de stabilisation)**, car la politique conjoncturelle peut atténuer les effets du cycle économique. De plus, cette stabilisation macroéconomique doit permettre de faire face à des chocs qui éloignent l'économie de l'équilibre : il s'agit de réduire les écarts par

rapport à cet équilibre en agissant sur la demande globale (consommation, investissement), par les politiques monétaire et budgétaire.

En effet, la croissance économique est un phénomène instable dans le temps : à des périodes de croissance soutenue (expansion) succèdent des périodes de forts ralentissements, voire de recul de la production pendant une période plus ou moins longue (récession et dépression). L'analyse des fluctuations économiques repose sur l'analyse des écarts pouvant exister entre, d'une part, le PIB effectif et la croissance effective (constatée) d'un pays et, d'autre part, le PIB potentiel et sa croissance potentielle. Le taux de croissance potentielle se calcule en mesurant la croissance maximale de la production qui peut être réalisée sans tensions inflationnistes compte tenu des facteurs de production disponibles. L'écart (*output gap*) entre le taux de croissance effective et le taux de croissance potentielle permet d'identifier la phase du cycle dans laquelle l'économie se trouve.

Une croissance effective supérieure à la croissance potentielle correspond à une période d'expansion où existent des tensions inflationnistes (« surchauffe »). Quand la croissance potentielle est supérieure à la croissance effective, l'économie est en situation de récession où le chômage augmente.

Les pouvoirs publics peuvent mettre en œuvre des politiques contra-cycliques (qui luttent contre l'instabilité cyclique) pour « lisser » les fluctuations de l'économie. Ces politiques n'ont d'effet qu'à court terme et ne permettent pas d'augmenter la croissance potentielle d'un pays.

La croissance économique n'est pas un processus stable et continu, mais s'accompagne de récessions : il y a alors ralentissement de la croissance de la production et montée du chômage. À l'inverse, si la croissance économique est trop forte, l'inflation et le déséquilibre du commerce extérieur menacent. La crise correspond au retournement brutal de la

conjoncture économique qui marque la fin d'une période d'expansion dans un cycle économique. Par extension, il s'agit d'une phase de contraction durable et de dépression de l'activité économique.

La politique économique conjoncturelle a une fonction essentielle de *régulation* du cycle économique.

En matière d'instruments de régulation conjoncturelle, on peut retenir deux indicateurs principaux :

- l'évolution des taux d'intérêt directeurs de la banque centrale qui conditionne le refinancement des banques commerciales et l'évolution du volume du crédit accordé aux ménages et aux entreprises (politique monétaire) ;
- l'évolution du solde budgétaire public qui dépend des dépenses de l'État et des prélèvements obligatoires (politique budgétaire). Ces deux indicateurs influencent le niveau de la demande globale.

On doit donc distinguer deux instruments de la politique conjoncturelle : la politique monétaire et la politique budgétaire

- La **politique monétaire** consiste pour la banque centrale à maintenir le pouvoir d'achat de la monnaie tout en se préoccupant de la stabilisation de l'activité économique en contrôlant le taux d'intérêt à court terme. C'est l'intervention des autorités monétaires en vue d'obtenir par l'action sur les variables monétaires et financières (taux d'intérêt, masse monétaire, taux de change, crédit) une modification des conditions de fonctionnement de l'économie à court comme à long terme.

La représentation courante en économie est que la politique monétaire agit sur la demande globale à court terme, et que la variation de la quantité de monnaie n'a qu'une influence sur le niveau général des prix à plus long terme.

- La **politique budgétaire** cherche à agir sur la situation macroéconomique par l'intermédiaire des finances publiques. Le budget des administrations publiques, qui

recouvre les budgets de l'État, des administrations locales et des administrations de sécurité sociale, remplit à la fois les fonctions d'allocation, de redistribution et de stabilisation. Il ne faut pas oublier que la politique budgétaire peut contribuer aux politiques structurelles en favorisant l'accumulation du capital humain, du capital technologique et du capital public (infrastructures).

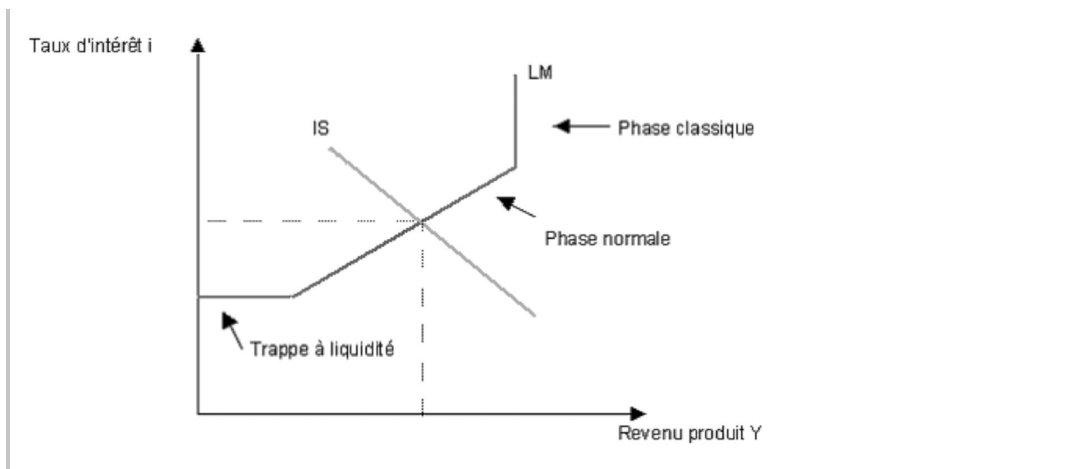
Toutefois, le terme de politique budgétaire renvoie essentiellement à la politique de stabilisation : selon l'économiste Paul Samuelson, c'est le « processus consistant à manipuler les impôts et les dépenses publiques aux fins : 1° de contribuer à amortir les oscillations du cycle économique, et 2° de favoriser le maintien d'une économie progressive, assurant un degré d'emploi élevé, affranchie de toute poussée excessive d'inflation ou de déflation ». De tout temps, les gouvernements ont engagé des dépenses publiques (de fonctionnement, d'investissement public, de transferts sociaux), qu'il est nécessaire de financer. Ils disposent pour cela de trois instruments : les impôts, la création monétaire, et l'emprunt public.

L'impact de la politique budgétaire sur la conjoncture demeure un sujet controversé. La conjoncture correspond à l'activité économique à un moment donné et se mesure par le niveau du PIB effectif à ce moment-là. Pour un niveau de PIB potentiel donné, l'état de la conjoncture dépend de la demande globale. Dans une perspective keynésienne, il existe donc un effet multiplicateur entre une variation initiale de la demande et la variation de l'activité économique, et une politique budgétaire de relance peut viser une augmentation de l'activité économique par une hausse de la demande globale. L'influence du budget sur la conjoncture peut ainsi s'exercer de façon passive, par les **stabilisateurs automatiques**, ou de façon active, par une politique budgétaire **discrétionnaire** et une stratégie volontariste d'accroissement des dépenses publiques ou de baisse des taux d'imposition.

La nécessité de réformer l'État et de rationaliser la dépense publique en France est plus que jamais d'actualité. Si la globalisation financière a permis le financement des déficits publics, elle a également accru la dépendance de la nation aux marchés internationaux de capitaux : toutefois, en particulier grâce à des politiques monétaires accommodantes (taux d'intérêt bas), la France en 2018 continue de bénéficier de taux de financement relativement bas (mais jusqu'à quand ?), proches de ceux de l'Allemagne, mais surtout en raison d'une défiance plus grande vis-à-vis des pays du sud de la zone euro.

18.3.2 *La politique conjoncturelle dans le cadre du modèle IS/LM et la trappe à liquidités*

Le modèle IS-LM de J. R. Hicks (1904-1989) et A. H. Hansen (1887-1975) constitue un compromis entre les analyses de J. M. Keynes et celles des « classiques ». Il permet de modéliser le fonctionnement de la politique conjoncturelle. Ce modèle envisage un équilibre sur les deux marchés (marché des biens et marché de la monnaie) et met en relation niveau du produit national (revenu nominal) et niveau du taux d'intérêt nominal. La courbe investissement/épargne (IS), décroissante, représente l'ensemble des situations d'équilibre possibles sur le marché des biens. Elle représente l'ensemble des couples (i, Y) pour lesquels il y a égalité entre investissement et épargne. La courbe demande de monnaie/offre de monnaie, croissante, représente l'ensemble des situations d'équilibre sur le marché de la monnaie. Elle représente l'ensemble des couples (i, Y) pour lesquels il y a égalité entre demande et offre de monnaie. L'offre de monnaie M est, dans ce modèle, une variable exogène qui dépend de la politique de l'État en matière de création monétaire. L'intersection des deux courbes détermine la combinaison du taux d'intérêt et du revenu nominal qui assure simultanément l'équilibre sur le marché de la monnaie et sur le marché des biens.



GRAPHIQUE 18.3.
Schéma IS/LM et trappe monétaire

Ce modèle décrit **l'interdépendance des politiques budgétaire et monétaire** : une augmentation du montant des dépenses publiques pousse la courbe IS vers la droite ; le revenu croît d'autant plus que l'offre de monnaie est en même temps accrue, limitant ainsi la hausse des taux d'intérêt. Une politique monétaire expansive déplace la courbe LM vers la droite, générant un nouvel équilibre (déplacement du point d'équilibre du marché des biens sur la courbe IS), fondé sur une augmentation du revenu nominal et une baisse du taux d'intérêt. Inversement une politique monétaire restrictive (réduction de la création monétaire), pour faire face à une situation marquée par un plein emploi et des tensions inflationnistes, déplace la courbe LM vers la gauche (restriction du revenu et augmentation du taux d'intérêt). Dans la même perspective d'une réduction des tensions inflationnistes, une politique budgétaire restrictive génère un déplacement de la courbe IS sur la gauche (réduction du revenu et du taux d'intérêt).

Selon le modèle IS/LM, l'économie peut se retrouver dans une **situation de trappe à liquidités**. En dessous d'un certain niveau de taux d'intérêt, l'élasticité de la demande de monnaie au taux d'intérêt est infinie et l'élasticité de l'investissement est dans le même temps nulle. Lorsque l'économie se trouve dans la trappe à liquidité, seule une

politique budgétaire peut permettre un retour vers le plein emploi (déplacement de la courbe IS vers la droite).

En poursuivant des stratégies de relance monétaire, les pays peuvent se retrouver piégés par une « trappe à liquidités », dans laquelle les politiques monétaires perdent toute efficacité. Lors des récessions, les banques centrales tentent généralement de stimuler l'économie en abaissant les taux d'intérêt directeurs, afin d'encourager les investissements, mais, si les taux d'intérêt sont déjà très bas, alors il peut s'avérer impossible de les baisser encore davantage. Dans ce cas, la stratégie de relance monétaire bute sur le fait que la monnaie supplémentaire distribuée n'est pas dépensée, mais simplement détenue sous la forme de liquidités, au lieu d'être prêtée pour financer l'investissement. Pour sortir de cette situation de « trappe à liquidités », il faut alors déployer d'autres moyens d'action, comme les politiques d'expansion monétaire dites « non conventionnelles », ou recourir à l'inflation, qui, lorsque les taux d'intérêt nominaux sont à zéro, réduit les taux d'intérêt réels (taux d'intérêt nominaux – taux d'inflation) et soutient l'investissement des entreprises.

18.3.3 *Les contraintes sur la politique monétaire en économie ouverte*

La politique monétaire perd de son autonomie dans la mondialisation de l'économie : si la banque centrale décide de baisser son taux d'intérêt directeur pour stimuler le crédit à l'économie et la demande intérieure, les marchés financiers peuvent craindre et anticiper une accélération de l'inflation qui dévalorise le pouvoir d'achat des titres financiers dont ils ont la charge et exiger, en conséquence, des taux d'intérêt plus élevés sur le remboursement de la dette publique. De plus, la baisse du taux d'intérêt domestique, si ce dernier devient inférieur aux taux d'intérêt pratiqués sur les autres places financières, peut déclencher des sorties de capitaux, provoquer des difficultés financières de l'État et favoriser une dépréciation de la monnaie nationale. En effet, les sorties de

capitaux se traduisent par une vente de la monnaie nationale sur le marché des changes, ce qui induit une baisse du taux de change. Toutes choses égales par ailleurs, cette dépréciation du change peut alors stimuler les exportations, mais renchérir le coût des importations, et élever le coût des investissements directs à l'étranger des entreprises nationales.

Les banques centrales doivent aujourd'hui concilier l'objectif du soutien à l'activité économique et celui de la stabilité monétaire pour garantir le pouvoir d'achat de la monnaie et éviter l'inflation. Le contrôle de l'évolution du prix des actifs financiers, éminemment complexe, constitue aujourd'hui un nouvel enjeu pour les banques centrales dans le cadre de la mondialisation financière, pour prévenir les effets délétères des crises financières.

Depuis la crise financière internationale de 2007-2008, la politique monétaire a été sollicitée de manière très importante par des baisses des taux d'intérêt directeurs, afin d'éviter le rationnement du crédit (« credit crunch »), très nocif pour la croissance économique. Par la suite, les banques centrales, face à l'aggravation de la crise, ont mis en œuvre des politiques dites « non conventionnelles » de forte création monétaire pour soutenir l'activité économique (à l'instar de la Réserve fédérale des États-Unis, qui a décidé de la mise en place des stratégies de « quantitative easing »). Durant la crise des dettes publiques dans la zone euro, en 2009-2010, la Banque Centrale européenne (BCE) a joué également un rôle déterminant pour soutenir les États défaillants et a poursuivi une politique monétaire expansionniste, même si elle ne peut racheter directement la dette des États. Dans le cadre de l'Union économique et monétaire (UEM), la politique monétaire est unique (les États ont transféré le pouvoir monétaire à la BCE), mais les politiques budgétaires demeurent nationales et sont étroitement encadrées par des règles budgétaires.

18.4 LES CONTRAINTES QUI LIMITENT LES EFFETS DE LA POLITIQUE BUDGÉTAIRE

18.4.1 *La politique budgétaire face à la « contrainte extérieure »*

Si la politique budgétaire constitue un instrument pour stabiliser l'économie, encore faut-il que l'État dispose d'une marge de manœuvre pour la mettre en application. Avec la mondialisation de l'économie, la politique budgétaire est soumise à de nouvelles contraintes économiques et financières. La mondialisation modifie l'environnement dans lequel la politique budgétaire se déploie : son efficacité est réduite face à la contrainte extérieure et financière. En économie fermée, la politique budgétaire d'inspiration keynésienne permet de soutenir la demande globale en cas de ralentissement conjoncturel : l'État augmente les dépenses publiques (investissements publics, prestations sociales, aides aux entreprises, grands travaux) ou baisse les impôts pour augmenter le revenu disponible des ménages et soutenir l'accumulation du capital productif et l'investissement privé des entreprises.

Or, la mondialisation économique réduit l'efficacité de ce type de politique budgétaire de relance : l'ouverture aux échanges incite les consommateurs à acheter des biens étrangers si la production nationale ne satisfait pas les besoins du marché domestique. La hausse du revenu des ménages se traduit alors par une hausse des importations de biens de consommation étrangers, et les entreprises peuvent également décider d'importer des biens de production (machines, par exemple). Pour des exportations constantes, le déficit commercial se creuse et peut pousser l'État à faire machine arrière pour éviter une dépréciation de la monnaie nationale, et revenir à une politique budgétaire de rigueur qui

freine la croissance économique (le cas de la France entre 1981 et 1983).

La mondialisation des échanges commerciaux a donc pour effet d'amoinrir les retombées économiques d'une politique de relance en raison de la « contrainte extérieure ».

18.4.2 La politique budgétaire face à la contrainte financière : controverses sur le niveau de la dette publique

La mondialisation a accru le pouvoir des marchés financiers sur les États, en raison de l'importance accordée aux agences de notation (Standard & Poor's et Moody's, par exemple), qui attribuent des cotes de crédit aux gouvernements. Ainsi, les politiques des États sont surveillées et évaluées par ces agences, qui n'hésitent pas à abaisser la cote de crédit d'un gouvernement lorsque sa situation financière se dégrade. Cette décote financière se traduit par la hausse du taux d'intérêt, en raison de la prime de risque exigée par les investisseurs financiers. La charge de la dette publique augmente alors et les grandes entreprises qui empruntent pour investir sont pénalisées par des taux d'intérêt plus élevés. Dans le cadre de la globalisation financière, les États lèvent des capitaux sur les marchés financiers pour financer les dépenses publiques, et les opérateurs financiers et les investisseurs étrangers sont très vigilants sur la soutenabilité à long terme de la dette publique. S'ils estiment que la politique budgétaire de relance peut accélérer l'inflation (qui détériorera le pouvoir d'achat des actifs dont ils ont la charge), ils peuvent alors imposer aux États d'emprunter à des taux d'intérêt plus élevés face à une perte de pouvoir d'achat de la monnaie et au risque de crédit (la probabilité plus forte que l'État fasse défaut sur le remboursement de sa dette).

L'État est alors forcé de consacrer davantage de ressources au service de la dette, laquelle comprend le remboursement du capital et des intérêts. Récemment, la crise des dettes

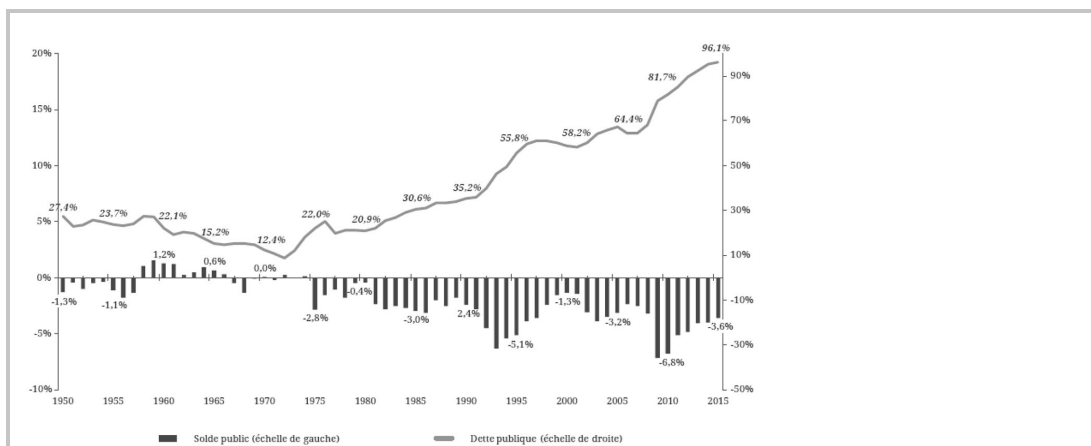
publiques dans la zone euro en 2009-2010 a montré tout l'enjeu, pour un État, de convaincre les marchés financiers de la crédibilité de sa politique budgétaire pour rassurer les créanciers. Si certains considèrent que les marchés financiers exercent une tutelle excessive sur les politiques macroéconomiques des États, en limitant leur marge de manœuvre et l'espace des choix démocratiques, d'autres font valoir que les marchés financiers jouent un rôle nécessaire de garde-fou pour empêcher les États d'adopter une gestion laxiste de leurs finances publiques.

Comme les capitaux et les entreprises se déplacent librement d'un pays à l'autre, les États se livrent une concurrence acharnée pour les attirer ou les retenir sur leur territoire. D'un côté, les entreprises exercent des pressions pour réduire l'impôt sur les profits, ce qui affecte la marge de manœuvre de l'État au chapitre des recettes fiscales. D'un autre côté, les investisseurs financiers recherchent les meilleurs rendements et fuient les risques associés à l'endettement croissant d'un État.

Selon le contexte économique, la dynamique de la dette publique peut devenir incontrôlable, ce qui oblige les États à mettre en œuvre des politiques budgétaires de rigueur.

Les facteurs à l'origine de la dégradation des comptes publics des principaux pays développés (dont la France) sont multiples : à partir des années 1970, le ralentissement de l'activité économique a généré un « effet de ciseaux », avec un tarissement des recettes fiscales et une augmentation des dépenses publiques (en particulier des dépenses sociales, sous l'effet de l'aggravation de la crise) ; de plus, certains facteurs structurels, comme le vieillissement de la population (financement des retraites et du système de santé), ont pesé sur les comptes publics ; à partir des années 1980, la hausse des taux d'intérêt réels (sous l'effet des politiques monétaires de rigueur et de la désinflation) a accru la charge de l'endettement public. S'enclenche alors un « effet boule de neige de la dette publique », lorsque le poids de la dette publique (ratio dette/PIB) augmente, du fait que le taux

d'intérêt réel demeure durablement supérieur au taux de croissance du PIB en volume. Dans ce cas, il y a auto-accumulation de la dette, il faut alors que le solde primaire du budget (hors charge de la dette, c'est-à-dire non compris les intérêts versés sur la dette) soit excédentaire. En revanche, si le taux d'intérêt est inférieur au taux de croissance, le poids de la dette peut ne pas augmenter, malgré l'existence d'un déficit primaire. La réduction du poids de la dette impose donc soit que l'État dégage un excédent budgétaire primaire important (cela implique une baisse des dépenses ou une hausse des impôts), soit qu'il y ait accélération de la croissance du PIB, soit que les taux d'intérêt réels baissent. L'État doit donc contrôler la dynamique de la dette publique. Celle-ci verra son poids grossir si le taux d'intérêt progresse plus vite que la croissance nominale du PIB, à moins que l'État n'améliore son solde primaire (et ce d'autant plus que l'écart entre le taux d'intérêt et le taux de croissance est important). Pour souligner l'importance du rôle de l'écart entre le taux d'intérêt et le taux de croissance dans la dynamique de la dette, on parle d'ailleurs d'« écart critique ». Lorsque le niveau de dette est rapporté à la taille de l'économie et exprimé en pourcentage du PIB, ce qui est généralement le cas, la condition pour le stabiliser est d'autant moins exigeante que la croissance du PIB est rapide. En effet, pour stabiliser le ratio d'endettement (dette/PIB), il suffit que numérateur (dette) et dénominateur (PIB) croissent au même rythme.



Source : France Stratégie, d'après Insee, complété par les données de Reinhart et Rogoff (2011) pour les données de la dette publique avant 1995

GRAPHIQUE 18.4.

France : Solde budgétaire et dette publique depuis 1950

La hausse des déficits et des dettes publics soulève la question de la *soutenabilité de la dette*. On dit que la situation des finances publiques est insoutenable si, sur la base de la politique économique suivie et des prévisions disponibles, l'évolution probable de la dette publique débouche sur une situation d'insolvabilité, c'est-à-dire sur l'incapacité de l'État à faire face à ses échéances compte tenu de ses ressources. Une politique budgétaire est insoutenable si elle conduit à une accumulation de dettes qui, sans changement de cap, ne pourrait pas être couverte à l'avenir par des excédents budgétaires. Si un État peut toujours, en théorie, assurer le service de la dette, dans la mesure où il peut lever l'impôt, dans la réalité, il existe des limites politiques et sociales à la hausse de la pression fiscale pesant sur la population (avant les prochaines élections). Il n'existe toutefois pas de seuil de soutenabilité de la dette publique universellement accepté, et il faut notamment se poser deux questions : quelle est la part de la dette nationale détenue par les investisseurs étrangers ? Est-ce que l'accroissement du poids de la dette publique est le résultat d'investissements publics (accumulation de capital public favorable à la croissance future) ou n'est-il que le résultat d'une augmentation des dépenses courantes, sans effet sur la croissance potentielle ? Dans le cas où les marchés financiers sont efficaces, cela peut se traduire dans le second cas par une élévation des taux d'intérêt, résultat d'un plus grand risque d'insolvabilité de l'État. Dans le premier cas, au contraire, toutes choses égales par ailleurs, la dette publique reste soutenable. Enfin, une dette publique relativement élevée peut alors être compensée par un faible endettement des ménages (il y a alors un effet de compensation avec l'endettement public), la socialisation

des dépenses de santé ou d'éducation de l'État limitant les dépenses des agents privés.

Malgré toutes ces contraintes, la crise de 2008 a montré toute l'importance de l'intervention de l'État dans la régulation conjoncturelle, malgré le contexte de mondialisation économique. Aux États-Unis et en Europe, l'action massive des pouvoirs publics a été indispensable pour stabiliser les marchés financiers, restaurer la confiance dans tout le système bancaire et fournir à nouveau aux ménages et aux entreprises les liquidités nécessaires à leur activité. Par la suite, la politique budgétaire a été à nouveau sollicitée : les principaux pays développés ont mis en œuvre des plans de relance pour substituer la demande publique (de l'État par ses commandes) à la demande privée (celle des ménages et des entreprises) faiblissante. Les dépenses publiques, sous la forme d'aides à l'indemnisation des chômeurs, de soutien financier aux entreprises, de politiques de grands travaux, ont permis de soutenir la demande globale (soit la consommation et l'investissement) et, donc, la croissance économique. Des réductions d'impôts ont également été mises en œuvre pour stimuler le revenu disponible des ménages et favoriser l'investissement des entreprises en période de crise. Des aides ont notamment été ciblées sur le soutien aux bas revenus, qui ont une forte propension à consommer, et la conjonction des plans de relance budgétaire a permis pour l'heure d'éviter la spirale de la déflation et de la dépression que les pays développés avaient connues durant les années 1930. À l'heure actuelle, il demeure un débat entre les économistes sur l'opportunité de poursuivre dans la zone euro des politiques budgétaires tournées vers la rigueur budgétaire, malgré le niveau élevé de l'endettement de nombreux États, alors qu'il subsiste un écart important entre la croissance effective du PIB et la croissance potentielle, et un contexte de lente reprise économique.

En février 2017, dans le cadre d'un rapport dirigé par Michel Pébereau, l'Institut de l'Entreprise a publié, peu avant l'élection présidentielle de 2017, une nouvelle note intitulée

« Dépense publique : l'état d'alerte », sur la situation jugée alarmante de nos finances publiques. Le rapport indique qu'entre 1989 et 2012, le solde structurel entre dépenses et recettes hors effet de conjoncture s'est dégradé de 2,8 points de PIB. Selon cette étude, cette dégradation résulte de l'augmentation de la dépense publique (+6,1 points de PIB), qui atteint 57 % du PIB en France en 2015 (contre 48,5 % en moyenne dans la zone euro), au second rang du classement de l'OCDE derrière la Finlande. C'est la raison pour laquelle, malgré les économies réalisées depuis 2012, l'étude stipule que la dépense publique continue de croître en volume (entre 2007 et 2015, elle a augmenté de plus de 250 milliards d'euros). Or, pour faire face à cette progression de la dette publique, la France a dû « massivement » augmenter ses prélèvements obligatoires, en affaiblissant la compétitivité des entreprises et en pénalisant l'emploi et l'attractivité. En frôlant les 100 % du PIB, la dette publique pourrait devenir difficilement soutenable en cas de remontée des taux d'intérêt (le rapport indique qu'avec le niveau des taux d'intérêt de 2005, « la charge de la dette publique française doublerait »). L'étude de l'Institut de l'Entreprise préconise alors une réduction de la dépense publique immédiate, transversale et pérenne, avec une maîtrise des dépenses sociales, une réduction des dépenses de fonctionnement de l'État, et une mobilisation contractuelle des collectivités territoriales pour programmer la réduction des dépenses. Pour les économistes de l'Observatoire français des conjonctures économiques (OFCE), le niveau d'endettement de la France n'a rien d'exceptionnel au sein des pays avancés : il est surtout dû à la crise économique (qui a tari les recettes fiscales et accru les dépenses sociales), et non au « laxisme » de la gestion publique. Même si la dette publique atteint 97,6 % du PIB à la fin du premier trimestre de 2018, ce niveau élevé n'est pas un argument, selon l'OFCE, pour sacrifier l'investissement public, nécessaire à la modernisation du système éducatif ou à la transition énergétique. Une trop forte compression de la dépense

publique dans le cadre de l'austérité budgétaire pourrait d'ailleurs s'avérer contre-productive et affaiblir le niveau de la croissance potentielle de l'économie à long terme.

18.4.3 *La question controversée des multiplicateurs de la politique budgétaire*

Certains macro-économistes rappellent qu'un ajustement budgétaire trop brutal peut exercer des effets récessifs sur l'activité économique : cette question reste d'ailleurs âprement controversée depuis la dernière crise de la dette publique en zone euro. Un ajustement budgétaire consiste en une augmentation de la capacité de financement des administrations publiques : il correspond soit au passage d'un déficit à un excédent, soit, *a minima*, à une réduction du déficit public. Cet ajustement budgétaire a pour objectif de freiner la hausse de la dette publique, et il prend la forme soit d'une augmentation des prélèvements obligatoires, soit d'une diminution de la dépense publique. Mais envisager un ajustement budgétaire suppose toutefois de tenir compte de ses effets sur l'activité économique. Les paramètres qui influencent les effets d'un ajustement sur l'activité économique sont les mêmes que ceux qui déterminent l'efficacité d'une politique budgétaire de relance, mais ils fonctionnent en sens inverse : d'un point de vue keynésien, un ajustement budgétaire a des effets récessifs, car il entraîne une baisse de la demande globale (suite à la hausse des impôts ou à la baisse de la dépense publique). Mais dans un cadre néoclassique, en cas d'effet d'éviction, ces effets récessifs sont amoindris (une baisse du déficit public entraîne une baisse des taux d'intérêt, qui favorise une hausse de la dépense privée). Il en va de même en cas d'effet Ricardo-Barro (une baisse du déficit budgétaire entraîne une baisse de l'épargne et une hausse de la consommation...). Mais d'une façon générale, les effets récessifs d'un ajustement

budgétaire sont d'autant plus importants que la valeur du multiplicateur budgétaire est élevée.

D'autre part, si certaines études ont soutenu que les ajustements budgétaires menés pendant les années 1980-1990 dans certains pays (Canada, Suède...) avaient eu des effets expansionnistes, conformément à l'effet Ricardo-Barro, Olivier Blanchard, l'ancien chef économiste du FMI, a récemment affirmé que personne ne croyait vraiment à cet effet sur le plan pratique tant ses conditions théoriques de déclenchement sont restrictives. Pour autant, au cours des années 1990, de nombreux économistes attribuaient en général une valeur faible au multiplicateur budgétaire (ce qui revenait à considérer qu'il était possible d'effectuer des ajustements budgétaires sans que ceux-ci ne se traduisent par des effets récessifs) et accordaient plutôt leur confiance à la politique monétaire pour stabiliser l'économie. Ce point de vue a toutefois été modifié à partir des années 2010. Le FMI a ainsi révisé en 2013 son évaluation du multiplicateur budgétaire grec en déclarant avoir fortement sous-estimé l'impact récessif de l'ajustement budgétaire imposé à l'économie grecque : les scénarios élaborés pour la Grèce en 2009 dans le cadre de l'austérité budgétaire étaient basés en effet sur un multiplicateur budgétaire de l'ordre de 0,5 alors qu'au bout du compte il aurait été de 1,5...

L'EFFET MULTIPLICATEUR

Le modèle keynésien de détermination de la production repose sur la fonction de consommation selon laquelle, quand le revenu augmente, la consommation croît, mais moins que proportionnellement à la hausse du revenu. Le niveau de l'épargne augmente donc tendanciellement avec le revenu, et la propension marginale à consommer c (rapport entre variations de la consommation et du revenu) est positive et inférieure à 1. En situation de sous-emploi et de capacités de production inutilisées, l'existence d'un investissement autonome et la stabilité de la fonction de consommation impliquent donc un effet multiplicateur. En effet, une hausse autonome de l'investissement signifie une distribution de revenu en partie consommée pour une fraction c . Ceci entraîne alors une nouvelle hausse de la production qui engendre une nouvelle distribution de revenu, dont une partie c est à nouveau consommée. Le processus se poursuit sous la forme d'une succession de vagues dont l'ampleur est décroissante. Le rapport entre le supplément de revenu et le supplément d'investissement est égal au multiplicateur :

$$k = \frac{1}{1-c}, \text{ et on a } k > 1 \text{ dans la mesure où } 0 < c < 1.$$

Le mécanisme du multiplicateur met l'accent sur l'importance des dépenses d'investissement des entreprises dans la fixation du niveau de l'emploi, mais il est possible d'introduire le rôle des dépenses de l'État (G) qui vont enclencher le même processus vertueux. C'est la propension à épargner qui ralentit progressivement l'impact de la hausse autonome des dépenses publiques sur la production,

tandis qu'en économie ouverte, la propension à importer réduit aussi la valeur du multiplicateur et son impact sur la production nationale, puisqu'une partie de la hausse de la demande se tourne vers les biens étrangers.

Les politiques budgétaires peuvent donc avoir des *effets procycliques* nocifs : si la récession se traduit par une hausse des déficits publics, des mesures budgétaires trop restrictives prises pour contrer cette hausse des déficits publics (hausse des impôts ou baisse de la dépense publique) peuvent générer en retour des effets récessifs qui aggravent la crise. De « l'erreur de 1937 » aux États-Unis à l'effet récessif des politiques restrictives menées en Europe dès 2010 en passant par « l'erreur de 1997 » au Japon, l'histoire a, de fait, montré les limites de la thèse de « l'austérité expansionniste », au moins dans les phases récessives du cycle. Le débat sur l'utilisation du budget a rebondi récemment aux États-Unis autour de la *Job guarantee* (« garantie d'emploi ») dans un prolongement des analyses hétérodoxes d'Abba Lerner et Hyman Minsky. Pour les tenants de la *Modern Monetary Theory* (Pavlina Tchernova, Stephanie Kelton), dans la mesure où le plein emploi est un bien commun, l'État pourrait, au-delà d'un rôle de soutien de la demande, utiliser son budget pour jouer celui d'un « d'employeur en dernier ressort » en fournissant une « garantie d'emploi ».

18.5 LE « POLICY-MIX » ET LE PILOTAGE CONJONCTUREL

18.5.1 *Le modèle Mundell-Fleming*

Le concept de « policy-mix » fait référence à l'ensemble des combinaisons possibles des politiques macroéconomiques conjoncturelles que sont la politique budgétaire et la politique monétaire. D'un point de vue théorique, le « policy-mix », en économie ouverte peut être analysé à partir du modèle Mundell-Fleming. On peut distinguer les combinaisons de politiques macroéconomiques conjoncturelles convergentes des « policy-mix » divergents.

Les politiques budgétaire et monétaire sont convergentes lorsque les deux sont expansionnistes ou restrictives. Elles sont divergentes lorsque l'une est expansionniste et l'autre restrictive. Cela donne quatre possibilités.

– **Cas n°1 : Politique budgétaire expansionniste et politique monétaire expansionniste.** La combinaison d'un soutien monétaire et d'une relance budgétaire doit théoriquement permettre de soutenir la croissance et l'emploi et d'éviter l'effet d'éviction lié à la hausse du taux d'intérêt provoquée par le creusement des déficits budgétaires. Toutes choses égales par ailleurs, ces soutiens cumulés doivent entraîner un effet sur le revenu global plus élevé que dans le cas d'une relance sur la base d'un seul instrument.

– **Cas n°2 : Politique budgétaire expansionniste et politique monétaire restrictive.** La combinaison d'une politique budgétaire expansionniste et d'une politique monétaire restrictive doit, théoriquement, provoquer une forte hausse des taux d'intérêt et avoir un effet faible sur le revenu global. Dans le cadre de ce « policy-mix », les effets se cumulent pour entraîner une hausse des taux d'intérêt, car il y a réduction de l'offre de monnaie (politique monétaire restrictive) et hausse de la demande de monnaie (politique budgétaire expansionniste). Cette forte hausse des taux d'intérêt peut se traduire par une période de croissance économique faible.

– **Cas n°3 : Politique budgétaire restrictive et politique monétaire expansionniste.** La combinaison d'une politique budgétaire restrictive et d'une politique monétaire expansionniste doit, théoriquement, entraîner une baisse des taux d'intérêt et aboutir à un effet affaibli sur le niveau de revenu global. Les deux politiques cumulent ici leurs effets pour entraîner une forte baisse des taux d'intérêt, car, d'une part, l'offre de monnaie croît avec la politique monétaire expansionniste et, d'autre part, la demande de monnaie des agents baisse en raison de la politique budgétaire restrictive.

– **Cas n°4 : Politique budgétaire restrictive et politique monétaire restrictive.** Le « policy-mix » doublement restrictif, politique monétaire et politique budgétaire restrictives, a principalement pour objectif la réduction des déficits budgétaires, de la dette publique, et la stabilité des prix. Les objectifs relatifs aux variables réelles comme la croissance et l'emploi sont secondaires. La contraction de la demande globale qui en découle entraîne une forte baisse du revenu global via notamment un maintien à un niveau durablement élevé des taux d'intérêt.

18.5.2 *La coordination face aux chocs symétriques et aux chocs asymétriques*

La question de la **coordination** des politiques économiques est fondamentale. Pour le comprendre, on distingue traditionnellement les **chocs asymétriques** et les **chocs symétriques** : les chocs asymétriques sont des perturbations aléatoires concernant l'offre ou la demande, mais qui affectent spécifiquement une branche d'activité, une région, un pays ou une spécialisation. Les chocs asymétriques posent un problème dans une union monétaire puisque deux pays de cet ensemble peuvent être affectés de manière divergente par ce type de choc alors qu'ils ont renoncé à utiliser le taux de change comme instrument d'ajustement.

On appelle choc symétrique un événement exogène ayant un impact similaire sur la demande globale et/ou l'offre globale des différents pays de la zone. Les variations du prix du pétrole constituent des chocs d'offre symétriques pour les pays de la zone euro ; les fluctuations de l'activité aux États-Unis constituent des chocs de demande presque symétriques.

En l'absence de coordination des politiques budgétaires, et si les externalités budgétaires sont positives, chaque pays est incité à ne pas tenter de stabiliser son activité, mais à attendre que ses partenaires le fassent. Ainsi, les politiques

budgétaires sont insuffisamment actives. D'où l'importance de la coordination des politiques économiques, que ce soit dans le cadre d'un processus d'intégration régionale (Union européenne et zone euro) ou dans le cadre de la mondialisation. En effet, en économie ouverte, les choix faits dans certaines économies ne sont pas sans effets sur les autres et réciproquement. De surcroît, dans une économie donnée, l'orientation de la politique monétaire n'est pas sans conséquence sur les marges de manœuvre budgétaires et *vice versa*. Enfin, l'ensemble de ces interdépendances se cumulent, avec de nombreuses autres. Il est donc évident que la coordination entre les politiques économiques d'un pays (ou dans le cadre d'une Union) reste un impératif et qu'un autre degré de coordination, ou de coopération, doit exister entre économies interdépendantes du fait de leurs échanges multiples.

18.6 LES POLITIQUES ÉCONOMIQUES STRUCTURELLES

Les politiques économiques structurelles regroupent les actions de l'État qui visent à agir à long terme sur les caractéristiques fondamentales de l'économie. Il peut s'agir des stratégies de planifications, de nationalisations, de politique industrielle ou d'aménagement du territoire. Les politiques économiques structurelles cherchent à stimuler la croissance potentielle, en particulier grâce à l'accumulation de différents capitaux (capital technologique, capital humain, capital public).

18.6.1 Des politiques structurelles pour agir sur la croissance potentielle

En effet, en matière de politique économique, on distingue généralement l'action conjoncturelle, qui exerce une action de régulation de la demande globale pour rapprocher la croissance effective du PIB de son niveau potentiel, et une action de plus longue haleine, qui consiste à réformer et à

adapter les institutions et les structures de l'économie pour augmenter la croissance potentielle du PIB (dite « structurelle »).

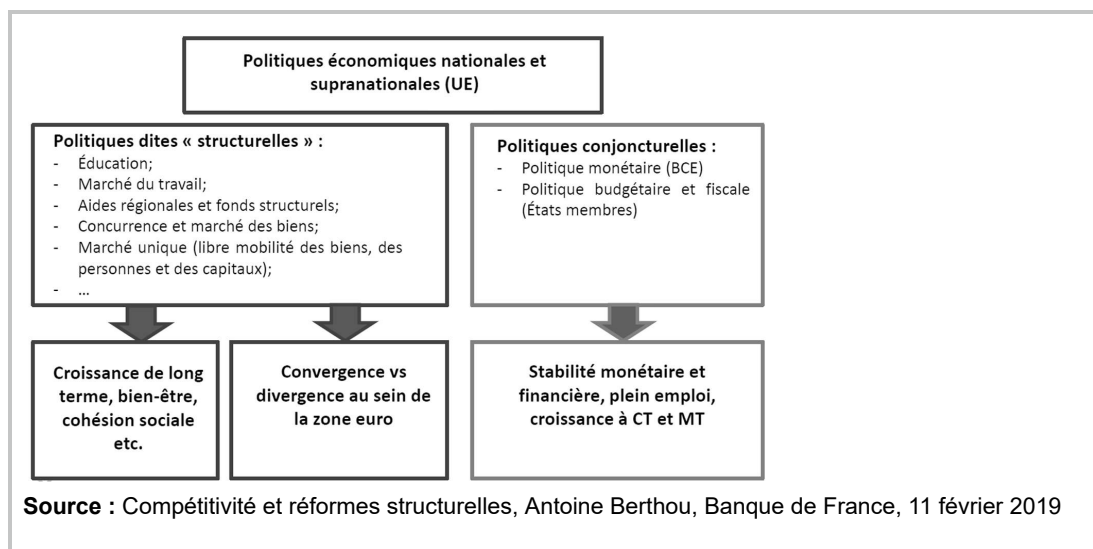
Cette politique structurelle est orientée vers le long terme et cherche à améliorer l'efficacité de l'appareil productif et la compétitivité des entreprises en créant un environnement macroéconomique favorable. Il s'agit notamment de rendre le marché du travail et le marché des biens plus concurrentiels, de réformer la fiscalité pour créer des incitations à l'investissement des entreprises, d'aider les secteurs stratégiques au titre de la politique industrielle et de stimuler l'attractivité du territoire et la productivité en renforçant le système éducatif (capital humain), l'innovation (capital technologique) et les infrastructures publiques (capital public).

18.6.2 *Les domaines d'action des politiques structurelles*

Les domaines traditionnels des politiques structurelles sont les politiques industrielle et sectorielle (aéronautique et automobile, par exemple), la politique territoriale (aménagement du territoire), la politique technologique (promotion de la recherche-développement et de l'innovation). Les politiques structurelles peuvent avoir comme objectif l'indépendance nationale et le développement économique (pour favoriser les secteurs comme l'énergie, le transport, l'agriculture), la lutte contre les défaillances du marché (par des instruments économiques et réglementaires de protection de l'environnement, une politique de la concurrence sur les marchés) et la justice sociale (politique de réduction des inégalités de revenus, politiques d'éducation pour favoriser la mobilité sociale, etc.). On peut aussi intégrer les politiques du marché du travail qui visent à accroître la qualité et le volume de l'emploi (action sur les réglementations du travail, la nature du contrat de travail, le salaire minimum, etc.).

18.6.3 Les enjeux des politiques structurelles en économie ouverte

Aujourd'hui, dans une économie globalisée, l'enjeu des politiques structurelles est aussi de créer un environnement fiscal et réglementaire propice aux activités économiques (et suffisamment attractif pour les investisseurs étrangers) : l'État peut donc favoriser la compétitivité des entreprises et attirer les entreprises étrangères grâce à la qualité de l'environnement offert (proximité des marchés, coût et qualité de la main-d'œuvre, ressources technologiques, sécurité des investissements, qualité des infrastructures de transport, de télécommunications et des services publics), tout en évitant d'accroître les charges fiscales et sociales.



GRAPHIQUE 18.5.

La structure des politiques économiques

Ainsi, dans le cadre d'une économie ouverte et d'un contexte de mutations technologiques rapides, les leviers de la politique économique se portent davantage sur les incitations fiscales (pour favoriser l'épargne, l'investissement et la création d'entreprises), le capital humain (l'éducation et la mobilité sociale, la recherche universitaire) et l'attractivité des territoires (pour inciter les investisseurs étrangers à créer de

l'activité sur le sol national). Mais la politique économique cherche également à promouvoir une plus grande ouverture à la concurrence des marchés des biens (réforme des professions réglementées, par exemple) et du marché du travail (pour donner davantage de flexibilité aux entreprises tout en sécurisant les parcours professionnels). Dans une économie de la zone euro où la dévaluation monétaire n'est plus possible, l'arme de la dévaluation fiscale (et de la détaxation des facteurs de production comme le travail) peut également être efficace pour stimuler la compétitivité des entreprises.

20 Insee : Institut national de la statistique et des études économiques

CAE : Conseil d'Analyse économique

OFCE : Observatoire français des conjonctures économiques

CEPII : Centre d'études prospectives et d'informations internationales

COE-Rexecode : Centre d'observation économique et de Recherche pour l'Expansion de l'économie et le Développement des Entreprises

BDF : Banque de France

GLOBALISATION FINANCIÈRE, INSTABILITÉ FINANCIÈRE ET RÉGULATION DES MARCHÉS FINANCIERS

« Le risque varie en raison inverse de la connaissance. » - Irving Fischer, 1930

« Le système de détermination des relations financières et d'évaluation des actifs par le marché donne des signaux qui amènent au développement de relations favorables à l'instabilité et à la réalisation de l'instabilité. Les périodes de stabilité (ou de "tranquillité") dans une économie capitaliste moderne sont transitoires. » - Hyman Minsky, 1986

« La globalisation financière est le nom donné à des transformations qui ont affecté les principes de fonctionnement de la finance. Ce sont des transformations très profondes qui associent étroitement la libéralisation des systèmes financiers nationaux et l'intégration internationale. » - Michel Aglietta, 2000

« Les marchés financiers libres sont l'élixir qui alimente le processus de la destruction créatrice, rajeunissant continuellement le système capitaliste. De fait, ils sont aussi la première cible des intérêts puissants qui craignent le changement. » - Raghuram G. Rajan, Luigi Zingales, 2004

19.1 Intégration des marchés de capitaux et globalisation financière

19.2 Le processus de globalisation financière : la règle des 3 « D »

19.3 Les conséquences de la globalisation financière

19.4 De l'instabilité financière à la nécessaire régulation

19.5 Le « shadow banking system » : opportunités et risques

19.6 De régulation micro-prudentielle à la régulation macro-prudentielle

19.1 INTÉGRATION DES MARCHÉS DE CAPITAUX ET GLOBALISATION FINANCIÈRE

19.1.1 *Les différents compartiments du marché des capitaux*

Le marché des capitaux est composé de différents segments interdépendants. On distingue généralement quatre grands compartiments :

– Le marché des changes

Le marché des changes est un marché où se confrontent l'offre et la demande de devises. Le cours des devises est coté en continu par des opérateurs reliés en réseau informatique dans le monde entier (que l'on appelle les « cambistes ») et travaillant dans une salle des marchés bancaires. Le marché des changes est donc un marché virtuel, qui n'a pas de localisation physique. Les entreprises utilisent le marché des changes soit dans le cadre de leurs opérations d'exportation ou d'importation, soit dans le cadre d'opérations financières (achat ou vente de titres étrangers).

– Le marché des actions

Le marché des actions est divisé en un marché primaire, sur lequel les entreprises émettent les titres et un marché secondaire, sur lequel les titres émis sont échangés. Une action est un titre de propriété, dont la rémunération est appelée dividende. L'émission d'actions contribue au financement à long terme de l'entreprise, en complément de l'émission d'obligations ou de l'endettement bancaire.

– Le marché des matières premières

Sur le marché des matières premières, on distingue trois grandes familles de matières premières qui sont échangées

sur les marchés : l'énergie (pétrole, électricité, gaz...), les métaux (or, argent, cuivre...) et les matières premières agricoles (blé, riz, cacao, café...).

– Le marché des taux d'intérêt

Le marché des taux d'intérêt est divisé en de nombreux segments. On distinguera principalement le marché monétaire et le marché obligataire.

*Dans son acception économique, le **marché monétaire** désigne le marché sur lequel sont fixés des taux d'intérêt à court terme (moins d'un an). Il est lui-même subdivisé en trois compartiments : le marché interbancaire, sur lequel les banques se prêtent et s'empruntent entre elles des liquidités sous forme de monnaie banque centrale ; le **marché des TCN** (titres de créance négociables) : **bons du Trésor** (titres à court terme émis par les États et négociés par les investisseurs sur le marché secondaire), **billets de trésorerie** (sur lequel les entreprises émettent des titres de financement à court terme) et **certificats de dépôts**, titres émis par les banques.*

*Le **marché obligataire** est le compartiment à **long terme** du marché des taux d'intérêt. Une obligation est un titre de financement à long terme (maturité à l'émission supérieure à sept ans), dont la rémunération versée à l'acheteur est appelée « coupon » et s'exprime sous la forme d'un taux d'intérêt. De nombreuses formes d'obligations existent, dont les obligations indexées sur l'inflation, souvent émises par les États, offrant aux investisseurs un taux d'intérêt plus bas en contrepartie d'une indexation sur l'inflation du capital remboursé à l'échéance.*

*La plupart des actifs financiers (devises, actions, contrats sur les matières premières, obligations, etc.) sont cotés sur le **marché comptant et sur le marché à terme**. Sur un marché comptant, la cotation est simultanée au transfert de propriété de l'actif. Sur un marché à terme se cotent des prix relatifs à des transactions différées dans le futur. Le rôle originel des marchés à terme est de permettre aux agents économiques*

exposés à un risque de variation de prix de se couvrir. Ainsi, un fabricant européen d'avions attendant le paiement dans deux ans de 70 millions de dollars peut souhaiter les vendre à terme pour se protéger contre la baisse de la devise américaine sur le marché des changes.

L'interdépendance entre les différents segments des marchés de capitaux, associée à la mobilité internationale des capitaux, illustre le phénomène de globalisation financière.

19.1.2 Bref historique de l'intégration des marchés de capitaux

19.1.2.1 Une « première mondialisation » du capital financier (1870-1914)

L'intégration des marchés financiers est un moteur essentiel de la convergence entre des économies riches en capital et des pays ayant des besoins de financement.

À la fin du XIX^e siècle et jusqu'en 1914, le mouvement d'internationalisation de l'économie, tant au niveau des biens que des capitaux, a été aussi important que dans la phase de mondialisation entamée depuis les années 1980.

De 1870 à 1914, les flux financiers internationaux étaient considérables et on a assisté à cette époque à un puissant mouvement d'intégration financière : la Grande-Bretagne et la France sont, en 1913, les principaux exportateurs de capitaux, les États-Unis et la Russie représentent les principaux bénéficiaires. En ce qui concerne la France, si 40 % de la richesse mobilière était déjà constituée de capitaux, près de la moitié était placée à l'étranger avant 1914 (particulièrement en Russie, dans l'Empire ottoman et en Amérique latine). Les investissements directs à l'étranger (IDE) se développent et de nombreuses entreprises anglaises (Dunlop, Unilever), allemandes (Siemens, Bosch, Daimler-Benz), françaises (Saint-Gobain) et suisses (Nestlé) sont déjà implantées à l'étranger.

En 1913, le stock d'investissements directs à l'étranger des pays d'Europe occidentale représente 20-22 % du PIB, soit

un niveau seulement rejoint dans les années 1990. En 1850, la Grande-Bretagne est en position de leadership sur les plans industriel, commercial et financier. Cette période se caractérise par une série de mutations décisives dans l'économie internationale qui tendent à créer un seul marché mondial pour les biens et services, le travail et le capital (accords de libre-échange, système d'étalon-or à même de réduire les incertitudes sur la volatilité des taux de change, législations sur les sociétés anonymes à responsabilité limitée pour favoriser les investissements). On parle d'ailleurs de « première mondialisation » pour désigner la période 1870-1914.

19.1.2.2 La marche vers la globalisation financière à la fin du xx^e siècle

Le gonflement considérable de la sphère financière à partir des années 1970 marque toutefois une nouvelle étape de la mondialisation. Le système monétaire international, fondé lors de la conférence de Bretton Woods en 1944, organisait une fixité des taux de change et une étroite régulation du secteur bancaire et financier. Or, la libéralisation financière à partir de cette époque a créé un environnement monétaire et financier radicalement différent : la dislocation du système monétaire international, l'ère des changes flottants et la révolution informatique ont fortement contribué à l'accroissement du volume et de la vitesse des transactions financières en créant un marché unique de l'argent intégré à l'échelle planétaire, caractérisé par une **double unité de temps** (fonctionnement des marchés financiers en continu 24h sur 24) **et de lieu** (interconnexion des places financières).

La globalisation financière implique un recours accru aux marchés financiers dans le financement des économies, une concurrence plus forte entre les acteurs financiers (banques, institutions financières, fonds spéculatifs) et la possibilité élargie de participer à tous les compartiments de ces marchés. Cette transformation du système financier mondial a été permise à la fois par des initiatives individuelles (les

banques et les institutions financières ont introduit de nouveaux types d'instruments financiers) et publiques (les États ont mis en œuvre de profondes réformes de leur secteur bancaire et financier à partir des années 1980). Progressivement, à l'échelle internationale, les marchés de capitaux se sont standardisés (mêmes produits, mêmes procédés, mêmes intervenants) et interconnectés, grâce en particulier aux technologies de l'information et de la communication (TIC), jusqu'à créer un vaste marché mondial des capitaux.

Le rôle de la finance directe s'est ainsi considérablement accru dans l'économie, tant au niveau des modalités de financement des entreprises qu'à celui des possibilités de placements financiers des ménages.

19.1.2.3 La globalisation financière : un choix politique

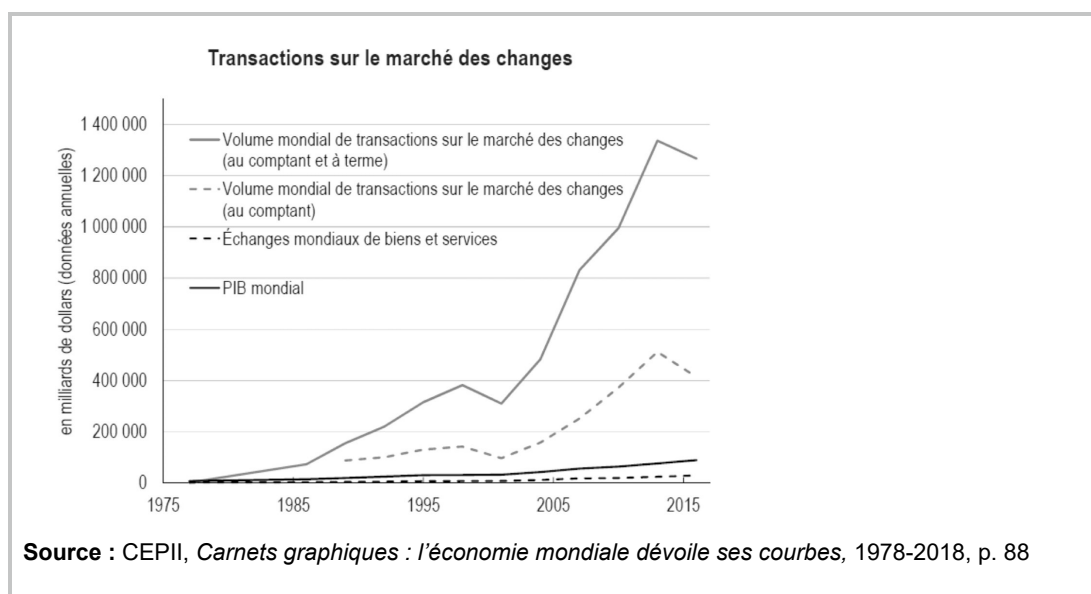
La globalisation financière s'est déployée à partir des années 1980 sous l'impulsion de la finance anglo-saxonne et a fait l'objet d'un consensus entre les gouvernants dans un contexte d'approfondissement des autres dimensions de la mondialisation de l'économie (échanges internationaux de biens et services, multinationalisation des firmes et nouvelle division internationale du travail) et de regain d'intérêt pour les idées libérales, doctrine promue par les grandes organisations financières internationales.

L'économiste John Williamson a ainsi pu évoquer le « consensus de Washington », doctrine défendue par le Fonds monétaire international (FMI) et la Banque mondiale qui ont leur siège situé au sein de la capitale politique américaine. Ces institutions ont diffusé un certain nombre de principes de bonne gouvernance à portée universelle :

- désengagement de l'État de la sphère économique et financière et primauté donnée aux mécanismes du marché (baisse des dépenses publiques et des prélèvements fiscaux, en particulier sur le capital) ;
- incitation à la concurrence, promotion du secteur privé (privatisations des entreprises publiques) ;

- libéralisation de l'environnement financier par la suppression des réglementations afin de favoriser la pénétration des investisseurs étrangers (ouverture des marchés de capitaux).

Les États-Unis, première puissance économique mondiale, ont concilié la mondialisation d'une finance déréglementée (sous l'impulsion de Wall Street et de la concurrence avec la place de Londres) avec l'ouverture aux échanges commerciaux : les modes d'organisation de la finance américaine (déréglementation financière et fiscale), la suprématie du dollar au sein des relations monétaires internationales et la domination des acteurs globaux de la finance américaine (marchés, banques d'affaires, investisseurs institutionnels, etc.) ont conduit à édifier aux États-Unis un modèle de développement économique tiré par la finance et l'endettement.



GRAPHIQUE 19.1.

Transactions sur le marché des changes

De plus, les performances économiques américaines à partir des années 1990 (forte croissance, faible chômage) ont incité les autres États à converger vers les structures du modèle américain, présenté comme une référence en matière

d'efficience par les institutions internationales, au moyen d'une série de profondes réformes de leur système financier (déréglementation du secteur bancaire, innovations financières, abaissement de la fiscalité du capital pour accroître l'attractivité du territoire, etc.). Si l'hégémonie de la finance américaine a ainsi conduit à une certaine homogénéisation du système financier mondial, elle n'a toutefois pas pour autant altéré la variété des systèmes bancaires et financiers nationaux.

Le graphique, ci-dessus, montre la déconnexion croissante entre l'augmentation des échanges de biens et services dans la sphère réelle et la hausse, bien plus forte, des transactions sur le marché des changes, qui accompagne et permet les flux de capitaux.

19.2 LE PROCESSUS DE GLOBALISATION FINANCIÈRE : LA RÈGLE DES 3 « D »

La globalisation financière est un processus d'intégration financière au niveau mondial. La globalisation financière implique un recours accru aux marchés financiers dans le financement des économies, une concurrence plus forte entre les acteurs qui interviennent sur ces marchés et la possibilité élargie de participer à tous les compartiments de ces marchés. Cette évolution a été permise autant que favorisée par trois changements structurels nommés habituellement règle des **3 « D » : Décloisonnement, Déréglementation, Désintermédiation**. La logique des 3 « D » (« décloisonnement » et « dérèglement » des marchés des capitaux et « désintermédiation » des financements) a été mise en évidence par l'économiste français Henri Bourguinat auteur d'un manuel de référence en « Finance internationale ».

Ce processus de globalisation à trois dimensions conduit à la constitution d'un marché unique des capitaux à l'échelle

mondiale, au sein duquel les différents marchés sont complètement interdépendants, en particulier grâce à l'utilisation de réseaux informatiques sophistiqués qui réduisent les coûts de transaction.

19.2.1 *La déréglementation*

La déréglementation se traduit par la suppression de divers obstacles réglementaires aux transactions financières. La **déréglementation** des marchés de capitaux doit favoriser l'essor des flux financiers, l'émergence de nouveaux types de placements et l'influence des investisseurs institutionnels (fonds mutuels, fonds de pension, compagnies d'assurances et fonds spéculatifs ou *hedge funds*). Les pays ont supprimé ou allégé un certain nombre de réglementations : la suppression des contrôles des changes (afin de laisser les mécanismes de l'offre et de la demande agir librement sur le marché des changes), la levée des mesures d'encadrement du crédit, la libéralisation des tarifs bancaires et des taux d'intérêt pratiqués par les banques commerciales. Ce mouvement de déréglementation a accentué la concurrence sur le marché du crédit et dans tout le secteur bancaire (en poussant à l'émergence de « banques universelles » réalisant à la fois des activités de banque de détail et de banques d'investissement, afin d'accroître leur rentabilité). Avec cette déréglementation, les banques ont également dû faire face à l'essor de nouveaux concurrents, tels que les sociétés d'assurance (en particulier dans le domaine de l'assurance-vie), les OPCVM et les fonds de pension.

19.2.2 *Le décloisonnement*

Le décloisonnement consiste en l'ouverture des marchés financiers nationaux à deux niveaux : **sur le plan externe**, par la suppression des entraves réglementaires afin de garantir une complétude des marchés, c'est-à-dire la possibilité donnée à un agent économique de disposer du contrat financier qu'il souhaite (à partir de 1985 a été ouvert en France un marché de titres de créances négociables,

ouvert à tous les agents économiques, qui offre la possibilité à un investisseur ou un emprunteur de choisir la durée de son placement) ; **sur le plan interne**, par une déspecialisation des activités bancaires (effacement de la séparation entre activités de banques de dépôt et de banque d'affaires depuis la fin des années 1990 aux États-Unis).

Le décloisonnement est un processus qui consiste à faire tomber les cloisons, les murs, entre les différents marchés de capitaux afin d'améliorer la possibilité de passer plus facilement d'un produit à un autre. Il y a suppression des barrières à la fois entre les divers marchés nationaux, mais aussi entre les différents compartiments des marchés de capitaux. Cette transformation unifie des marchés jusque-là séparés (marché monétaire, marché obligataire, marché des changes, marché à terme, etc.). Ce décloisonnement des marchés se traduit alors par une interdépendance des prix des actifs financiers, tant au niveau domestique qu'au niveau international. Aux compartiments étanches entre différents marchés (marché monétaire, marché obligataire, marché boursier) se substituent, grâce aux innovations financières (certificat de dépôts, billets de trésorerie, etc.), des possibilités d'arbitrage entre ces différents marchés. La suppression des obstacles à la mobilité interne et internationale des capitaux élargit les possibilités de diversifier son portefeuille de placements financiers.

Ce décloisonnement a été rendu possible par la déréglementation. Celle-ci a renforcé la concurrence financière entre les acteurs et abouti à une grande liberté de circulation de capitaux, en quête des meilleurs rendements au gré des variations des différents paramètres financiers (taux d'intérêt, mais aussi taux de change dans le cadre d'un système de changes flottants à partir des années 1970). Cette évolution a d'ailleurs suscité des innovations financières afin de gérer la double instabilité des taux d'intérêt et des cours de change (en particulier les produits dérivés).

19.2.3 *La désintermédiation*

La désintermédiation est un processus qui consiste à favoriser et développer les relations directes entre agents à capacité et à besoin de financement sans passer par l'intermédiaire des banques. La désintermédiation doit permettre aux agents à besoin de financement de recourir directement aux marchés financiers plutôt qu'au crédit des banques ; en effet, la globalisation financière a favorisé le passage d'une économie dominée par la finance « indirecte », dans laquelle les entreprises sont essentiellement financées par les banques au moyen de crédits bancaires classiques, à une situation de finance « directe », dans laquelle les entreprises se financent davantage par apport de fonds propres ou par émission de titres de créances négociables, directement sur les marchés financiers. Dans le même temps, les banques ont accru leurs interventions sur les marchés financiers, en émettant elles-mêmes des titres financiers. Cette présence accrue des banques sur les marchés financiers (on parle à ce titre de « marchéisation ») a rendu d'ailleurs assez arbitraire la frontière entre l'intermédiation et la désintermédiation.

La désintermédiation donne le choix aux agents non financiers (ménages, entreprises) dans le financement de leurs activités. Ces derniers vont pouvoir comparer les coûts (taux d'intérêt, coûts de transaction, de collecte et de traitement de l'information) et les risques des différents modes de financement.

Les économistes continuent pourtant de distinguer les systèmes fondés sur les financements « intermédiés », et donc sur le crédit bancaire (« bank based »), et ceux reposant principalement sur les marchés de capitaux (« market based »), même s'il existe une forte complémentarité entre les deux.

La désintermédiation financière renvoie aussi à la titrisation qui se traduit par la substitution, dans le bilan des banques, de l'activité traditionnelle de transformation des dépôts en crédits par des opérations de marché. Le mécanisme de la titrisation, soit la transformation des crédits bancaires (ou plus

largement d'actifs peu négociables) en titres financiers négociables sur des marchés secondaires, est l'une des dimensions de la désintermédiation (déplacement du crédit bancaire vers les titres financiers). L'intermédiation financière s'est par ailleurs transformée : si le mouvement de désintermédiation dans le financement des crédits est très important, les banques gardent toutefois un rôle essentiel dans le fonctionnement des marchés et de leur liquidité.

Ainsi, ces trois processus simultanés ont favorisé l'explosion des échanges sur les marchés financiers : il y a de plus en plus d'intervenants qui échangent de plus en plus de produits financiers (le progrès technique favorise l'innovation financière) au niveau mondial. La globalisation financière a abouti à un grand marché financier unique et à des marchés de plus en plus interdépendants. Les intervenants sur les marchés financiers peuvent changer d'actifs, de marchés, de plus en plus vite. Donc le processus de globalisation financière se traduit clairement par une très forte croissance des échanges sur l'ensemble des marchés entre agents à capacité de financement et agents à besoin de financement.

19.3 LES CONSÉQUENCES DE LA GLOBALISATION FINANCIÈRE

19.3.1 *Les avantages attendus de la globalisation financière*

La globalisation financière a entraîné un développement très important des transactions sur les marchés financiers et un essor sans précédent du stock d'actifs financiers, des opérations sur le marché international des capitaux (actions, obligations), ainsi que des transactions sur le marché des changes (devises) et les marchés dérivés.

Cette globalisation financière est censée apporter plusieurs avantages :

- Permettre un agrandissement du choix des épargnants et des investissements et améliorer, au niveau national et

surtout international, l'allocation des ressources financières.

- Favoriser une baisse du coût du capital avec la désintermédiation bancaire et la concurrence internationale entre les grandes banques. Ce mouvement est censé stimuler la croissance et l'emploi.
- Garantir une liberté d'accès aux différents marchés financiers pour permettre à tout opérateur de diversifier son portefeuille d'actifs financiers ; l'offre de nouveaux produits de placement doit permettre de diversifier et de mutualiser les risques pour les investisseurs.
- Accroître la pression des marchés financiers sur les politiques économiques pour favoriser la rentabilité du capital en limitant l'inflation et en contrôlant la dynamique de l'endettement public qui influence le risque de crédit.

La finance globalisée permet, en théorie, de centraliser et de diriger l'épargne disponible vers les projets d'investissement les plus productifs : les institutions financières gèrent les ressources des individus qui épargnent sous la forme de placements financiers (actions, obligations, instruments dérivés) et escomptent percevoir dans l'avenir un certain rendement, qu'il faut rapporter au risque de perte lié à chaque type de produit financier. Une des fonctions de la finance consiste à transférer les risques vers ceux qui souhaitent les assumer moyennant un gain supérieur : la variété des produits financiers disponibles permet aux agents économiques d'optimiser l'arbitrage entre le rendement et le risque, mécanisme sans lequel un certain nombre de projets industriels incertains ne verraient pas le jour, pour une production finale de biens, de services et d'emplois et un flux de revenus finalement plus faibles.

Par ailleurs, en établissant une information publique sur les performances des entreprises en Bourse (valeur des titres), les marchés financiers orientent les choix des individus : ils concourent ainsi à une meilleure affectation des ressources dans l'économie.

L'intensification de la concurrence entre les institutions financières doit ainsi permettre de stimuler les activités productives en améliorant l'allocation de l'épargne mondiale, dirigée vers les projets d'investissement les plus rentables.

19.3.2 *Les risques de la globalisation financière*

– **Une croissance tirée par l'endettement** : la globalisation financière a permis de faire croître le volume des crédits accordés par les banques, d'élever la richesse financière des ménages, de soutenir la consommation et la croissance par une montée de l'endettement (acquisition de biens immobiliers et de biens durables), en particulier aux États-Unis à partir des années 1990, dans un contexte de stagnation des salaires (compensation dette/salaire). Mais ce modèle de développement s'est aussi accompagné du retour de l'**instabilité financière**, d'une sous-évaluation des risques au sein de circuits financiers complexes et fragiles et d'une dynamique préoccupante de l'endettement, jusqu'à la crise des crédits *subprimes*, en 2007.

– **Une perte d'autonomie des politiques conjoncturelles** : la globalisation financière a modifié les conditions de l'action de l'État. Le besoin de financement croissant a poussé les États à lever des fonds prêtables sur les marchés de capitaux par le biais de la vente des titres de la dette publique (comme les bons du Trésor). Les politiques conjoncturelles ont été soumises à une nouvelle contrainte financière : la priorité donnée à la stabilité des prix s'est imposée pour garantir une valorisation suffisante du capital. Dans un contexte de forte mobilité des capitaux, l'accélération de l'inflation anticipée par les marchés financiers (qui détériore le pouvoir d'achat des titres) peut en effet déclencher de brutales sorties de capitaux vers les places financières qui offrent une meilleure rémunération. Les politiques fondées sur des règles (cibles d'inflation à atteindre, normes de déficits et de dette publique) ont pris

alors le pas sur les politiques discrétionnaires centrées sur la demande.

– **Un retour de l'instabilité financière et du risque systémique** : le retour périodique de graves crises financières à partir des années 1990 a mis en évidence un certain nombre de défaillances de la finance moderne. La crise financière internationale qui a éclaté en 2007 a montré que la finance globalisée n'est pas toujours capable de remplir efficacement deux de ses fonctions essentielles : l'allocation optimale des ressources financières à ses usages les plus productifs et la gestion des risques.

Les banques d'investissement prenant en charge la titrisation ont notamment mixé des risques de natures différentes en créant des produits financiers complexes, dont le risque global était difficile à évaluer. Ces innovations financières ont donc accru le niveau global de risque du système financier mondial en le concentrant sur un minimum d'acteurs, en contradiction avec l'argument qui défendait la diversification des risques et leur mutualisation pour les faire porter par les plus aptes et les mieux informés. Au-delà, parce qu'elle facilite l'accès au crédit, la libéralisation financière a contribué à l'accumulation des déséquilibres et accru la variabilité du cycle économique (phase de boom puis phase de dépression).

– **La crise financière internationale de 2007** a plus particulièrement démontré la nécessité de mieux prévenir ce risque systémique, soit le risque qu'un événement particulier (une faillite d'une grande institution financière, un défaut, un choc, etc.) entraîne, par réactions en chaîne, des effets négatifs considérables sur l'ensemble du système financier pouvant occasionner une crise générale de son fonctionnement. Le système financier est alors dans l'incapacité d'assurer ses fonctions habituelles d'allocation de l'épargne et des risques, au point d'affecter fortement la croissance économique et le bien-être social.

19.3.3 Les places financières et les secteurs financiers :

des acteurs stratégiques²¹ pour la croissance et l'emploi

Pour un pays développé inséré dans l'économie globale, posséder une place financière de premier plan constitue un atout essentiel sur le plan stratégique (85 % des flux mondiaux sont gérés par les vingt premières places localisées dans les grandes métropoles des pays les plus riches), afin de contrôler les centres de décision d'allocation des ressources vers le système productif (information, gestion et couverture des risques, assurance). Chaque grande place financière au sein des pays avancés recherche la taille maximale afin de devenir la place de référence dans un univers très concurrentiel, et tente d'attirer les flux de capitaux, les sièges sociaux des grandes sociétés industrielles, les centres de recherche en économie et en finance, et la main-d'œuvre hautement qualifiée (universités, grandes écoles), dans le but de bénéficier de la valeur ajoutée ainsi créée. L'existence d'une place financière de premier plan, dans la mesure où elle favorise le bon développement de l'industrie financière par les effets d'agglomération qu'elle génère, présente un intérêt réel pour l'économie nationale.

Le secteur financier exerce des effets d'entraînement importants sur l'activité et l'emploi : en France, la part du secteur financier dans la valeur ajoutée est ainsi comprise entre 4 % et 5 % depuis quarante ans, soit à un niveau comparable à celui observé en Allemagne, en Italie ou en Espagne. La contribution de l'industrie financière au produit intérieur brut (PIB) y est ainsi dix fois plus importante que celle de l'industrie automobile. Le secteur financier génère ainsi une contribution significative sur le plan budgétaire : les impôts et prélèvements sur la masse salariale acquittés par les entreprises du secteur financier représentent près de 5 %

de l'ensemble des prélèvements obligatoires. Le poids de l'industrie financière n'est pas négligeable en termes d'emploi : les effectifs de la banque et de l'assurance sont estimés à plus de 750 000 emplois temps plein en France.

19.4 DE L'INSTABILITÉ FINANCIÈRE À LA NÉCESSAIRE RÉGULATION

19.4.1 *Instabilité des marchés financiers : interdépendances, finance globalisée et risque systémique*

C'est un constat indiscutable qui n'apporte cependant aucune explication au phénomène. Comment expliquer l'instabilité des marchés financiers ? Quelles peuvent être les causes de l'instabilité financière, c'est-à-dire des prix des actifs financiers ? De manière générale, sur n'importe quel marché, ce sont les variations de l'offre et de la demande qui font varier les prix. Ce sont donc les variations relatives de l'offre et de la demande d'actifs financiers qui font varier les prix des actifs financiers sur les marchés financiers. Il existe de nombreux facteurs permettant d'expliquer l'instabilité de l'offre et de la demande d'actifs financiers et donc la volatilité des prix de ces actifs.

La globalisation financière implique un recours accru aux marchés financiers dans le financement des économies, une concurrence plus forte entre les acteurs financiers (banques, institutions financières, fonds spéculatifs) et la possibilité élargie de participer à tous les compartiments de ces marchés, de passer facilement d'un marché à un autre.

La globalisation financière s'est accompagnée d'une instabilité accrue des prix des actifs financiers liée à l'explosion des échanges de capitaux avec la croissance exponentielle du nombre d'intervenants et de produits financiers. Les très fortes variations observables de l'offre et de la demande des différents actifs financiers entraînent une

très grande instabilité des prix de ces actifs. De surcroît, du fait des interdépendances accrues entre produits financiers et marchés, l'instabilité sur certains marchés entraîne une instabilité sur d'autres marchés. Le crédit et donc la création monétaire peuvent alimenter et amplifier cette instabilité financière, car des agents empruntent pour acheter des actifs puis les vendre afin d'en acheter d'autres. Ce fut le cas avant la crise des *subprimes* de 2007. *In fine*, la combinaison excès de crédit-instabilité financière peut favoriser l'occurrence de crises financières qui provoquent ensuite des crises économiques et sociales qui préparent le terrain pour d'autres crises financières, dans un futur plus ou moins proche.

L'horizon souvent court des investisseurs, lié à l'incomplétude informationnelle à long terme (l'avenir est incertain), et la myopie des agents sur les différents marchés se traduisent par des comportements « panurgiens » (moutonniers ou mimétiques).

Le rôle central des anticipations qui sont donc de moins en moins rationnelles (sauf si la rationalité consiste à adopter le même comportement que les autres), la diversification des produits financiers et notamment la multiplication des produits dérivés (de produits financiers dits « sous-jacents ») conjuguent leurs effets et conduisent à exacerber les phénomènes de « surréaction » et « d'exubérance irrationnelle » (Robert Shiller), se traduisant par l'envolée des cours de différents produits.

Ces bulles spéculatives, alimentées par le fait que la hausse entraîne la hausse, finissent toujours par atteindre un seuil au-delà duquel la baisse prend la forme d'une chute brutale et vertigineuse des cours. C'est le krach. Après la phase de « boom », vient... le krach !

Malheureusement, le krach financier va déborder les frontières de la finance, inonder tel un tsunami toute la planète financière et engloutir ensuite l'économie réelle. Il en résulte que la crainte d'une véritable crise systémique, lorsqu'un déséquilibre profond apparaît dans un

compartiment du système avant de se propager, comme dans un jeu de domino, au reste du monde, devient tout à fait légitime.

Selon une définition de référence, proposée par les économistes Michel Aglietta et Philippe Moutot, le **risque de système** désigne « *l'éventualité qu'apparaissent des états économiques dans lesquels les réponses rationnelles des agents individuels aux risques qu'ils perçoivent, loin de conduire à une meilleure répartition des risques par diversification, amènent à élever l'insécurité générale* ».

Selon ces auteurs, cette définition permet d'insister sur la **non-neutralité des structures financières**, qui influencent le champ des possibles en matière d'équilibre macroéconomique. Autrement dit, de graves dysfonctionnements dans l'allocation des ressources financières (par exemple, l'octroi de crédits à des ménages peu solvables avec les *subprimes*...) peuvent conduire à des équilibres macroéconomiques inefficients. Mais ces équilibres ne proviennent pas (seulement) de l'irrationalité des agents, mais surtout des défauts de la coordination de milliers de choix individuels (microéconomiques) par les marchés financiers. D'où l'existence d'un résultat macroéconomique « anormal », c'est-à-dire la survenue d'un risque systémique.

19.4.2 Principes de la régulation des marchés financiers

19.4.2.1 Le contexte : crise financière et risque systémique

Les crises monétaires et financières internationales comme l'instabilité des taux de change posent la question d'une nouvelle architecture monétaire et financière internationale. Certains économistes, plutôt confiants dans le caractère régulateur du marché, considèrent qu'il faut simplement mieux encadrer le fonctionnement actuel du système financier international. Pour d'autres, il s'agit de mettre en place un nouveau régime monétaire international cohérent pour pallier

les insuffisances et le caractère déstabilisateur des marchés de capitaux.

La régulation des marchés financiers peut se définir par un **ensemble de règles, de pratiques et d'encadrement des activités financières** dont les objectifs prioritaires sont le bon fonctionnement du marché (transparence et intégrité), la protection des dépôts des épargnants et la stabilité du système financier. Cette régulation s'exerce par l'intermédiaire de diverses autorités de régulation nationales, européennes ou internationales. Elle peut être micro- ou macro-prudentielle : surveillance du marché et des intermédiaires financiers, qui repose sur la quantification du risque et la mise en place de réglementations destinées à minimiser les risques de faillite ou de dysfonctionnement du secteur financier.

En raison de la responsabilité particulière des institutions financières et des banques dans la crise économique et financière de 2007-2008, de nombreuses voix se sont élevées pour exiger une réglementation plus stricte des activités financières et une coordination plus forte à l'échelle internationale. Cette exigence a donné lieu à des réformes en faveur de la stabilité financière. La Banque Centrale européenne (BCE) définit la stabilité financière comme « *une situation où le système financier – comprenant les intermédiaires, les marchés et les infrastructures de marchés financiers – est capable de faire face à des chocs* ».

La stabilité financière comprend ainsi trois caractéristiques :

- Le système financier devrait être capable de transférer des ressources de façon régulière des épargnants vers les investisseurs.
- Les risques financiers devraient être évalués et valorisés de manière relativement précise et devraient être relativement bien gérés.
- Le système financier devrait être dans un état où il peut confortablement absorber des surprises et des chocs financiers qui impacteraient l'économie réelle.

L'objectif principal de la stabilité financière est de limiter le coût des crises financières et bancaires qui ont des conséquences potentiellement dévastatrices en raison du rationnement du crédit, à l'origine d'une chute de la demande globale.

Après la crise financière internationale qui a éclaté en 2007, le thème de la régulation financière est revenu sur le devant de la scène : cette crise majeure du capitalisme, la plus grave depuis 1929, a montré l'importance d'une coopération internationale lorsque les États se trouvent confrontés à ce type de crise systémique, qui menace l'intégrité de tout le système financier mondial. L'objectif général de la réglementation bancaire et financière est simple : créer et maintenir la confiance dans chaque pays et dans le système global. Cette **confiance** est ce qu'on appelle un ***bien public***, c'est-à-dire un bien dont tous les citoyens bénéficient, la confiance dans les banques permettant le financement de l'économie à moindre coût. La question de la régulation financière s'inscrit aujourd'hui dans le problème plus vaste de la gouvernance mondiale : la finance est globalisée, mais la réglementation financière ne l'est pas.

En définitive, l'idée de régulation financière suppose implicitement que chaque État se trouve dans l'incapacité d'apporter une réponse satisfaisante à certains risques globaux, comme celui de la propagation d'un choc financier au reste de la planète.

En matière économique et financière, la régulation peut être appréhendée à partir de trois paramètres :

- une intervention publique, mais indépendante du gouvernement ;
- une action sur un ou plusieurs segments des marchés financiers ;
- La création d'un cadre sécurisé de concurrence pour les activités financières.

Pour de nombreux experts, la crise financière qui a éclaté en 2007 est le résultat de la défaillance de la régulation et de la

multiplication de situations d'aléa moral sur les marchés financiers, lorsqu'il existe des incitations, pour les acteurs privés, à prendre des risques individuels si ceux-ci sont assurés, transférés ou supportés de manière collective, c'est-à-dire lorsqu'ils exercent des externalités négatives. Si l'environnement macroéconomique a pu créer un cadre favorable aux excès du secteur bancaire et financier durant les années 2000 (déficits de la balance des paiements courants des États-Unis, excès d'épargne en Asie à la recherche de placements financiers diversifiés, politique monétaire expansionniste et faibles taux d'intérêt favorables à l'endettement), ce sont également les dysfonctionnements des règles des marchés, du contrôle de leur application et les comportements des acteurs (banques, fonds d'investissement, opérateurs sur les marchés) qui expliquent les dérèglements financiers.

La crise financière a démontré les failles en matière de régulation du système financier mondial, notamment en ce qui concerne le contrôle des risques encourus dans le cadre des innovations financières, en plein essor depuis les années 1990-2000 (comme les produits dérivés et la titrisation).

19.4.2.2 La quête d'une meilleure régulation financière après la crise

La mise en évidence de l'instabilité des marchés financiers et la récurrence des crises financières ont conduit à un débat sur la réforme du système financier, voire, plus largement, sur la nécessité de mettre en place une nouvelle architecture financière internationale. Les propositions de réformes débattues au sein du G20, comme au sein de l'Union européenne depuis la crise des *subprimes*, ont penché en faveur d'une approche plus volontariste de la régulation financière. Le renforcement des mesures de contrôles des banques, des assurances et des marchés financiers en Europe s'inscrit dans une logique visant à limiter l'instabilité financière, à lutter contre les comportements d'aléa moral et

contre les externalités négatives qui résultent de l'instabilité financière. La concentration du système bancaire et le risque d'aléa moral se sont accrus : certaines banques sont, plus que par le passé, « trop grosses pour faire faillite » (*too big to fail*). La recherche d'une plus grande stabilité financière et d'un système financier davantage contrôlé pourrait passer par une coopération internationale plus étroite des politiques économiques afin de limiter les déséquilibres financiers internationaux, un renforcement des règles prudentielles, voire, dans une approche plus volontariste, la mise en œuvre d'une taxation des mouvements de capitaux pour juguler les mouvements spéculatifs, ou l'établissement d'un prêteur en dernier ressort international. Les pays avancés ont connu depuis les années 1980 une succession de crises d'une intensité inédite et croissante depuis la « Grande Dépression » des années 1930 : ces crises financières ont largement impacté l'économie réelle, avec des coûts importants en termes de croissance et d'emplois détruits. L'instabilité financière a conduit à s'interroger sur le poids pris par le secteur financier sur l'organisation et le rôle de celui-ci, mais elle pose également la question des modalités de régulation et de contrôle de la finance par la puissance publique, ainsi que sur certains déséquilibres macroéconomiques.

Néanmoins, des débats subsistent sur la pertinence d'une régulation plus étroite du secteur financier, comme le note l'économiste Pierre-Noël Giraud : « *Les évolutions depuis quarante ans ont accru l'instabilité. Pour quel bénéfice et au bénéfice de qui ? Les défenseurs de la finance de marché affirment qu'elle remplit mieux ses fonctions que la finance bancaire cloisonnée qu'elle a remplacée, et donc permet une croissance moyenne supérieure. Ses détracteurs soulignent son instabilité et nient, à cause des crises qu'elle provoque, qu'elle stimule la croissance moyenne.* » (*Principes d'économie*, La Découverte, 2016)

Un certain nombre de réformes ont eu lieu depuis la crise financière, avec les accords du comité de Bâle, les standards

du Conseil de stabilité financière et divers rapports comme le rapport Vickers au Royaume-Uni, tandis que diverses législations ont été votées, soit à l'échelle nationale (Dodd-Frank Act aux États-Unis, loi de réforme bancaire au Royaume-Uni, loi de régulation séparation et de régulation des activités bancaires en France en 2013, etc.), soit à l'échelle européenne (directives CRD IV, révision de la directive sur les systèmes de garantie des dépôts, directive sur le redressement et la régulation des crises bancaires, etc.) D'autre part, la nécessité de renforcer la coopération internationale en matière de régulation financière a donné naissance à certaines institutions comme, en 2009, le Conseil de stabilité financière (*Financial Stability Board*) qui vise à coordonner le développement de standards internationaux de régulation financière. Le G20 a renouvelé sa volonté de lutter contre l'opacité générée par le *shadow banking system*, qui concourt à accroître l'instabilité financière, ainsi que contre les paradis fiscaux.

On peut dès lors présenter les institutions de la régulation financière plus en détail.

19.5 LE « SHADOW BANKING SYSTEM » : OPPORTUNITÉS ET RISQUES

19.5.1 Les contours du « shadow banking system »

Face aux nouvelles régulations qui touchent une partie du système financier, on assiste à la croissance d'un système opaque et non régulé : le « shadow banking system ». L'expression « shadow banking system » a été utilisée pour la première fois et dans un sens péjoratif en 2007 par Paul Mc Culley, un analyste de Pim-co. Les régulateurs l'ont défini depuis de manière plus neutre comme **un système d'intermédiation de crédit englobant des entités et des activités qui ne font pas partie du secteur bancaire classique** (comme certains fonds d'investissement). Il relève

cependant de la « **finance de marché** » car il fournit à l'économie des ressources que les banques ne peuvent pas toujours apporter, souvent en raison de contraintes réglementaires (le « shadow banking system » a la possibilité de contourner les ratios qui limitent l'effet de levier). Les acteurs du « shadow banking system » sont les banques d'affaires, les hedge funds, les fonds de titrisation, les fonds monétaires, les fonds de pension, mutuels, d'assurance-vie, mais aussi les entreprises de capital-investissement, les sociétés de garantie de crédit, les trusts de gestion d'actifs (immobiliers, par exemple). Dans un sens encore plus large, le « shadow banking system » peut regrouper les établissements de crédit, les sites de crowdfunding (financements participatifs), les plateformes de monnaies virtuelles (bitcoins, par exemple). Ces entités, à la différence des banques traditionnelles, ne reçoivent pas de dépôts et ne fonctionnent que sur le crédit.

Difficile à cerner, on peut approcher le « shadow banking system » par les entités qui le constituent ou les activités qu'il développe. Ces entités 1) collectent des capitaux présentant des caractéristiques similaires à celles des dépôts ; 2) réalisent des opérations de transformation d'échéance ou de liquidité ; 3) permettent des transferts de risques de crédit et d) ont recours au levier d'endettement, directement ou indirectement. En 2015, le Conseil de stabilité financière (Financial Stability Board) a introduit dans son rapport annuel une mesure en fonction des fonctions économiques et non des entités concernées. Cette approche plus étroite du « shadow banking system » a conduit à limiter la taille estimée de ce segment de la finance (pour 2014, il passait ainsi de 80 000 milliards de dollars à 32 000 milliards).

19.5.2 *Les risques du « shadow banking system »*

Conformément aux recommandations du Conseil de stabilité financière, la Commission européenne a publié en 2012 un

livre vert consacré au « shadow banking system », qui évoquait les orientations prioritaires à concrétiser pour qu'il soit plus efficacement régulé. Ce document insistait en particulier sur l'interconnexion entre le « shadow banking system » et le système bancaire traditionnel, dans la mesure où le financement du premier dépend largement du second. De manière plus globale, les risques relatifs au « shadow banking system » ont quant à eux été détaillés de la manière suivante par le CSF : 1) risques de désengagements massifs et brutaux des structures de financement, similaires au risque de *runs* auquel les banques font face ; 2) accumulation importante de risques liés à l'effet de levier ; 3) risque d'arbitrages des banques vers le secteur bancaire parallèle pour échapper à la réglementation et aux ratios de liquidité et de solvabilité et 4) danger de défaillances non ordonnées affectant le système bancaire. Le Comité de Bâle, à la demande du CSF, a élaboré plusieurs mesures pour parvenir à maîtriser les risques dans les interactions entre le système bancaire classique et le « shadow banking system ».

Pour cela, de nouvelles exigences prudentielles ont été finalisées, tandis que le cadre de surveillance déjà en place pour mesurer et contrôler les expositions importantes aux contreparties a été élargi aux entités du « shadow banking system ». Il est également prévu de soumettre les banques à des limites d'allocation des fonds propres plus strictes d'ici le 1^{er} janvier 2019. Le régulateur devra trouver en la matière un équilibre subtil entre la réduction des risques que le « shadow banking system » fait peser sur la stabilité financière et le soutien à l'innovation financière, pour ne pas compromettre les sources additionnelles de financement qu'il représente pour la croissance de l'économie « réelle ».

19.6 DE RÉGULATION MICRO-PRUDENTIELLE À LA RÉGULATION MACRO-PRUDENTIELLE

Schématiquement, on peut représenter et classer les objectifs des politiques micro-prudentielles et macro-prudentielles de la manière suivante :

TABLEAU 19.1.

Les objectifs de la réglementation financière

• Protéger le déposant • Protéger l'investisseur	} Politique micro-prudentielle
• Résoudre les frictions/imperfections financières : – Asymétries d'information : transparence, publications des résultats, ratings – Concurrence imparfaite : lois anti-monopoles	
– Externalités : ventes à découvert ; ventes forcées d'actifs ; risque systémique	} Politique macro-prudentielle

Source : Banque de France

La régulation financière vise en particulier à lutter contre le déséquilibre né de l'essor des marchés de capitaux et leur internationalisation, et le caractère souvent national des règles existantes. Traditionnellement, on distingue la *régulation micro-prudentielle*, soit la quantification du risque de faillite d'une entité financière et les réglementations afin de limiter les risques de défaillance individuelle, de la *régulation macro-prudentielle*, c'est-à-dire la quantification du risque et définition d'un cadre réglementaire qui concerne le système financier pris dans son ensemble, afin de limiter les risques de crises financières et leurs conséquences sur la croissance économique.

19.6.1 *La régulation micro-prudentielle des risques financiers*

La crise financière de 2007-2008 a posé le problème de la qualité des règles prudentielles. Ces dernières sont des dispositions contraignantes adoptées par les autorités monétaires visant à prévenir les faillites bancaires et à

assurer le bon fonctionnement des marchés financiers. Ces règles prudentielles, en tant que composante de la politique monétaire, ont notamment pour objectif de consolider la solvabilité et la liquidité des opérateurs financiers et d'éviter les phénomènes de contagion liés aux paniques financières. Le renforcement de ces règles depuis le krach de 1987 et le développement de l'activité des autorités de régulation n'ont pas empêché la crise financière majeure de 2007. Celle-ci résulte de la mauvaise qualité de l'information et d'une grave insuffisance dans la gestion des risques au niveau de chaque banque ou institution financière, comme du système financier dans son ensemble. Les règles prudentielles ont été réformées par les accords de Bâle III de décembre 2010 : elles prévoyaient que les banques devaient respecter de nouveaux ratios afin de prévenir les risques de solvabilité, de pro-cyclicité et de liquidité. Ils reposent sur quelques mesures principales : des fonds propres renforcés en quantité et en qualité (le ratio atteindra 10,5 % de fonds propres en 2019, dont 7 % de fonds propres « durs », actions et réserves) ; un ratio de levier simple (non pondéré par les risques) de 3 % ; des ratios de couverture de liquidité de court terme et de long terme plus élevés. La consolidation des systèmes bancaires constitue une condition indispensable d'un retour à une croissance équilibrée dans les pays développés : un système bancaire efficient et stable est nécessaire à la production des services de paiement et à la mobilisation de l'épargne, aux meilleures conditions de qualité, de coût et de sécurité, et les banques jouent d'ailleurs un rôle majeur pour la distribution du crédit et l'intermédiation de marché dans l'allocation du capital. La régulation micro-prudentielle a pour objectif de faire en sorte que les banques évaluent, gèrent et tarifient correctement les risques - et les assument pleinement.

Les accords de Bâle III ont cherché à corriger les dispositifs de contrôle prudentiel par un accroissement du niveau et de la qualité des fonds propres, l'extension de la couverture des risques et la prise en compte du risque de liquidité.

- L'objectif de la *réglementation des fonds propres bancaires* est de préserver la capacité d'absorption des pertes et de garantir ainsi la pérennité des établissements de crédit. Il s'agit aussi de limiter l'aléa moral, car la crise financière a prouvé que la réglementation en vigueur était défailante puisque les États ont été contraints de mobiliser des ressources de la collectivité pour recapitaliser et offrir diverses garanties à leur secteur bancaire.
- En matière d'*évaluation des risques*, les précédents accords de Bâle II avaient établi une relation entre le montant du capital requis d'un établissement bancaire et le niveau des risques portés par ses actifs, afin de réduire l'incitation des banques à la prise de risques. Mais la crise financière a montré également leurs limites, non seulement le caractère pro-cyclique de la réglementation (avec moins de fonds propres quand la situation est bonne, et plus de fonds propres quand la situation se dégrade, avec des capitaux alors rares et plus coûteux), mais également la sous-pondération de certains risques (comme les risques de marché ou liés aux produits complexes et liés à la titrisation).
- Des *normes sur le risque de liquidité* sont également introduites par les accords de Bâle III : ainsi la banque doit sélectionner des actifs facilement cessibles sans perte de valeur pour alimenter sa trésorerie en cas de difficulté à cause de retraits massifs de sa clientèle ou de l'assèchement du marché interbancaire. Il est également demandé aux banques de pondérer leurs actifs selon la qualité du risque : ainsi, une augmentation du risque de contrepartie ou le développement des activités de marché devront être compensés par davantage de fonds propres. En outre, un ratio pour plafonner l'effet de levier, soit le rapport entre le total des actifs et les fonds propres, a été introduit afin de conjurer les risques liés à l'endettement (les capitaux propres doivent représenter au moins 3 % du total de l'actif d'une banque).

En matière de réglementation de la liquidité, une innovation importante de Bâle III a été l'introduction de deux contraintes de liquidité : la première, le *liquidity coverage ratio* (LCR), impose aux établissements de disposer d'actifs liquides suffisants pour assurer leurs engagements sur un mois dans l'hypothèse d'une situation de crise sur les marchés monétaires. La seconde, le *net stable funding ratio* (NSFR), entend faire respecter un équilibre entre la liquidité des actifs et les exigibilités des passifs (de bilan comme de hors bilan) à l'horizon d'un an. Dans les deux cas, il s'agit bien de limiter la dépendance des banques vis-à-vis du marché monétaire pour éviter que ne se reproduise une situation capable de bloquer le crédit et le marché interbancaire comme au plus fort de la dernière crise financière de 2007. Bâle III demeure un dispositif essentiellement micro-prudentiel axé sur la prévention des risques individuels des établissements bancaires, mais l'on sait que la dimension macro-prudentielle de la régulation financière est toutefois essentielle face au risque systémique.

19.6.2 Les objectifs de la régulation macro-prudentielle

19.6.2.1 La lutte contre le risque systémique

Si toutes les politiques économiques peuvent contribuer à promouvoir la stabilité financière, la politique macro-prudentielle a été développée dans le but explicite et principal d'assurer la stabilité de l'ensemble du système financier et d'éviter la matérialisation du risque systémique. Ce type de risque affecte généralement le système financier dans son ensemble (ou une large part de celui-ci) et peut se manifester au travers des effets de contagion, d'amplification des chocs (par exemple, en raison des expositions à un risque commun de plusieurs acteurs du système financier) ou des déséquilibres financiers (comme l'existence de bulles ou une croissance excessive du crédit). La crise financière a en effet

contribué à la prise de conscience de l'insuffisance des politiques jusqu'alors mises en œuvre pour prévenir ce type de risque. Il s'agit donc d'abord d'ajouter une perspective macroéconomique à la supervision du système financier, laquelle n'était jusqu'à présent abordée que sous un angle « micro ». Comme la crise l'a illustré, la stabilité financière ne dépend pas seulement de la stabilité des composantes du système financier prises individuellement : elle est liée également aux interdépendances et interactions complexes qui unissent ces composantes. Le terme « macro » renvoie à l'interaction entre les sphères réelle et financière, un risque ne devenant proprement « systémique » qu'à partir du moment où les déséquilibres ou les chocs qui affectent le système financier menacent, de façon significative, l'activité économique. L'Union européenne s'est ainsi dotée d'un Conseil Européen du Risque Systémique (CERS), qui dépend étroitement de la BCE. Pour mieux intégrer l'objectif de stabilité financière, elle est désormais responsable de la supervision des grandes banques de la zone euro dans le cadre de l'Union bancaire. Les banques centrales intègrent ainsi l'objectif de stabilité financière. La Banque Centrale européenne (BCE) définit ainsi la régulation macro-prudentielle comme une stratégie qui comporte trois grandes dimensions :

- prévenir l'accumulation excessive de risques, liée à des facteurs externes et des défaillances du marché, afin de lisser le cycle financier (dimension temporelle) ;
- renforcer la capacité de résistance du secteur financier et limiter les effets de contagion (dimension transversale) ;
- favoriser une perspective à l'échelle du système dans le cadre de la réglementation financière en vue d'instaurer un ensemble approprié d'incitations pour les intervenants de marché (dimension structurelle).

La question de la surveillance des évolutions du prix des actifs, et par là la prévention des bulles spéculatives, est désormais posée. Dans son rapport de 2015, le Haut Conseil

de Stabilité Financière (HCSF) français évoquait la régulation macro-prudentielle et les facteurs de risques qui, individuellement ou conjointement, peuvent conduire à une fragilisation du système financier. Il retenait notamment une expansion du crédit et un effet de levier excessifs au sein du système financier comme au sein des secteurs non financiers ; un trop fort degré de transformation (de maturité, de liquidité) opéré par les acteurs financiers ; une concentration excessive des expositions directes et indirectes des acteurs financiers ; des incitations inappropriées conduisant à un aléa moral, y compris les conséquences potentielles du caractère (individuellement ou collectivement) systémique de certains acteurs. Ces constats ont conduit à la mise en œuvre en France à partir de janvier 2016 du « coussin de fonds propres contra-cyclique », dans le cadre de la directive européenne CRD IV. C'est une surcharge en fonds propres qui s'ajuste dans le temps, de sorte que les exigences en fonds propres s'accroissent en période d'accélération du crédit au-delà de sa tendance et se desserrent en période de ralentissement (le taux additionnel atteint son niveau maximal de 2,5 % lorsque le « *credit-to-GDP gap* » - écart du ratio crédit/PIB – est supérieur ou égal à 10 points de pourcentage).

Les actions visant à assurer la stabilité financière sont devenues l'une des priorités du G20, dès le sommet de Londres en avril 2009. On retient généralement que la régulation macro-prudentielle a un double objectif : d'une part, lutter contre la pro-cyclicité des systèmes financiers, soit l'effet par lequel les systèmes financiers peuvent amplifier les cycles économiques, que ce soit en favorisant l'apparition de phases d'expansion excessives au cours desquelles les risques sont sous-évalués et s'accumulent ou, symétriquement, en amplifiant les perturbations dans les phases de contraction du fait d'une brutale et excessive aversion aux risques (la régulation macro-prudentielle n'a pas pour but d'éliminer tout cycle financier, mais d'éviter une volatilité et une ampleur excessives du prix des actifs

financiers) ; d'autre part, la politique macro-prudentielle vise à renforcer la résilience du système, soit sa capacité à absorber les chocs financiers. Pour cela, la politique macro-prudentielle cible notamment les institutions d'importance « systémique » (le Conseil de stabilité financière met ainsi à jour une liste de *Global Systemically Important Financial Institutions* [G-SIFIs], c'est-à-dire les institutions dont la faillite mettrait en danger l'ensemble du système financier), et elle s'attache à limiter les phénomènes de défaillances collectives, qu'ils soient liés à des interconnexions fortes ou à des expositions massives à un risque identique. La prise en compte du risque systémique conduit aussi à compléter le volet micro-prudentiel par des politiques macro-prudentielles.

19.6.2.2 L'interaction avec les autres politiques macroéconomiques

La politique macro-prudentielle n'est pas neutre pour les autres politiques économiques. Il peut exister des complémentarités ou des freins entre les différentes autorités et les niveaux d'intervention :

- La politique macro-prudentielle est ainsi très liée à la politique budgétaire : par exemple, l'endettement excessif du secteur public peut renforcer la vulnérabilité du système bancaire qui se trouve davantage exposé à un risque souverain ou bien les garanties d'intervention publique en cas de crise peuvent créer un phénomène d'aléa moral.
- La régulation macro-prudentielle est aussi liée à la politique monétaire, car en limitant la propagation des chocs sur le prix des actifs ou sur l'offre de crédit, elle peut renforcer le mécanisme de transmission de la politique monétaire, tandis que les politiques monétaires accommodantes, quant à elles, peuvent favoriser la formation de bulles spéculatives et compromettre la stabilité financière.
- Enfin, la politique macro-prudentielle est étroitement associée à la politique micro-prudentielle dont elle partage certains outils (ratios de capitaux ou de liquidité), même si

les objectifs des deux niveaux d'intervention sont différents : au cours d'une crise bancaire, l'autorité micro-prudentielle peut avoir intérêt à renforcer les exigences de certaines banques identifiées comme à risque. Mais cela peut amplifier la contagion en imposant un ajustement plus rapide et simultané du secteur bancaire, et renforcer le resserrement de l'offre de crédit. Ainsi une autorité macro-prudentielle privilégiera davantage un assouplissement des contraintes réglementaires pour éviter un ajustement trop brutal dans le court terme.

19.6.3 *Les limites de la régulation financière*

La globalisation financière, en organisant une hyper-connexion des marchés de capitaux, a accru la vitesse de propagation des crises à l'échelle mondiale et, malgré d'indéniables progrès, la régulation financière internationale peine à s'imposer dans un espace politique fragmenté en diverses législations nationales. Un obstacle important au développement d'une régulation financière internationale tient au principe de souveraineté qui régit les rapports entre États dans le cadre des relations internationales, aussi bien dans le cadre de l'Union européenne qu'au niveau mondial. Face à l'impossibilité d'élaborer une régulation financière totalement intégrée au niveau mondial, trop ambitieuse, les États ont choisi d'avancer par la concertation et le consensus grâce à des instances souples, qui cherchent avant tout à persuader plus qu'à imposer (conseils, forums, comités, etc.). Ces diverses instances ont abouti, de manière pragmatique, à l'élaboration d'un ensemble de référentiels et à de la *soft law*, c'est-à-dire des codes, des recueils de bonnes pratiques (de « bonne gouvernance »), un ensemble de lignes directrices à suivre, dont la mise en œuvre reste de la compétence des institutions nationales. En particulier, il n'est pas prévu de sanctions de nature juridique en cas de non-respect, ni par un quelconque organisme international, ni par des tribunaux internationaux. Ces codes de « bonne gouvernance » ne nécessitent pas de modification substantielle des traités

internationaux, mais n'offrent pas encore toutes les garanties de transparence ou d'obligation de rendre des comptes sur les progrès en matière de régulation financière.

À l'avenir, la nouvelle architecture financière internationale devra s'établir en maintenant la coopération entre les États, malgré des intérêts divergents face à l'instabilité financière, au risque systémique ou face aux conflits potentiels sur le front des taux de change.

L'analyse économique justifie la nécessité d'un « prêteur international en dernier ressort » (PIDR), dont le rôle pourrait être dévolu au FMI, en raison de l'existence d'effets de contagion qui s'expliquent par la volatilité des marchés financiers, l'importance des anticipations et l'absence de coordination entre les pays. Mais la création de ce prêteur international en dernier ressort bute sur le risque d'aléa moral si les grandes institutions financières « systémiques » acquièrent la certitude qu'elles seront secourues même en cas de risques financiers excessifs.

21 Le secteur financier est un déterminant de l'attractivité d'un territoire. Cf. chapitre 15.

CRISES FINANCIÈRES, CRISES ÉCONOMIQUES ET INTERVENTIONS DES POUVOIRS PUBLICS

« La finance détruit le commerce par ses injustices, par ses vexations, par l'excès de ce qu'elle impose. » - Montesquieu, 1748

« Lors des crises financières passées - l'effondrement de la bourse en 1987, les conséquences du défaut de paiement de la Russie en 1998 - la Fed a pu utiliser sa baguette magique et disperser les nuages sur les marchés. Mais, cette fois-ci, la magie ne fonctionne pas. Pourquoi ? Parce que le problème des marchés n'est pas seulement celui d'un manque de liquidité. Il y a également un problème fondamental de solvabilité. » - Paul Krugman, 2007

« La "finance de PONZI" est une situation où la solvabilité d'un agent économique (d'un emprunteur) n'est pas assurée par ses flux futurs de revenus, qui ne couvrent pas sa dette présente. Elle n'est assurée que s'il peut se réendetter périodiquement pour faire face au service de sa dette, par exemple s'il détient des actifs dont le prix augmente à un rythme supérieur au taux d'intérêt et qui servent de garantie à cet endettement. Mais cette situation n'étant pas durable, la finance de Ponzi finit normalement en un défaut de l'emprunteur. » - Patrick Artus, 2007

« La crise actuelle marque la fin d'une période d'expansion du crédit basée sur un dollar jouant le rôle de monnaie de réserve internationale. Les crises périodiques précédentes résultaient de processus d'emballement puis d'éclatement

(boom et bust). Celle-ci marque la fin d'un super-boom économique qui a duré 60 ans. » - Georges Soros, 2008

SOMMAIRE

[20.1 Des crises financières aux crises de l'économie réelle](#)

[20.2 Nature et mécanismes des crises financières](#)

[20.3 Les réponses des autorités publiques face à la crise financière et économique](#)

[20.4 LA ZONE EURO ET LA CRISE DES
DETTES souveraines : crise de financement des États](#)

[20.5 L'action DES États et de la BCE face aux différentes
dimensions de la crise des dettes souveraines](#)

[20.6 Les évolutions institutionnelles de la zone euro depuis
2010 et les perspectives](#)

20.1 DES CRISES FINANCIÈRES AUX CRISES DE L'ÉCONOMIE RÉELLE

20.1.1 *Interdépendances entre sphère financière et sphère réelle : l'illusoire déconnexion*

La question de la connexion et de la déconnexion entre la sphère financière et la sphère réelle de l'économie est pourtant une question récurrente, que se posent les économistes, les dirigeants politiques, les dirigeants d'entreprises, les salariés, etc. La réponse simple est que la déconnexion ne peut être que partielle, temporaire et contingente. Ce que nous observons, depuis 2007 et surtout 2008, semble nous montrer clairement et cruellement que la déconnexion n'est qu'une illusion d'optique qui s'explique davantage par les lunettes idéologiques que par les lentilles de la théorie confrontée aux analyses empiriques. Les erreurs de choix et de comportements des agents sur les marchés financiers et les déséquilibres financiers associés ont des conséquences majeures sur les acteurs et leurs opérations dans l'économie réelle. Les faillites bancaires, les défaillances successives d'organismes de crédit, les licenciements, les ménages devenus insolvable et en cessation de paiement, la montée du chômage... sont autant de preuves indiscutables de l'interdépendance entre les sphères réelle et financière de l'économie. Toutes les crises financières, qu'il s'agisse d'une crise de change, d'une crise bancaire, d'une crise de la dette ou d'une crise boursière sont susceptibles de provoquer une récession, voire une dépression, dans la sphère réelle de l'économie.

Ces interdépendances ne devraient pourtant surprendre personne, car toutes les activités économiques, dans une

économie capitaliste de marché, doivent être financées. La sphère réelle regroupe l'ensemble des activités de production et de répartition des biens et services, les opérations d'échanges des facteurs de production... La sphère de la finance correspond à l'ensemble des opérations et des institutions permettant d'organiser les relations entre les agents qui disposent de capacités de financement avec ceux qui ont des besoins de financement.

20.1.2 *Interdépendances au niveau microéconomique et au niveau macroéconomique*

– Au niveau microéconomique, les opérations réelles sur les biens et les services (production et échange) et les opérations sur les facteurs de production (le travail et le capital productif) nécessitent des moyens financiers. Toutes les activités économiques réelles nécessitent un financement préalable. Les échanges de capacité de financement entre les agents à besoin de financement (emprunteurs, débiteurs) et les agents à capacité de financement sont la contrepartie indispensable des opérations dans la sphère réelle de l'économie. Par exemple, une entreprise devra emprunter pour investir au moment de sa création ou au cours de son activité. Un ménage devra emprunter pour un achat immobilier...

– D'un point de vue macroéconomique, l'épargne des ménages (agent structurellement prêteur, c'est-à-dire disposant d'une capacité de financement) va permettre de financer les investissements des entreprises et des administrations (agents structurellement emprunteurs). Il y a donc dans notre économie des agents qui sont des créanciers structurels et d'autres qui sont des débiteurs structurels. Comme, de surcroît, nos économies sont des économies monétaires, les échanges financiers et réels sont libellés et réalisés en monnaie. D'où l'importance de la valeur interne et de la valeur externe de la monnaie pour

la stabilité du système. Les déséquilibres monétaires peuvent être la cause ou la conséquence de déséquilibres financiers ou réels. Dans ces conditions, lorsque la croissance économique ralentit, cela signifie que la croissance de la production intérieure brute (PIB) est plus lente. Les besoins en main-d'œuvre sont moindres et le chômage augmente, entraînant dans son sillage tout un cortège d'inégalités. *A fortiori*, dans le cas d'une dépression, la production baisse et les difficultés économiques et sociales se renforcent. En outre, le ralentissement, voire la baisse de l'investissement en capital productif, peut contribuer à accélérer le processus récessif, ou dépressif.

20.1.3 *Interdépendances accrues dans une économie de marchés financiers*

L'économie, nationale, régionale ou mondiale prend donc la forme d'un système dans lequel les diverses composantes et les multiples éléments (agents et opérations) sont plus ou moins en interaction avec les autres composantes ou éléments, et avec le système dans son ensemble. En fait, depuis le début des années 1980, diverses réformes structurelles ont contribué à multiplier les phénomènes interactifs et la (sur)réactivité de ces systèmes interconnectés en continu. En effet, les processus de déréglementation, désintermédiation, décloisonnement et innovation ont favorisé le passage d'une économie d'endettement (finance indirecte) à une économie de marchés financiers (finance directe et titrisation). Dans le cadre d'une économie d'endettement, les agents à capacité et à besoin de financement réalisent leurs opérations financières par le biais d'un intermédiaire financier (le plus souvent une banque) et dans le cadre d'une économie de marchés financiers, et les deux catégories d'agents se rencontrent directement sur les différents marchés de capitaux. En théorie, cela se passe ainsi, mais on

constate que les banques et les autres organismes financiers se sont adaptés à l'économie de finance directe et y jouent un rôle majeur et fondamental. Les banques restent au cœur de la finance et ont donc joué un rôle crucial dans le déclenchement et la propagation de la crise des *subprimes* aux États-Unis.

Cet essor de l'économie financière de marché est d'une telle ampleur que les montants échangés sur les différents marchés de capitaux (marchés financiers et marché monétaire) sont désormais sans commune mesure avec les opérations et les échanges de biens et services dans l'économie réelle. Et cette nouvelle réalité pose, *de facto*, de nombreux problèmes financiers et économiques. Partant de ce constat, on peut se demander dans quelle mesure la finance reste au service de l'économie réelle, ou dans quelle mesure les marchés financiers sont déconnectés, autonomes, de l'économie réelle. En fait, il semble que la déconnexion existe surtout à court terme, mais que la connexion se rétablisse, au plus tard, à long terme. Pour autant, ce découplage de court terme n'est pas sans incidences durables, par le biais d'un phénomène d'hystérésis selon lequel l'effet persiste alors que sa cause (le facteur déclencheur) a disparu. Dans le cas présent, les effets financiers et réels de la crise initiée par les défaillances du système du crédit *subprime*, et qui s'est propagée du fait d'une titrisation débridée au niveau mondial, persiste depuis plusieurs années, alors que les créanciers imprudents et les débiteurs très risqués auront déjà largement payé la note.

20.2 NATURE ET MÉCANISMES DES CRISES FINANCIÈRES

20.2.1 *D'une crise financière à une crise économique*

Une crise financière se traduit par la **chute brutale des prix d'actifs financiers**, qui entraîne un effet de richesse négatif.

Simultanément, cet effet de richesse négatif s'accompagne d'un excès d'endettement. En effet, quand la valeur de ce que les agents possèdent baisse alors que le montant de la dette reste le même, les agents économiques voient leur taux d'endettement augmenter. Logiquement, ces crises financières ont des **conséquences importantes sur la croissance et l'emploi**. En effet, cet appauvrissement patrimonial et ce surendettement vont nécessiter un désendettement pour les ménages et les entreprises. Ainsi, tout ce qui s'achetait par la dette ne va plus s'acheter, la demande de crédit baisse. Pour les agents financiers, notamment les banques, l'appauvrissement va entraîner une baisse de l'offre de crédit. Donc le désendettement nécessaire entraîne une baisse de la demande globale, ce qui va entraîner une crise économique, car une baisse de la demande globale va entraîner une baisse de l'offre globale, laquelle va provoquer une hausse du chômage, donc une nouvelle baisse de la demande globale, ce qui va renforcer la crise économique, mais aussi la crise financière. C'est le cercle vicieux de la crise économique.

Selon la thèse de la **déflation par la dette** (Irving Fisher, 1933), une crise financière (chute brutale des prix des actifs) provoque, *ex post*, une hausse de la dette réelle des agents économiques. Avant le krach et la crise, ils s'étaient endettés pour davantage d'actifs. Désormais, ils vendent en masse leurs actifs pour récupérer des fonds et se désendetter. Ce faisant, ils obtiennent le contraire de ce qu'ils recherchent. En effet, la hausse de l'offre par rapport à la demande renforce la baisse des prix des actifs... La déflation accroît alors la valeur réelle de la dette.

Compte tenu des effets financiers et économiques destructeurs des crises financières, il semble indispensable d'essayer d'en comprendre les causes profondes afin d'éviter l'occurrence des spirales négatives qui peut en résulter. Pour davantage structurer notre compréhension des mécanismes fondamentaux de ces crises, nous pouvons utiliser la distinction habituelle entre les causes microéconomiques et les causes macroéconomiques, sachant que les unes ne sont pas exclusives des autres et réciproquement.

La récente crise dite des *subprimes* (2007) et ses effets cumulés nous serviront d'illustration.

20.2.2 *Les causes microéconomiques des crises financières*

D'un point de vue microéconomique, on peut mettre en évidence plusieurs facteurs majeurs des crises financières : une information imparfaite qui conduit à une estimation erronée du degré de risque des actifs (actifs immobiliers), une insuffisante standardisation des actifs financiers qui brouille la visibilité et la compréhension de ceux-ci, des règles prudentielles et des normes comptables procycliques, une allocation (répartition) non optimale et dangereuse des risques via le mécanisme de la titrisation, une distribution de crédit à des ménages selon des critères de solvabilité inadaptés, un risque de liquidité accru pour les banques.

20.2.2.1 La place centrale de l'information

Sur les marchés financiers, maîtriser l'information est crucial. Le marché international des capitaux est un marché totalement intégré qui favorise une rapide diffusion de l'information : ce phénomène témoigne des interdépendances entre les différents marchés et du rôle croissant des agences de notation financière. Ainsi, les évolutions sur l'un des segments des marchés de capitaux ont des effets instantanés sur les autres compartiments. Par exemple, la modification du taux de change euro/dollar a des effets sur le comportement des acteurs sur les marchés obligataires, notamment si les titres de créances sont libellés en dollars ou en euros. Ces interconnexions renforcent l'instabilité des marchés financiers, et ce que l'on appelle la volatilité, soit la variation (à la hausse comme à la baisse) de la valeur d'un titre par rapport à une valeur théorique, c'est-à-dire la variation de son cours sur les marchés.

L'information financière est souvent obligatoire et réglementée, notamment pour les sociétés cotées (cf. communications trimestrielles ou annuelles du bilan, du compte de résultat), mais elle est aussi informelle, sélective, voire réservée aux initiés. L'information est une nouvelle

connaissance, un renseignement porté à la connaissance de quelqu'un, et cet élément est capital sur les marchés financiers pour arbitrer entre tous les choix possibles. Sa diffusion a d'ailleurs des effets importants sur les marchés puisqu'elle est une des sources principales de révision des anticipations des opérateurs. Les signaux qui influencent les acteurs (acheteurs ou vendeurs) sur les marchés sont de nature différente. Ces informations peuvent être des indicateurs macroéconomiques (taux de croissance du PIB, taux de chômage, taux d'inflation, déficits et excédents commerciaux, etc.) ; des annonces de résultats d'entreprises (résultats financiers, annonce de la signature de contrats, cours boursiers, etc.) ; des décisions de banque centrale (taux directeurs, volume de refinancement, discours des gouverneurs des banques centrales, etc.) ; des décisions de politique économique (législation ou évolution du salaire minimum, revalorisation des minimas sociaux, modification de la fiscalité, etc.), mais aussi des rumeurs reprises ou diffusées par les acteurs de la finance. Ces rumeurs, fondées ou non, peuvent entraîner des comportements mimétiques et, ainsi, avoir des chances de se réaliser. Cette propriété auto-réalisatrice des informations diffusées sur les marchés financiers est source d'instabilité.

20.2.2.2 Le rôle décisif des agences de notation

Les agences de notation sont des entreprises privées qui alimentent les marchés financiers en proposant une information particulière : le « rating » ou appréciation de la qualité de signature d'un émetteur de titres financiers (entreprise, État, collectivité locale) ou des instruments financiers concernés. Les notations financières, élaborées à partir de données factuelles, d'hypothèses et de scénarios sur l'évolution des risques de non-remboursement, s'établissent de A à D. Plus la note est élevée (par exemple « AAA »), plus la solvabilité de l'émetteur est supposée forte ; plus la note est faible (C ou D), plus la solvabilité de l'émetteur est jugée fragile par l'agence de notation, qui diffuse alors l'information

d'un fort risque de non-remboursement des intérêts ou du capital. La globalisation financière a renforcé le rôle des agences de notation, car elles produisent une information facile à interpréter pour les investisseurs qui ne connaissent pas les émetteurs de titres dans un univers financier de plus en plus complexe. L'analyse et l'information sur les risques de crédit sont ainsi déléguées à un secteur oligopolistique puisque trois agences (Standard & Poor's, Moody's Investor Service et Fitch Ratings) détiennent plus de 90 % du marché mondial de la notation financière, dont 80 % pour les deux premières. L'information proposée (contre rémunération par les émetteurs) par les agences de notation a des effets sur l'entreprise émettrice : la note étant un indicateur de risque, une note élevée (ex. AAA) réduit le taux d'intérêt sur le titre et, à l'inverse, une note plus faible, informant d'un risque (de défaut, de difficultés ou de retards de paiement, etc.), augmente le taux d'intérêt demandé par les acheteurs des titres de créance, donc le coût de financement pour l'émetteur de dette. La notation financière attribuée par ces agences est devenue une référence pour les gestionnaires de fonds (compagnies d'assurances, fonds de pension, OPCVM), voire une norme utilisée par les régulateurs (cf. les accords de Bâle II, titres acceptés par la Banque Centrale européenne en contrepartie de refinancement des banques, etc.). Cette notation produit donc une information qui a, pour certains économistes, un caractère de bien public qui rendrait discutable leur production marchande, puisqu'il s'agit de sociétés privées.

20.2.2.3 Imperfection et asymétrie de l'information

En situation d'imperfection de l'information, on dit qu'il y a asymétries d'information. En général, les agents économiques prennent des décisions, font des choix, en situation d'asymétrie d'information, au sens où ils ne disposent pas tous des mêmes informations, ni de la même capacité à traiter ces informations pour prendre leurs décisions. Ces asymétries peuvent induire des

comportements non vertueux : comportements d'antisélection (sélection adverse) et comportement d'aléa moral. L'existence de ces comportements menace l'existence même des marchés concernés, car les échanges peuvent s'arrêter. Or, s'il n'y a plus d'échanges, il n'y a plus de marché !

– **Les comportements d'antisélection**

En situation d'asymétrie d'information, on observe des comportements d'antisélection (sélection adverse) quand les agents économiques font collectivement des choix qui vont conduire à la défaillance du marché. Sur le marché du crédit, les prêteurs (banques), ne sont pas certains de la solvabilité des emprunteurs. En relevant les taux d'intérêt afin de dissuader et d'éliminer les mauvais emprunteurs, les banques risquent d'éloigner les emprunteurs peu risqués et de se retrouver seulement avec les emprunteurs risqués. Il y a sélection adverse ou antisélection, car les échanges vont se réduire et le marché peut potentiellement disparaître, au détriment de chacune des parties.

– **Les comportements d'aléa moral**

En situation d'asymétrie d'information, on observe des comportements d'aléa moral quand un agent économique prend davantage de risques parce qu'il sait qu'il sera sauvé pour protéger l'économie d'un risque systémique. C'est souvent le cas des grandes banques jugées « too big to fail ». Le dilemme aléa moral versus risque systémique est un problème complexe posé aux économistes et aux autorités publiques.

Les situations d'aléa moral peuvent contribuer à l'instabilité financière, car l'action des banques centrales fournit une sorte d'assurance aux acteurs financiers qui savent que les choses ne vont pas trop mal tourner si le système devient instable. Cependant, quand les banques centrales interviennent pour limiter les conséquences des crises financières, elles rassurent, certes, les acteurs financiers, mais elles les incitent aussi, indirectement, à prendre des risques, puisqu'ils seront aidés en cas de récession. Ainsi, ce

sont les stratégies d'aide au secteur financier menées par les banques centrales qui sont les racines de la prise excessive de risque suivante et donc de la prochaine crise. Les banques centrales sont alors en quelque sorte des pompiers-pyromanes. Cet argument est différent d'une autre forme d'aléa moral, souvent citée en économie bancaire, qui est le « trop gros pour faire faillite ». Les acteurs systémiques dont la faillite mettrait en péril tout le système financier savent qu'ils bénéficieront de l'aide publique dans tous les cas et sont donc incités à prendre trop de risques.

20.2.3 *De la crise des subprimes à la crise financière*

20.2.3.1 Une crise peut en cacher une autre

Il est admis qu'une partie de la décennie 2000 fut une période de transition entre la bulle technologique (2001-2003) et la bulle immobilière (2006-2007), ponctuée par deux crises. Pour de nombreux observateurs, certaines des origines profondes de la crise financière de 2007-2008 se trouvent dans la lutte contre les effets dévastateurs de la bulle technologique au tournant des années 2000, avec un usage excessif et bien trop durable de l'expansion monétaire, qui a produit, graduellement, les conditions d'occurrence de la crise de l'immobilier américain en 2006-2007. En réaction à la crise de la « nouvelle économie », qui se traduit par une très forte hausse des défauts (de remboursement des dettes) des entreprises, les banques centrales américaine (FED ou Réserve fédérale) et européenne (BCE) mènent des politiques monétaires très expansionnistes, surtout aux États-Unis, afin de contrecarrer l'éventualité d'une situation d'assèchement du marché interbancaire et donc du crédit : le « credit crunch ». Le problème est que cette politique, si elle permet d'éviter le pire des scénarios à court terme, va ouvrir la voie, avec environ deux ans de décalage, à une forte poussée de l'endettement des ménages, en particulier l'endettement hypothécaire pour l'achat immobilier, puis celui

des entreprises. Cette hausse de l'endettement hypothécaire des ménages alimente la demande sur le marché immobilier (demande dans le secteur de la construction), laquelle provoque inéluctablement une hausse des prix des biens immobiliers. De surcroît, étant donné le fonctionnement du système d'endettement pour l'achat immobilier aux États-Unis (cf. le cercle vicieux de la finance de Ponzi et ses effets potentiellement explosifs), les plus-values en capital réalisées sur les reventes (effectives ou potentielles) de biens immobiliers augmentent les possibilités d'endettement des ménages.

20.2.3.2 De la solvabilité fondée sur le revenu à la solvabilité basée sur le prix, instable, d'un actif

En substituant une forme de solvabilité actif (patrimoine) à une forme de solvabilité fondée sur le revenu, les banquiers ont commencé à fabriquer les germes de leurs futures pertes financières. Une situation dans laquelle la solvabilité n'est assurée que dans la mesure où l'actif acheté voit son prix augmenter n'est *de facto* pas durable, car, dès l'instant où les prix baissent, les défauts des emprunteurs augmentent... Lorsque l'effet de richesse se retourne, la pyramide tend à s'effondrer comme un vulgaire château de cartes.

En effet, la crise des *subprimes* est la crise d'un marché de prêts hypothécaires à risque aux États-Unis. Les crédits *subprimes* s'adressent à des ménages peu, voire non solvables, car leurs revenus sont trop faibles pour rembourser. La solvabilité est le fait de pouvoir faire face à une dette avec des revenus suffisants. Les banques rendent solvables ces ménages via l'hypothèque. La valeur de la maison achetée à crédit sert de garantie d'emprunt. Ce système est tenable tant que le prix du logement reste supérieur au montant de la dette. Tant que le marché immobilier est haussier, le crédit hypothécaire semble être une bonne idée. Mais le jour où les prix des maisons baissent, le système s'effondre. En effet, tant que le prix de l'actif augmente, les gens sont incités à s'endetter. Or,

globalement, plus il y a de crédits *subprimes* accordés, plus le prix de l'immobilier augmente. Mais le crédit *subprime* est un crédit à taux variable. Les deux ou trois premières années, les taux étaient faibles, ainsi que les mensualités de remboursement. Les ménages à faibles revenus pouvaient faire face. Cependant, au moment du *reset* (quand le taux d'intérêt augmente), les mensualités augmentent et les ménages font défaut. Aucun système ne peut fonder durablement la solvabilité sur la valeur d'un actif, car le prix de celui-ci peut augmenter ou baisser très vite. La seule solvabilité solide est celle conférée par le revenu, car il est beaucoup plus stable dans le temps.

20.2.3.3 Modèle de crédit « originate and distribute » et titrisation débridée

De plus, le fait de privilégier le modèle de crédit « originate and distribute » via la pratique d'une titrisation débridée des créances (afin d'obtenir une meilleure et plus grande répartition des risques en regroupant des créances à risque faible isolément, en se soustrayant aux contraintes prudentielles) a contribué à faciliter et à accroître la distribution du crédit, à soutenir la dynamique de gonflement des bulles immobilière et financière. En effet, la diffusion du risque par la titrisation a rendu l'information sur le degré d'exposition à celui-ci d'une grande opacité. C'est le revers de la médaille de la titrisation, le côté obscur du système. Comme plus personne ne savait qui possédait quoi, ni combien, les marchés se sont asséchés au sens où lorsque la liquidité de marché tend vers zéro, plus personne ne peut vendre, car personne ne souhaite acheter. Or, quand il n'y a plus d'échange, plus de marché, les prix des actifs immobiliers et financiers s'effondrent. Ceux-ci entraînent *ipso facto* dans leur sillage les bilans des sociétés concernées qui chutent successivement comme dans un jeu de dominos. Les pertes en capital se cumulent, la solvabilité se dégrade, les défauts se multiplient, la liquidité disparaît, etc. La baisse entraîne la baisse comme la hausse suscitait de nouvelles

hausse quelques semaines ou mois, plus tôt. Si on ajoute à ceci la grande complexité des produits (CDO, ABS, MBS, CDS) qui ont porté la titrisation, on comprend que les acteurs concernés (banques et autres institutions financières) ont joué un rôle majeur dans le déclenchement et la propagation de cette crise.

20.2.3.4 De la crise des *subprimes* à la crise des produits financiers structurés

La crise des *subprimes* s'est accompagnée d'une crise sur les marchés de produits financiers structurés. Ces crises, interactives, sont à l'origine de nombreuses faillites aux États-Unis. Dans un premier temps, les défaillances d'organismes de crédit immobilier se sont multipliées dans ce pays. Dans un second temps, la crise des prêts immobiliers à risque américains, d'abord circonscrite au marché américain, s'est propagée à d'autres économies pour atteindre les banques australiennes, britanniques, allemandes... L'exposition de ces institutions centrales au marché de la titrisation et des produits structurés exotiques, devenus toxiques, car constitués en partie de crédits *subprime*, s'est révélée être une faiblesse fondamentale. Dès l'été 2007, BNP Paribas a gelé trois fonds de placement composés de titres adossés à des créances immobilières après avoir annoncé que la disparition de toute transaction sur certains segments du marché de la titrisation aux États-Unis avait conduit à une absence de prix de référence et à une illiquidité quasi totale des actifs figurant dans les portefeuilles des fonds, quelle que soit leur qualité ou leur notation. BNP Paribas reconnaissait qu'elle ne pouvait plus établir une juste valorisation des actifs contenus dans ces fonds pour justifier sa décision. Le ver était dans le fruit et le signal d'alarme avait retenti. La crise avait vraiment commencé.

En effet, cette vague de titrisation débridée, via des produits structurés avec la caution des grandes agences de notation (Standard & Poor's, Fitch et Moody's) et l'assurance obtenue par l'achat simultané de CDS (*credit default swap*), etc., a

favorisé le processus de contagion au reste de la finance et de l'économie mondiale. Les banques et courtiers ont pu vendre autant de crédits *subprimes*, car ils ont pratiqué la titrisation à l'excès. La titrisation consiste à transformer une créance en un actif négociable. Les banques peuvent augmenter leur capacité à accorder des crédits et à créer de la monnaie, grâce à la titrisation, laquelle permet de faire disparaître des créances de leur bilan. Pour parvenir à sortir du bilan des créances douteuses, risquées, les banques ont construit des produits structurés dits « exotiques » en mélangeant des créances risquées et des créances non risquées. Cette titrisation permet de répartir les risques combinés de différents actifs dans un nouveau produit structuré, et de partager les risques entre plusieurs détenteurs, de façon à ce que chacun soit convaincu de posséder un portefeuille d'actifs supposé en moyenne moins risqué. On a donc, d'une certaine manière, cherché à diluer le risque pour le camoufler, même s'il était toujours présent, mais personne ne savait vraiment où il se cachait et quelle était son ampleur.

Les banques ont développé ces pratiques de titrisation, car elles devaient faire face à une forte demande de crédit et que, de surcroît, ce montage financier rendait l'opération très rentable. En effet, les agences de notation, en attribuant la note maximum, le fameux triple A (AAA) à ces produits structurés, ont incité les investisseurs du monde entier à les acheter. *De facto*, le rendement de ces produits s'est élevé avec la forte demande dont ils ont fait l'objet, validant ainsi les notes des agences. Quand tout le monde achète, les prix augmentent, donc les agents ont de plus en plus envie d'en acheter. Cela produit une bulle des prix de ces titres, qui entretient la bulle des prix de l'immobilier. Ce faisant, les investisseurs qui achetaient ces produits structurés ont incité les banques à faire encore davantage de crédits et à acheter ces produits structurés. Petit à petit, les banques ont racheté des créances qu'elles avaient elles-mêmes participé à titriser.

De fait, elles ont tout simplement repris dans leur bilan des risques qu'elles avaient fait sortir pour l'assainir.

Dans le même temps, alors que les bulles atteignaient des sommets, quelques investisseurs mieux informés se doutaient de plus en plus du risque sous-jacent. Ils ont donc cherché à s'assurer contre d'éventuelles et probables pertes par l'achat de CDS (*credit default swap*), alimentant encore un peu plus les bulles sur les marchés de l'immobilier et des produits structurés. Pendant un temps, beaucoup y ont gagné, mais, comme « les arbres ne montent jamais jusqu'au ciel », selon l'adage en vigueur dans le monde de la finance, tout s'est effondré quand le marché immobilier s'est retourné et que le prix de l'actif sous-jacent (l'immobilier) a chuté. Il existe une règle intangible en finance selon laquelle la valeur d'un actif ne vaut que ce que vaut sa partie la plus faible. Malheureusement, l'emballement collectif est propice à « l'aveuglement au désastre » (André Orléan).

Le problème épineux de la contagion, du fait de la titrisation dans le monde entier, vient aggraver la crise et ses conséquences. La crise des *subprimes* s'est rapidement propagée à l'ensemble des marchés financiers : celui des crédits, des dérivés de crédit, des actions, des CDO, des ABS et du marché interbancaire. Cette transmission s'est faite par plusieurs canaux, en particulier la titrisation. La présence de crédits *subprimes* titrisés dans des produits structurés a fini par favoriser la montée de l'aversion pour le risque des banques et des investisseurs, et a alimenté les ventes massives en urgence par des fonds à levier d'endettement. Plus tard, en septembre 2008, le choix des autorités américaines de laisser Lehman Brothers faire faillite a déclenché une crise financière mondiale, du jour au lendemain. En choisissant de privilégier la lutte contre l'aléa moral, les autorités américaines ont renoncé à choisir la protection contre le risque systémique.

Pour autant, au-delà de ces comportements microéconomiques excessifs, il fallait que des conditions macroéconomiques existent.

20.2.4 *Les causes macroéconomiques des crises financières*

D'un point de vue macroéconomique, on peut mettre en évidence deux causes essentielles des crises financières : les politiques monétaires pro-cycliques, excessivement et durablement expansionnistes, des banques centrales des grands pays de l'OCDE depuis 2001 et la gestion de la crise technologique (bulle technologique) ; et l'excès de liquidité mondiale lié à l'accumulation de réserves de change par les banques centrales des pays émergents et exportateurs de pétrole.

20.2.4.1 Les politiques monétaires excessivement et durablement expansionnistes

Les banques centrales ont mené des politiques monétaires durablement, voire excessivement, expansionnistes qui ont créé un environnement favorable à l'apparition de bulles financières. Les réactions initiales de pompier des banques centrales ont fini par transformer progressivement celles-ci en pyromane.

L'hypothèse est que la crise financière de 2008 serait le résultat différé d'une politique monétaire excessivement et durablement expansionniste. Autrement dit, le maintien d'une telle politique monétaire serait à l'origine d'une « dynamique infernale » qui permet de sortir d'une crise, mais à condition d'en payer le prix, plus tard, en acceptant d'entrer dans une autre crise financière, ailleurs.

Le facteur déclenchant, le détonateur, est le maintien des taux d'intérêt directeurs si bas par rapport au taux de croissance du PIB, qui, en facilitant l'endettement, va conduire à la survalorisation des actifs sur les différents marchés. Cette survalorisation se traduit par l'apparition d'une bulle spéculative, qui, en éclatant, exige l'alimentation du marché en liquidité par le biais d'une politique monétaire très expansionniste... Il s'agit d'abord d'éviter le « credit crunch » (le rationnement du crédit pour les ménages et les

entreprises) sans pour autant créer les racines d'une bulle financière future.

Entre 1998 et 2000, puis entre 2003 et 2006, les banques centrales des États-Unis et des pays européens ont conservé des taux directeurs bas même dans les périodes de croissance, de progression du crédit et de hausse forte des prix des actifs. Ce comportement anormalement expansionniste des banques centrales dans les périodes d'expansion (donc pro-cyclique) favorise l'excès d'endettement et les bulles des prix des actifs. Ce comportement aux États-Unis et en Europe s'ajoute à la politique monétaire durablement expansionniste du Japon.

20.2.4.2 Excès de liquidité mondiale : accumulation de réserves de change et croissance excessive de la base monétaire

L'excès de liquidité mondiale s'explique aussi par l'accumulation de réserves de change par les banques centrales des pays émergents et exportateurs de pétrole et entraîne une croissance excessive de la base monétaire qui circule entre les banques commerciales et les banques centrales. Cet excès est lié à une absence de coordination internationale des politiques monétaires et des politiques de change. Dans ces conditions, chacun fait ce qu'il veut et une situation dangereuse s'installe du fait des interdépendances et des externalités des politiques économiques en économie ouverte. Ce manque de coordination internationale contribue à la création d'une très grande quantité de monnaie centrale et de monnaie. Les États et les banques centrales sont donc largement responsables de la persistance de déséquilibres financiers et monétaires internationaux qui portaient en eux, depuis des années, un potentiel de crise systémique.

Les excédents commerciaux de la Chine ont eu pour contrepartie un gonflement exponentiel des réserves de change de la banque centrale de Chine qui n'a pas hésité à acheter les dollars qu'elle recevait et des actifs en dollars en créant de la monnaie domestique pour éviter une appréciation

du renminbi par rapport au dollar. Ce comportement en matière de change étant une conséquence de la crise monétaire et de change qu'ont connu de nombreuses économies asiatiques à la fin des années 1990. L'objectif est non seulement de maintenir un haut degré de compétitivité prix, mais de surcroît, une plus grande capacité de défense du taux de change contre une éventuelle dépréciation brutale. Mais ces dollars accumulés vont revenir vers les États-Unis. Ils vont venir gonfler la base monétaire et donc y alimenter la création monétaire, la dette et les bulles immobilières, financières... Cet afflux des dollars en provenance de Chine vers les États-Unis a produit un excès durable de liquidité macroéconomique (monétaire) qui a entretenu les conditions pour le crédit facile, la titrisation, la demande d'actifs, le gonflement de bulles spéculatives... et finalement la crise de 2008 dans toutes ses dimensions. Le fait que le pays le plus riche, ayant un taux d'épargne quasiment nul, puisse consommer et investir de manière durablement excessive parce que des économies beaucoup moins développées épargnent énormément, mais investissent relativement peu au regard des besoins d'investissement qu'ils ont, et transfèrent leurs capitaux vers les États-Unis où les agents peuvent ainsi continuer à consommer et investir massivement, fut une véritable énigme contemporaine.

En définitive, c'est bien la combinaison de déséquilibres monétaires et financiers internationaux et de comportements inappropriés au niveau microéconomique qui est responsable du déclenchement de la crise, de sa propagation et de son ampleur.

Nous pouvons clairement déduire, de cet épisode, qu'un régime de croissance basé sur l'endettement privé est insoutenable à moyen terme. Il n'est pas possible de soutenir durablement la croissance d'une économie dans laquelle les inégalités de revenu se sont accrues, une économie où le seul moyen de maintenir un haut niveau de consommation, de croissance, est de permettre l'endettement des plus pauvres.

20.3 LES RÉPONSES DES AUTORITÉS PUBLIQUES FACE À LA CRISE FINANCIÈRE ET ÉCONOMIQUE

Depuis 2007, dans la plupart des grands pays développés et depuis 2010 dans la zone euro, les banques centrales ont dû agir pour contrecarrer les effets négatifs de l'enchaînement des crises financières et des crises économiques qui les accompagnent inévitablement. Cette période ouverte en 2007-2008 est qualifiée de « Grande Récession » compte tenu de l'ampleur de la baisse des taux de croissance (voire du PIB), de la hausse des taux de chômage, mais également de sa persistance et de sa durée.

Afin de réussir à atteindre ses objectifs, la politique monétaire expansionniste va prendre des formes non conventionnelles pour relancer le crédit bancaire, faciliter le désendettement et le ré-endettement des agents économiques privés, faciliter l'endettement public et le financement des politiques budgétaires, favoriser la reprise, lutter contre les risques de déflation et de dépression, restaurer la stabilité financière sans porter préjudice à la stabilité des prix.

En effet, les politiques de relance traditionnelles n'ont pas permis de stabiliser les anticipations sur les marchés financiers, et les banques centrales ont dû déployer de nouveaux instruments, notamment pour intervenir sur les marchés obligataires.

Pour la BCE, la stabilité financière est une situation dans laquelle le système financier — qui englobe les intermédiaires, les marchés et les infrastructures de marché — est capable de résister aux chocs, en réduisant la probabilité d'une interruption du processus d'intermédiation financière, suffisamment importante pour perturber l'allocation optimale des ressources.

20.3.1 *L'action des banques centrales et des États face à la crise de 2007-2010*

En ce qui concerne l'action à court terme, on peut distinguer trois types de mesures :

- D'abord, les mesures immédiates des banques centrales face aux problèmes de liquidités des banques. La banque centrale doit soutenir les banques en leur donnant un accès à toutes les liquidités dont elles ont besoin. Il s'agit de son rôle de prêteur en dernier ressort. Mais cela n'a pas suffi, il a fallu qu'un autre « pompier » intervienne : l'État.
- Dans un second temps, les plans de sauvetage bancaires et financiers sont indispensables pour faire face à des risques de solvabilité. Les États ont dû utiliser de l'argent public pour sauver les banques et le système financier, afin d'assurer la poursuite du financement des économies et de pouvoir sauver *in fine* l'économie réelle.
- Enfin, les plans de relance de l'économie réelle concernent l'ensemble des mesures de relance macroéconomique de type keynésien. Ces mesures ont été mises en place pour l'essentiel en 2008-2009.

20.3.1.1 Les mesures de politique monétaire des banques centrales face aux problèmes de liquidité

Il y a crise de liquidité bancaire quand les banques ne peuvent plus se refinancer, c'est-à-dire quand elles ne peuvent plus obtenir la monnaie centrale dont elles ont besoin pour créer de la monnaie en accordant des crédits... Une crise de liquidité est d'abord une crise localisée sur le marché interbancaire. Face à cette crise interbancaire, les banques centrales baissent les taux d'intérêt directeurs. Les cinq plus grandes banques centrales l'ont fait en même temps en octobre 2008, dans le cadre d'une vraie coopération internationale, car les banques des différents pays sont toutes liées par la globalisation financière et elles doivent pouvoir

échanger de la monnaie et des devises. En baissant les taux directeurs, les banques centrales espèrent que les banques se feront davantage confiance pour se prêter entre elles, car elles pourront toujours obtenir de la liquidité à des taux très bas. Mais cela ne va pas vraiment fonctionner. Les banques centrales vont devoir mettre en place des politiques monétaires non conventionnelles afin de fournir une plus grande quantité de liquidité aux banques.

20.3.1.2 Les plans de sauvetage du système bancaire

Face aux pertes des banques et aux risques d'insolvabilité, les États et les banques centrales vont mettre en place des plans de sauvetage du système bancaire et financier. Les banques réalisent des pertes en se débarrassant à prix réduit de leurs produits devenus toxiques. Elles sont donc en risque de faillite. D'où l'intervention des États, avec les plans de sauvetage. Dans le cadre de ces plans, les États vont garantir les emprunts des banques. Si celles-ci peuvent emprunter et faire face à leurs dettes passées, elles seront jugées plus solvables et pourront donc prêter aux agents non financiers (ménages, entreprises). La création de structures de défaisance ou de cantonnement (il s'agit de Fonds ad hoc) va permettre aux banques de libérer leur bilan des actifs toxiques qui s'y trouvent encore, ceux dont la valeur de marché est devenue très faible. Une structure de défaisance permet aux banques de sortir ces produits toxiques de leur bilan. Le but est de redonner confiance dans la solvabilité de banques... pour que les prêts interbancaires puissent avoir lieu et que les banques puissent financer l'économie. Les États ont pu aussi recapitaliser des banques, voire les nationaliser. L'État peut ainsi devenir copropriétaire de l'entreprise. Cela peut rassurer les investisseurs sur la solvabilité de la société. Les États ont enfin augmenté la garantie des dépôts dans les banques. L'objectif est d'éviter les « bank runs », c'est-à-dire les files d'attente des clients devant les banques afin de retirer leur argent en espèce. Cette situation entraîne automatiquement la faillite des

banques concernées, qui ne peuvent honorer la demande de tous leurs clients simultanément. De surcroît, en évitant les paniques bancaires, l'État ne protège pas seulement les banques, mais les agents économiques et donc l'économie. Ici, la confiance joue un rôle majeur. Or, la monnaie est fondée sur la confiance. Dans une situation de crise de confiance, il faut absolument restaurer cette confiance de tous les agents économiques, sans laquelle rien ne peut être résolu.

20.3.1.3 Les plans de relance des économies des États

Il aurait été peu efficace de mettre en place des plans de relance de la croissance par le soutien de la demande globale sans avoir préalablement traité la crise de liquidité interbancaire et mené des plans de sauvetage bancaires et financiers. Faire préalablement cela était une condition nécessaire avant d'essayer de relancer les économies. En retour, ces plans de relance, en facilitant la résolution durable des problèmes de solvabilité des banques, rendent possible le rééquilibrage du système bancaire et financier. Les plans de relance consistent à relancer la croissance et l'emploi en soutenant la consommation des ménages et les investissements des entreprises. Pour y parvenir, les dépenses publiques augmentent et les prélèvements baissent. Le creusement du déficit budgétaire entraîne une hausse du montant de la dette publique. Les plans de relance par les politiques budgétaires expansionnistes sont souvent accompagnés de politiques monétaires expansionnistes. Le problème est qu'une politique monétaire expansionniste n'est pas efficace quand les agents non financiers sont globalement trop endettés, que leur priorité est le désendettement et quand, du côté de l'offre de crédit, les banques restent frileuses. Nous voyons que la réussite des plans de relance est étroitement dépendante des succès des plans de sauvetage des banques et réciproquement. Enfin, si les mêmes causes produisent toujours les mêmes effets, alors on peut légitimement se demander si la lutte présente

contre les crises du moment et contre leurs conséquences ne porte pas en elle les germes de crises du même type dans les années suivantes.

20.4 LA ZONE EURO ET LA CRISE DES DETTES SOUVERAINES : CRISE DE FINANCEMENT DES ÉTATS

La crise des dettes souveraines dans la zone euro, à partir de 2010, illustre les difficultés de financement que rencontrent les États en période d'incertitude. Ces problèmes de financement des administrations publiques peuvent être la conséquence et la cause de ceux du secteur privé. La crise des dettes souveraines dans la zone euro permet aussi de mettre en évidence les interactions et interdépendances entre politique monétaire et politiques budgétaires. Cela oblige les autorités publiques, nationales et européennes à repenser le « policy-mix » et les institutions de la zone euro.

20.4.1 *Les causes économiques de la crise des dettes souveraines*

On peut considérer que la crise de la zone euro a eu principalement trois types de causes : la crise financière et économique mondiale de 2007-2010, qui correspond à un choc symétrique aux effets asymétriques sur des économies aux structures plus ou moins hétérogènes, la grande hétérogénéité des économies de la zone euro en termes de structures et les défaillances institutionnelles associées aux défauts de gouvernance économique de la zone euro depuis 1999.

20.4.1.1 Les effets déstabilisateurs de la crise financière et économique mondiale

La zone euro a, dans un premier temps, été affectée par la crise financière américaine de 2007. La crise des *subprimes* a touché les banques européennes en septembre 2008 avec des pertes en capital, une crise de liquidité et une baisse du

produit net bancaire. Un certain nombre de banques de la zone euro ont ainsi été acculées à la faillite (mais en moins grand nombre qu'aux États-Unis).

La crise financière venue des États-Unis s'est ensuite transmise à l'économie réelle, et la zone euro a dû faire face à des menaces de récession à partir du printemps 2008. Le choc financier a entraîné une restriction de l'offre de crédit : puisque les banques ont eu des difficultés à trouver des liquidités et à reconstituer leurs fonds propres, elles ont réduit l'octroi de prêts aux ménages et aux entreprises et le coût des emprunts – le taux d'intérêt – a augmenté (prime de risque). La crise financière a entraîné une chute de la valeur des actifs immobiliers et boursiers des ménages, qui ont épargné davantage par précaution, ce qui a provoqué une baisse de la consommation (ou un report des achats envisagés). De surcroît, la crise a provoqué un mouvement de désendettement rapide des ménages et des entreprises.

Par ailleurs, les pays de la zone euro ont également été touchés par une contraction des échanges internationaux et un renchérissement du coût des matières premières en 2007-2008. En raison de la crise financière puis de la transmission de ses effets à l'économie « réelle » (ce que certains économistes ont appelé la « Grande Récession »), les États ont cherché à compenser ce désendettement privé par une croissance de l'endettement public afin de soutenir la demande globale et d'éviter l'aggravation de la récession. Cet endettement élevé des États a entraîné un creusement des déficits publics, un gonflement de leur dette publique et donc une montée des primes de risques imposées par les marchés financiers.

La récession s'est conjuguée aux politiques de relance budgétaire et aux plans de sauvetage des établissements bancaires menacés de faillite pour accroître les charges financières des États, jusqu'à susciter la défiance des marchés financiers internationaux face à leurs capacités de remboursement, alors que le vieillissement démographique

risque d'accroître encore les besoins de financement dans l'avenir.

En règle générale, les États empruntent de l'argent en émettant des obligations, titres de créance qui constituent pourtant les actifs de référence des marchés de capitaux, car elles sont garanties par des États qui ne peuvent en théorie pas faire faillite. Mais lorsque le risque d'un défaut de remboursement sur la dette souveraine intervient et que l'incertitude augmente, dans un contexte de ralentissement économique, les acheteurs potentiels demandent des taux d'intérêt sur les obligations d'État plus élevés (ce qu'on appelle la prime de risque). La détérioration de la situation des finances publiques de certains pays membres de la zone euro a avivé l'inquiétude des marchés financiers sur la probable incapacité des pays les plus endettés à honorer leur dette, et le coût de l'emprunt a alors fortement augmenté. En 2007, avant la crise, la Grèce payait sur ses obligations un taux d'intérêt supérieur d'environ un quart de point au taux allemand. En mars 2010, cette prime de risque sur les emprunts souscrits par le gouvernement grec a atteint 325 points de base (3,5 %), avant d'augmenter encore par la suite. Ainsi, la Grèce devait payer des taux d'intérêt nettement plus élevés pour emprunter de l'argent, ce qui, en retour, accroissait encore ses besoins d'emprunter pour faire face aux échéances.

20.4.1.2 Éclatement et contagion : de la crise de la dette souveraine grecque à une crise généralisée des dettes souveraines

La crise des dettes souveraines se caractérise par un risque de défaut de paiement d'un État sur sa dette. Dans la zone euro, celle-ci est née d'une anticipation partagée sur les marchés financiers d'un risque futur, voire imminent, mais non certain, d'un défaut de paiement d'un État. Il s'agit donc d'une crise de confiance, liée à l'opinion des marchés financiers, qui peut avoir des effets très pervers, car l'État débiteur doit faire face à un net

durcissement de ses conditions de refinancement, avec une forte hausse des taux d'intérêt. Un mécanisme de prophéties auto-réalisatrices peut alors déclencher un défaut de paiement et le pays peut être confronté à des sorties de capitaux, des fuites de dépôts bancaires et à une baisse des investissements étrangers qui réduisent ses capacités de remboursement futures.

Pourtant, lorsque la crise de la dette souveraine grecque débute à l'automne 2009 dans la zone euro, la Grèce ne fait pas défaut sur le service de sa dette, mais la révélation d'une grave déficience dans la présentation de ses comptes publics a fait naître l'idée sur les marchés de capitaux que le pays pourrait ne plus assurer le paiement à l'échéance de ses créances. Il est apparu que la dette grecque était en réalité beaucoup plus élevée que ce qui avait été annoncé initialement. Les agences de notation ont alors pris la mesure de la situation financière réelle de la Grèce, avec une dette publique proche des 115 % du PIB et un déficit budgétaire proche de 13 % du PIB, et cela a eu pour effet de renchérir les taux d'intérêt demandés sur les marchés et aggravé d'autant le problème de financement de la Grèce. La crise de la dette souveraine grecque s'accompagne d'un risque de contagion et d'aggravation de la récession.

Après la Grèce, la perte de confiance des marchés s'est propagée à d'autres États de la zone euro comme l'Italie, l'Espagne, le Portugal et l'Irlande. La prise de conscience de ce risque de contagion a alors aggravé la situation des finances publiques des États de la zone euro les plus exposés, mais également fragilisé le système bancaire et financier mondial, puisque les grandes banques internationales détentrices de ces titres de dettes souveraines ont vu apparaître un risque de défaut de paiement qu'elles n'avaient pas anticipé. Pour rassurer les marchés financiers, la Grèce a dû mettre en œuvre un premier plan d'austérité, complété par un prêt de l'Union européenne et du FMI, avant que l'Union ne décide d'un nouveau plan de secours plus

important. Mais en l'absence d'un budget fédéral permettant de réellement rassurer les marchés financiers sur la solidarité budgétaire des États, la zone euro reste soumise à de potentielles tensions spéculatives. Plusieurs pays se sont engagés à remettre en ordre leurs finances publiques (Grèce, Italie) en décidant de diverses mesures d'austérité. Le risque est que ces plans de rigueur budgétaire ne freinent encore davantage la croissance du PIB et finalement amputent les possibilités de remboursement des États membres de la zone euro. Mais la crise des dettes souveraines a surtout constitué un révélateur de déséquilibres plus profonds au sein de la zone euro, et mis en évidence ses fragilités institutionnelles, notamment en matière de solidarité budgétaire.

La crise de la dette souveraine grecque fut donc un révélateur des différentes limites du fonctionnement interne et de la gouvernance de l'Union économique et monétaire. La crise de la dette souveraine grecque n'est, en fait, pas uniquement grecque, car elle résulte aussi des défaillances fonctionnelles de la zone euro. Ce phénomène peut être illustré par la métaphore de « l'arbre qui cache la forêt ». Derrière la crise dite grecque se cache une crise bien plus profonde, celle de la zone euro... au péril de l'intégrité de la zone. Ce sont probablement les dysfonctionnements internes de la zone euro, bien davantage que les abus divers du gouvernement grec, qui expliquent la crise grecque, laquelle entraîne un risque de crise systémique par effet de contagion.

20.4.1.3 Les conséquences de la grande hétérogénéité des économies de la zone euro

Une union monétaire, dont les économies sont très hétérogènes, sans discipline budgétaire partagée et respectée, ni intégrations financière et bancaire suffisante (via une union bancaire) est vouée *ipso facto* à l'échec ! En effet, une simple union monétaire va avoir des effets contradictoires selon la situation conjoncturelle des économies constitutives.

Dans des périodes calmes, l'union monétaire va avoir un effet anesthésiant. Une sorte de douce négligence, de nouveau paradoxe de la tranquillité, s'installe et rend possible la formation de déséquilibres durables de compétitivité, de dette privée, de dette publique, de croissance, d'emploi, de chômage, de revenus... Depuis 1999 et la création de l'euro, les taux d'intérêt ont baissé dans la zone euro et favorisé un emballement du crédit et la formation de bulles spéculatives, en particulier sur le marché de l'immobilier, où les prix ont fortement augmenté durant les années 2000 : comme le taux d'intérêt est le même pour l'ensemble de la zone, les pays où l'inflation était plus élevée ont vu leurs taux d'intérêt réels (*on obtient le taux d'intérêt réel en faisant la différence entre le taux d'intérêt nominal et le taux d'inflation*) baisser encore davantage, ce qui a stimulé artificiellement la demande intérieure et gonflé l'endettement privé (la dette combinée des ménages et des entreprises atteignait ainsi près de 380 % du PIB en Irlande et 215 % du PIB en Espagne en 2007). Une bulle de crédit s'est formée, puis a explosé avec la crise financière ; le risque de défaut de remboursement a brutalement augmenté avec la chute des prix de l'immobilier, la récession et la stagnation des revenus, contraignant les pays à de difficiles plans de rigueur, à l'instar de la Grèce (et de l'Espagne). *A contrario*, dans les périodes agitées, troublées, l'union monétaire a eu un effet accélérateur de la crise dans la mesure où le cumul des plans de rigueur budgétaire, et même d'austérité, a un effet contre-productif via les effets multiplicateurs, la croissance molle, la hausse du chômage et, *in fine*, le creusement des déficits budgétaires, qui entraînent la hausse des ratios de dette publique...

LES DIFFÉRENTES DIMENSIONS DE LA CRISE DES DETTES SOUVERAINES

Une crise des dettes souveraines peut présenter différents degrés en matière de gravité : une **crise de liquidité** (capacité de refinancement pour couvrir le service de la dette) peut devenir une **crise de solvabilité** (capacité à faire face à ses engagements en matière de remboursement du montant de la dette souscrite) et remettre en cause, à long terme, la **soutenabilité de la dette publique** au regard de la trajectoire de l'économie réelle et de la croissance potentielle. De plus, cette crise de la dette souveraine grecque et la contagion qui a suivi ont clairement mis en évidence **un cercle vicieux très**

dangereux entre risque bancaire et risque souverain, en raison de la forte détention de dette des différents États par les banques des pays de la zone euro. **Crises des dettes souveraines et crises bancaires sont intimement liées.** En 1999, les architectes de la zone euro avaient cru que le marché unique conduirait spontanément à l'Union bancaire et financière, mais ce ne fut pas le cas. *A contrario*, les systèmes bancaires et financiers sont devenus de plus en plus fragmentés, ce qui a été un facteur d'amplification de la crise de l'euro.

20.4.2 Les causes institutionnelles : défaillances de la gouvernance de la zone euro

Les défauts de construction institutionnelle sont l'indépendance de la BCE, la clause de non-renflouement (« no bail out »), l'échec de l'article 126.1 du Traité sur la discipline budgétaire, le manque d'intégration bancaire et budgétaire... et leurs conséquences cumulées.

20.4.2.1 Les effets de l'indépendance de la BCE peuvent être négatifs

L'indépendance de la BCE vis-à-vis des gouvernements interdit la monétisation des dettes publiques ou souveraines (la dette publique désigne la dette des administrations publiques, la dette souveraine désigne la dette de l'État central). La BCE n'a pas le droit de « prêter » aux États en achetant des titres de la dette publique sur le marché primaire (monétisation). C'est pour cette raison que les banques centrales sont devenues indépendantes. Cela empêche les États qui ont des besoins d'endettement importants de s'endetter à des taux très bas, de faire fonctionner la « planche à billets », d'avoir des comportements durablement irresponsables en matière budgétaire. Mais, en période de crise et de doute sur la viabilité de la zone euro, le signal qui est envoyé aux créanciers internationaux qui ont prêté aux États de la zone euro (lesquels doivent soutenir leurs économies en période de crise, quitte à creuser les déficits budgétaires) fabrique de la défiance. Les États, garants de la stabilité de l'économie nationale, ne seront pas aidés par la banque centrale, même si la situation l'exige ! Dans ces conditions, une crise de la dette publique a de fortes chances de s'aggraver rapidement.

20.4.2.2 La clause de non-renflouement peut favoriser la méfiance des investisseurs

La clause de « *no bail out* » va venir renforcer ce sentiment de défiance. Selon cette clause, un État ne peut pas en sauver un autre. Cette clause a été mise en place pour éviter les comportements d'aléa moral, pour que les États ne comptent pas trop sur les autres. Mais, en luttant contre l'aléa moral, on a favorisé l'émergence du risque systémique... D'autant plus que les créanciers internationaux savent déjà que la BCE ne viendra pas au secours des États. Même si la doctrine européenne a évolué depuis 2010, et s'est partiellement orientée vers le « *bail in* », cette clause était également susceptible de renforcer les craintes des créanciers des États externes à la zone euro. Pour eux, détenir des créances en euros et donc de l'euro devenait de plus en plus risqué. L'enjeu à court terme est donc d'introduire un instrument permettant d'immuniser définitivement les États contre le risque bancaire. C'est précisément la finalité du *bail in*, qui impose aux créanciers d'une banque insolvable la conversion de leurs créances en actions de cette banque ou la réduction du montant de ces créances. Dans ce schéma, le coût du sauvetage ne pèse plus uniquement sur le contribuable, mais aussi sur les investisseurs, ce qui permet théoriquement de rompre partiellement le cercle vicieux banques-États.

20.4.2.3 L'échec de l'article 126.1 du Traité sur la discipline budgétaire

L'échec de l'article 126.1 du Traité sur la discipline budgétaire est une troisième forme de faille institutionnelle. Encore une fois, au départ, la règle se veut collectivement protectrice. Les excès des uns ne devaient pas avoir d'effets négatifs sur les autres. Il s'agit d'éviter les comportements de passager clandestin en union monétaire. En effet, dans une zone monétaire, il y avait d'autant plus de risques que l'effet d'éviction lié aux excès des uns (endettement excessif d'un

État) se répercute sur les autres. Pourtant, dès 2003, on a assisté à des problèmes de non-respect des règles par l'Allemagne et la France, pays fondateurs et piliers de la construction économique et monétaire européenne. Alors, en 2005, quelques changements ont été opérés dans les règles. Plus tard, en 2009, on a découvert que l'État grec (la Grèce est une petite économie relativement à la zone euro dans son ensemble) ne respectait absolument pas la discipline budgétaire et que ses dirigeants avaient même falsifié les comptes transmis à Eurostat, l'office de statistique européen. Le rappel à l'ordre fut limpide : les règles doivent être respectées. La discipline budgétaire partagée était donc prévue, mais elle n'a pas été complètement respectée. Un message propice à créer de la défiance chez les créanciers et autres investisseurs en euros dans la zone euro...

UN POLICY MIX NON OPTIMAL : POLITIQUE MONÉTAIRE UNIQUE ET POLITIQUES BUDGÉTAIRES NATIONALES

La zone euro est un modèle de gouvernance totalement original qui combine une politique monétaire unique et onze puis dix-neuf politiques budgétaires souveraines, encadrées par des règles. Cela se traduit donc par dix-neuf « policy mix », compte tenu du cadre réglementaire. Concrètement, la gestion macroéconomique de la zone euro depuis 1999 n'a pas pu être optimale, car il n'existait qu'une seule et même politique monétaire pour des pays membres qui se trouvaient dans des situations économiques très différentes, du fait de l'hétérogénéité structurelle originelle et croissante entre économies. En effet, la politique monétaire unique se traduit par la conjonction d'un taux d'intérêt nominal unique et de taux d'inflation différents. *Ipso facto*, cela entraîne des écarts importants de taux d'intérêt réels qui modifient le coût relatif de la dette, les incitations et les arbitrages.

20.4.2.4 Le manque d'intégration bancaire

Le manque d'intégration bancaire s'est révélé être une faille majeure dans l'édifice institutionnel de la zone euro. On pensait que la solidité des systèmes bancaires nationaux pouvait être assurée par les seuls États. Autrement dit, qu'un État aurait les moyens financiers suffisants pour sauver les banques domestiques en difficulté. Or, ce n'est pas le cas. Les tailles des bilans de certaines grandes banques sont des multiples du PIB de certaines économies et donc du budget des États. La crise de la zone euro a servi de laboratoire en permettant de montrer que la solidité des systèmes bancaires nationaux ne peut désormais plus être assurée par les seuls États. Une union bancaire est nécessaire pour

assurer le bon fonctionnement des banques sur l'ensemble de la zone euro et éviter les « Bank runs ». L'union bancaire doit aussi permettre une meilleure transmission des effets des décisions de la politique monétaire de la BCE (voir chapitre 6, point 6.5).

20.5 L'ACTION DES ÉTATS ET DE LA BCE FACE AUX DIFFÉRENTES DIMENSIONS DE LA CRISE DES DETTES SOUVERAINES

20.5.1 *Les mesures budgétaires des États : hausse des déficits budgétaires et de la dette souveraine*

Les remèdes à cette crise multidimensionnelle sont du même type que ceux déjà présentés pour la crise financière et économique mondiale. La crise économique et financière en zone euro entraîne *ipso facto* le creusement des déficits budgétaires et la hausse des taux d'endettement en raison des actions des États par le jeu des stabilisateurs automatiques. Les plans de relance, via des politiques budgétaires discrétionnaires selon une logique keynésienne, se traduisent par un creusement des déficits publics et la hausse de la dette publique. Les plans de sauvetage bancaires et financiers correspondent à des dépenses publiques nécessaires pour sauver le système bancaire et financier en évitant le cercle vicieux entre risque bancaire et risque souverain. Mais le cumul des effets sur les finances publiques de ces diverses mesures se traduit par une très forte hausse des ratios dette publique/PIB pour de nombreux États ayant l'euro comme monnaie. Or, en période de méfiance sur la liquidité des titres de dette publique et sur la solvabilité des États souverains, cela ne fait que renforcer

les craintes des investisseurs. Et la méfiance favorise la contagion... surtout dans le cadre institutionnel et de gouvernance de la zone euro tel qu'il est en cette période.

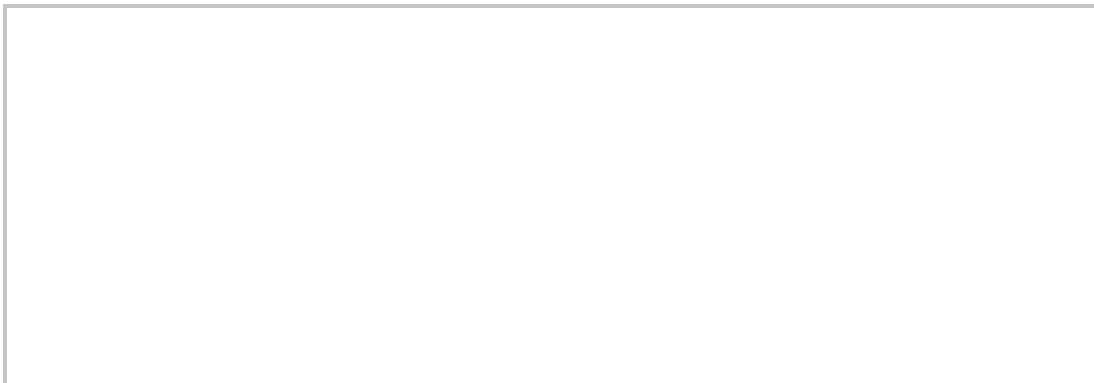
20.5.2 *Les mesures de politique monétaire de la BCE*

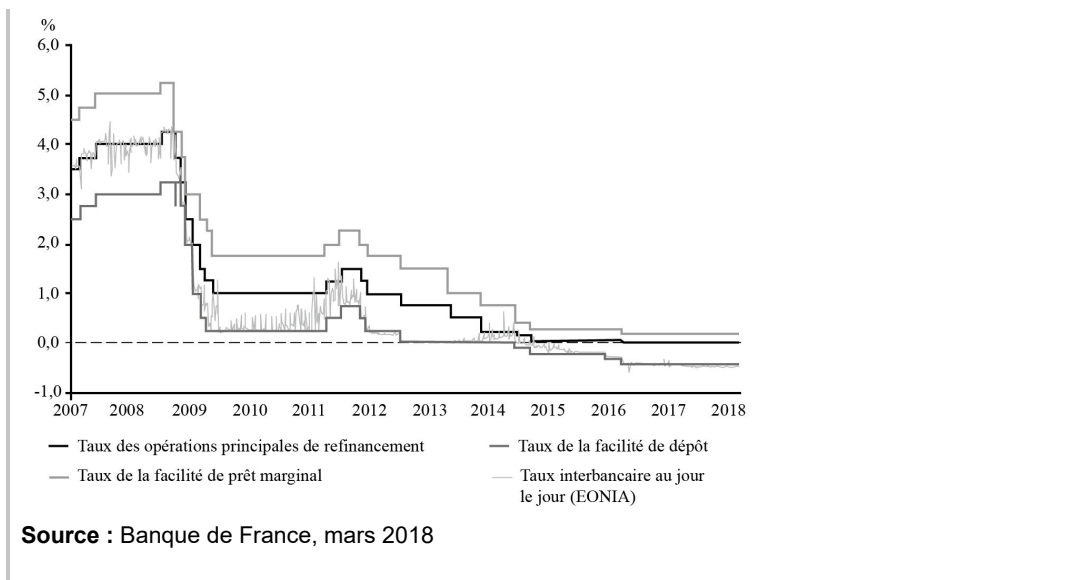
Il convient de différencier les mesures standards (ou conventionnelles) des mesures non standards (ou non conventionnelles).

20.5.2.1 Les objectifs et instruments de la politique monétaire conventionnelle de la BCE

Le rôle des taux d'intérêt directeurs

L'action de la Banque Centrale européenne (BCE) a eu pour principal objectif de rompre la spirale négative et d'éviter l'éclatement de la zone euro. Pour atteindre son but, la BCE a utilisé divers instruments, conventionnels et non conventionnels. Dans la zone euro, la BCE utilise principalement le levier des taux d'intérêt. Néanmoins, pour atteindre ses objectifs, l'Eurosystème dispose de multiples instruments : les opérations d'open market, les facilités permanentes et la constitution de réserves obligatoires rémunérées sur des comptes ouverts aux banques par les banques centrales nationales (BCN). Les opérations « d'open market » et les facilités permanentes permettent de couvrir l'essentiel du besoin de refinancement des établissements de crédit qui résulte des réserves obligatoires et des « facteurs autonomes de la liquidité bancaire ».





GRAPHIQUE 20.1.

Évolution des taux directeurs de la BCE et du taux du marché monétaire sur la période 2007-2018

Les opérations d'open market (qui s'adressent à toutes les banques sans discrimination) consistent en un pilotage des taux d'intérêt pour la gestion des liquidités. Ce sont des opérations de cession temporaires (des prêts garantis par la prise en pension d'actifs ou collatéraux et non des achats fermes d'actifs) à l'initiative de la BCE par la voie d'appels d'offres hebdomadaires. Elles permettent de déterminer le principal taux d'intérêt directeur : le taux principal de refinancement. La BCE propose d'augmenter les liquidités aux plus offrants via les opérations principales de refinancement (MRO : Main Refinancing Operations). Ces concours sont d'une durée d'un jour à une semaine. Le guichet lié aux appels d'offres est donc ouvert temporairement. Les appels d'offres sont réalisés en général à taux fixe. La BCE contrôle donc le prix et peut ensuite allouer sa monnaie en quantité limitée ou illimitée, selon les besoins et les circonstances. Dans les cas des appels d'offres à taux variable, la BCE prête les liquidités selon un mécanisme d'enchère.

Les facilités permanentes correspondent à un guichet ouvert en permanence. Elles permettent de déterminer les deux autres taux directeurs de l'Eurosystème. Les facilités permanentes permettent de fournir ou de retirer des liquidités, par le biais de prêts ou de dépôts à 24h, de manière automatisée ou à l'initiative des établissements de crédit contreparties de la Banque de France. Le taux des facilités de prêt marginal détermine le taux plafond. Les BCN accordent un prêt rémunéré à la journée, toujours contre collatéral, les prêts en blanc étant interdits. Le taux des facilités de dépôt que les établissements de crédit peuvent utiliser pour effectuer des dépôts à 24 heures auprès des BCN de l'Eurosystème détermine le taux plancher. Les taux des facilités permanentes délimitent le corridor de fluctuation du loyer de l'argent au jour le jour sur le marché monétaire. Le niveau de l'EONIA (Euro Overnight Index Average), un taux au jour le jour pratiqué sur le marché interbancaire pour les échanges (offre et demande de monnaie centrale), en résultera et oscillera autour du taux de refinancement, dans le corridor entre les deux taux de facilités marginales.

Mais, constatant « *des tensions aiguës sur certains segments des marchés qui entravent le mécanisme de transmission de la politique monétaire* », la BCE a d'abord cherché à assurer le refinancement des banques pour rétablir le financement de l'économie. La politique monétaire de la BCE est d'abord très expansionniste via une baisse durable et importante des taux d'intérêt directeurs. L'objectif étant de réduire le coût de refinancement des banques.

Le graphique montre que dans le cas de la zone euro, la baisse des TID (et des taux d'intérêt de marché) a atteint ses limites, car les taux directeurs de la BCE et l'EONIA sont proches de zéro. Depuis juin 2014, **le taux des facilités de dépôt** de la BCE et de l'Eurosystème, le taux plancher du corridor des taux directeurs, est **négatif**. Il s'établit à – 0,10 % en juin 2014 et -0,20 % en septembre 2014. C'est, d'une part, le prix de la sécurité pour la banque qui ne souhaite pas prêter alors qu'elle dispose d'importantes

liquidités en monnaie centrale et c'est, d'autre part, une incitation à prêter dans la mesure où un prêt peut lui rapporter de l'argent alors que l'alimentation de son compte à la banque centrale lui coûte de l'argent avec le taux négatif...

Il existe encore d'autres formules ponctuelles de refinancement via des appels d'offres : les opérations à plus long terme, les opérations de réglage fin et les opérations structurelles. Par ailleurs, via l'ELA (assistance d'urgence à la liquidité), les BCN peuvent prêter aux banques (qui restent viables), mais auxquelles la BCE ne prête plus. Quant à l'instrument des réserves obligatoires, il reste peu utilisé désormais.

La Banque Centrale européenne (BCE) lancera le 2 octobre 2019 son nouveau taux interbancaire au jour le jour, dénommé €STR (*Euro-Short-Term-Rate*). L'€STR, un temps dénommé Ester, est appelé à remplacer l'Eonia comme référence des taux des prêts entre banques de la zone euro et donc à servir de référence pour la fixation des prix de milliers de milliards d'euros d'instruments financiers. En réalité, la BCE a élaboré ce nouveau taux après l'échec d'une tentative de définition d'un nouvel instrument par le secteur bancaire lui-même.

20.5.2.2 Les objectifs des politiques monétaires non conventionnelles

Les politiques monétaires non conventionnelles correspondent à des modifications dans l'utilisation des instruments traditionnels, mais se traduisent également par l'emploi de nouveaux instruments et procédures pour rétablir durablement l'accès à la liquidité des banques, renforcer la confiance et retrouver un fonctionnement stable du marché monétaire. Les politiques monétaires non conventionnelles ont pour objectif de rétablir le fonctionnement normal du marché monétaire et l'efficacité des mécanismes de transmission de ses effets quand les canaux traditionnels sont bloqués.

– **Garantir la liquidité sur le marché interbancaire**

L'objectif premier est le maintien de la liquidité sur le marché interbancaire. Les mesures immédiates des banques centrales, dans leur rôle de prêteur en dernier ressort, face aux problèmes de liquidité que peuvent rencontrer les banques, sont cruciales et incontournables, mais insuffisantes.

– Maintenir la stabilité des prix et la stabilité financière

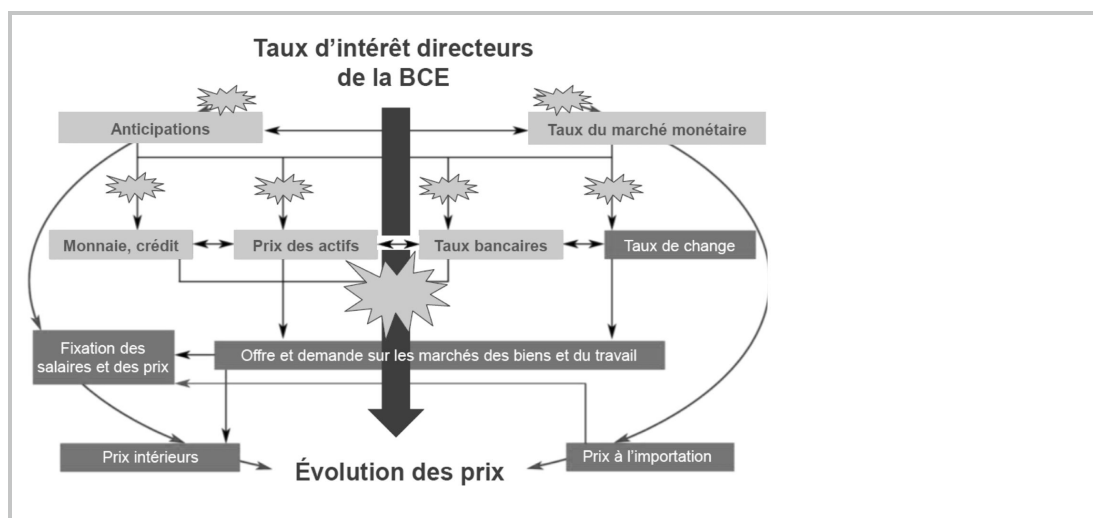
L'objectif de stabilité des prix par la lutte contre l'inflation reste un objectif de long terme, mais pour sortir les économies de la récession, d'autres objectifs sont devenus prioritaires dans les faits : éviter la déflation, relancer la croissance et l'emploi pour faire baisser le chômage, rétablir et maintenir la stabilité financière. En effet, le cumul des crises se traduit par une place plus grande accordée à l'objectif de stabilité financière par les banques centrales dans leur rôle de superviseur. Cependant, à ce jour, il n'y a pas dans la littérature économique de définition consensuelle de la stabilité financière. De même, il n'y a pas de mesure consensuelle de cette stabilité financière. Pour autant, selon la Banque des règlements internationaux (BRI), la stabilité financière peut être mesurée par divers indices synthétiques et composites, le ratio crédit/PIB, le Credit gap de la BRI et de la BCE, le niveau de levier des agents non financiers (FRB Chicago), le spread du crédit...

– L'interdépendance entre stabilité des prix et la stabilité financière

La stabilité des prix sur le marché des biens et services n'entraîne pas forcément la stabilité financière. En effet, la faiblesse de l'inflation peut inciter les banques centrales à mener des politiques monétaires expansives qui accroissent le volume du crédit et la volatilité du prix des actifs financiers. La doctrine relative au lien entre stabilité des prix et stabilité financière a nettement évolué. Selon l'hypothèse monétariste qui faisait consensus, la stabilité des prix est une condition suffisante de la stabilité financière. « La sagesse commune est qu'il n'y a pas de conflit entre stabilité des prix et stabilité

financière. » (Bini-Smaghi, 4 avril 2008) La crise financière et ses conséquences ont abouti à la remise en cause de l'hypothèse monétariste formulée par Anna Schwartz, car la Grande Modération de l'inflation entre 1995 et 2005 s'est terminée par la crise financière de 2007. La stabilité des prix n'est plus considérée comme une condition suffisante de la stabilité financière. Il existe un arbitrage, à court terme, entre stabilité des prix et stabilité financière. Avec les crises et l'incertitude croissante, l'arbitrage entre inflation, endettement et stabilité financière devient de plus en plus difficile pour les banques centrales. Les autorités monétaires ne peuvent plus se contenter de contrôler l'inflation sur le marché des biens et des services. Elles doivent prendre en considération avec beaucoup de rigueur l'évolution comparée de l'endettement et des prix des actifs financiers, immobiliers... Ainsi, pour éviter ou au moins réduire les effets sur l'économie réelle des crises financières, il convient de répondre sans attendre à la forte demande de liquidités des banques pour éviter une crise systémique, une situation de « credit crunch », une dégradation des bilans bancaires... Néanmoins, si le principe de séparation entre politique monétaire et stabilité financière semble de fait remis en cause, les principes d'une conciliation concrète paraissent bien difficiles à établir.

20.5.2.3 Les instruments et mécanismes de transmission de la politique monétaire non conventionnelle



GRAPHIQUE 20.2.

Les mécanismes de transmission des effets de la politique monétaire enrayés

Avec l'approfondissement et l'inscription dans la durée des crises, la BCE va devoir étendre l'éventail de ses interventions. En effet, quand les effets des politiques monétaires conventionnelles s'affaiblissent durablement, car les canaux de transmission des effets des mesures de politique monétaire sont bloqués, les banques centrales n'ont pas d'autres choix que d'utiliser d'autres instruments non standards.

À partir de fin septembre-début octobre 2008, la BCE décide d'élargir les modalités de refinancement et de soutien à l'économie. Ces mesures dites « non conventionnelles » auront pour principales conséquences d'accroître la taille de son bilan et de modifier sa composition. Des **risques d'inflation** et de **distorsions des prix** sur les marchés d'actifs sont évoqués par les observateurs.

STABILITÉ DES PRIX, ZERO LOWER BOUND ET FORWARD GUIDANCE

« Lorsque, suite par exemple à une crise financière, l'inflation très faible et que la banque centrale ne dispose que de faibles marges de manœuvre pour réduire son taux directeur, instrument usuel de la politique monétaire, le risque de heurter le plancher à zéro des taux d'intérêt (Zero lower bound - ZLB) se présente. [...] une spirale déflationniste peut alors s'enclencher. » (*Politique monétaire*, De Boeck Supérieur, 2015, p. 294)

Face à cette menace, les banques centrales peuvent agir sur les anticipations et les taux d'intérêt longs en utilisant le guidage des taux d'intérêt (Forward Guidance - FG) et les achats massifs de titres (Assouplissement quantitatif et qualitatif).

– Les différentes opérations de refinancement avec collatéral

L'évolution des modalités de refinancement de la BCE :

MRO (Main refinancing Operations) LTRO (Long Term Refinancing Operations), VLTRO (Very Long Term Refinancing Operations) puis TLTRO (Targeted Long Term Refinancing Operations).

On a : MRO => LTRO => VLTRO => TLTRO

1 semaine 3 mois 3 ans 4 ans

Ainsi, depuis 2010, avec la crise des dettes souveraines, face à la fragmentation du marché monétaire de la zone euro, au risque d'éclatement de la zone et de baisse des prix (déflation) dans certains pays, la mise en place de politiques monétaires non standards est venue compléter la panoplie des outils traditionnels de la BCE. Les TID étant quasiment nuls, les marges de manœuvre étaient inexistantes. La Banque Centrale européenne a donc conduit, en décembre 2011 et en février 2012, deux opérations massives de refinancement des banques européennes (VLTRO) en leur accordant des prêts à trois ans à des taux d'intérêt très bas, pour un montant total estimé à 1000 milliards d'euros. Mais cela n'a toujours pas suffi. En 2014, la BCE a décidé de mettre en place huit opérations de refinancement ciblées à plus long terme (TLTRO ou Targeted Longer Term Refinancing Operations), entre septembre 2014 et mai 2016..., interventions consistant à refinancer les banques sous conditions de prêts aux ménages et entreprises. Les TLTRO sont des opérations de refinancement au taux de 0,15 % jusqu'en janvier 2015, puis de 0,05 %, avec une échéance pouvant aller jusqu'à septembre 2018. Ces LTRO visent à renforcer l'activité de prêts bancaires aux entreprises et aux ménages.

– La fourniture de liquidités en devises

La BCE a aussi augmenté et facilité la fourniture de liquidités en devises, en particulier les échanges d'euros contre dollars via le programme *Liquidity Swap Lines*. Ce programme d'échange de devises entre la BCE et la FED vise à surmonter les difficultés que certaines banques opérant sur les marchés internationaux rencontraient pour se financer en devises, notamment en dollars. Ces échanges de devises répondent, à partir de la décision de la BCE du 13 octobre 2008, aux besoins temporaires de liquidités du système bancaire dans ces devises. La BCE cherche ainsi à assurer aux banques la fourniture de liquidités en devises

étrangères. Des *swaps* de change euros/francs suisses ont également été effectués.

– **Les programmes d'achats fermes de titres par la BCE**

En ce qui concerne les achats fermes de titres, il convient de distinguer six programmes d'achats : 1) de titres privés (obligations sécurisées) avec CBPP 1 (terminé) ou *Covered Bonds Purchase Programme* ; 2) de titres privés avec CBPP 2 (terminé) ; 3) de titres privés et publics avec SMP (terminé) ou *Securities Market purchases* ; 4) un SMP conditionnel avec l'OMT (jamais lancé) : *Outright Monetary Transactions* ; 5) titrisation de titres privés avec CBPP 3 (depuis fin 2014) ; 6) titrisation de titres privés avec ABSPP (depuis fin 2014) ou *Asset-Backed Securities Purchase Programme*.

Pour garantir la liquidité du marché de la dette souveraine, la BCE a mis en œuvre (mai 2010) le SMP, programme limité à l'achat de titres de dette publique sur le marché secondaire, car certains États devaient payer des taux très élevés pour pouvoir emprunter et s'enfonçaient donc dans la crise de la dette. Surtout, cela faisait peser un risque considérable sur la pérennité de l'euro et de la zone euro. Le programme OMT (opérations monétaires sur titres) devait lui succéder à partir de 2013. Mais ce dernier n'a jamais fonctionné, compte tenu des conditions très strictes et restrictives de son fonctionnement.

Conformément au Traité sur le fonctionnement de l'Union européenne (TFUE), les **achats d'obligations d'État par l'Eurosystème** doivent être strictement **limités au marché secondaire** de la dette et ne pas augmenter le niveau de la monnaie centrale, pour ne pas affecter l'orientation de la politique monétaire de la BCE, dont l'objectif principal demeure la stabilité des prix dans la zone euro (pour garantir le pouvoir d'achat de la monnaie).

Concrètement, via ces différents **programmes d'achat d'actifs**, la BCE cherche à stabiliser le fonctionnement des

marchés, en particulier celui des dettes souveraines. Elle joue alors un rôle de teneur de marchés. En effet, pour réduire le ratio dette/PIB, les États doivent réduire les déficits budgétaires (ce qui est très difficile en période de récession), sans pouvoir compter sur une accélération de l'inflation et sans parier sur l'annulation d'une partie de la dette, car cela serait potentiellement nuisible à long terme en risquant de dissuader les investisseurs de prêter, et doivent permettre de faire baisser les taux d'intérêt de long terme et de faciliter le désendettement privé et public. De surcroît, la faiblesse des taux d'intérêt, du fait de la politique monétaire durablement expansionniste de la BCE, en l'absence d'inflation, facilite le désendettement à long terme des États et leur capacité à se refinancer à court terme. En période de croissance durablement « molle » (autour de 1 % du PIB), de chômage élevé et de perspectives pessimistes, les taux à long terme sont bas, car il n'y a pas d'anticipation de l'inflation et les investisseurs sont plutôt gourmands en titres de dette publique !

– **L'assouplissement quantitatif** ou « quantitative easing » (QE) **de la BCE**

En 2015, la BCE passe au « quantitative easing » (QE), sur le modèle de la stratégie américaine de la Réserve fédérale des États-Unis, pour renforcer le caractère expansionniste de la politique monétaire. Comme d'autres banques centrales et à l'instar de la Réserve fédérale américaine, qui a mis en œuvre plusieurs plans successifs de « quantitative easing », la BCE va racheter de la dette, à la fois publique (des obligations souveraines émises par les États de la zone euro et les institutions européennes) et privée. Dans le cadre de ce programme inédit, la BCE a injecté en 2016 l'équivalent de 60 milliards d'euros par mois, puis 80 milliards d'euros par la suite. En **mars 2018**, la BCE a toutefois renoncé à la possibilité d'augmenter le volume mensuel de ses rachats de dettes publiques et privées. Par un communiqué officiel, la BCE a amorcé une décrue irréversible du « *quantitative easing* »

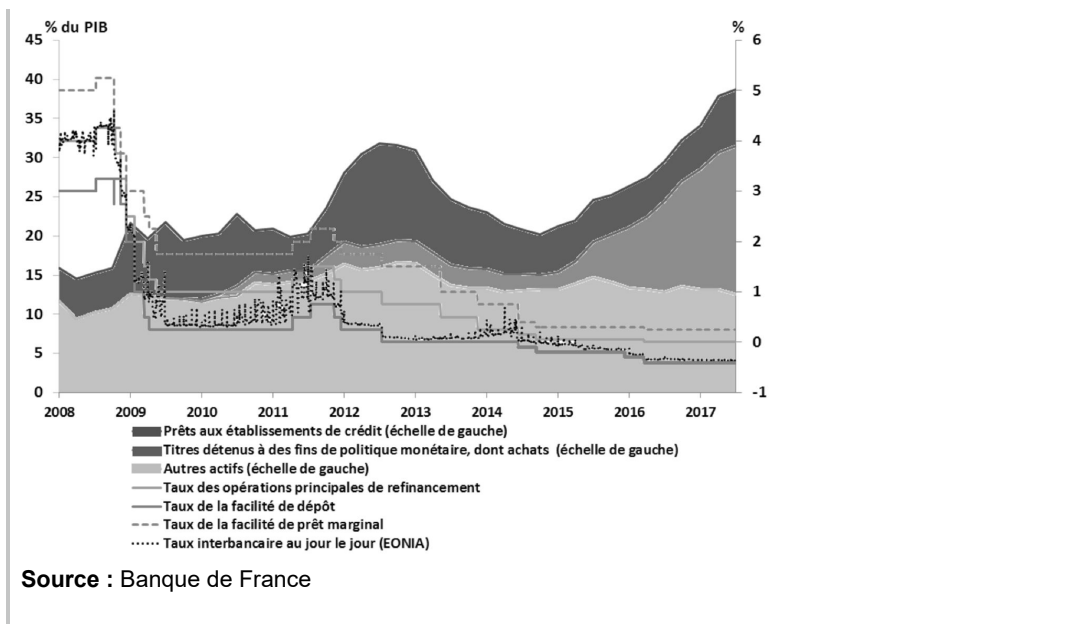
(QE), ce que les experts nomment le « tapering », passé en avril 2016 de 80 milliards à 60 milliards d'euros mensuels, puis tombé à 30 milliards d'euros par mois depuis janvier. Ce volume de rachat d'actifs représente, depuis son lancement, en mars 2015, un volume d'environ 2 400 milliards d'euros.

La BCE a progressivement stoppé le QE à la fin de l'année 2018, mais en mars 2019, face au risque d'un ralentissement de l'économie en zone euro, Mario Draghi, le président de la Banque Centrale européenne, a annoncé que la BCE avait décidé de repousser la hausse des taux d'intérêt à 2020 et d'accorder des prêts massifs aux banques dans le cadre de nouvelles mesures de soutien. En effet, à partir du mois de septembre 2019 et jusqu'en mars 2021, la BCE lancera un nouveau programme de prêts massifs aux banques avec pour objectif d'injecter des liquidités dans le circuit économique. Ces prêts, à très faible taux, de 0 à 0,40 %, vont permettre de soutenir les prêts à l'économie réelle et d'accorder des prêts aux particuliers, pour relancer la consommation, mais également aux entreprises pour leur permettre de poursuivre les investissements productifs.

Depuis 2008, on a pu observer la concomitance entre, d'une part, la baisse des taux d'intérêt directeurs et du taux à très court terme du marché interbancaire et, d'autre part, l'augmentation de la taille du bilan de la BCE liée à la croissance des achats fermes de titres dans le cadre des opérations d'assouplissement quantitatif. Les signes du caractère très expansionniste de la politique monétaire de la BCE sont très clairs.

Le bilan de la BCE s'élevait à 1154 milliards en janvier 2007, et atteindra un maximum de 3099 milliards en 2012. Depuis, la taille du bilan de la BCE est redescendu à 2029 milliards (novembre 2014). En janvier 2019, la taille du bilan de la BCE atteint 4694,3 milliards d'euros, soit 42 % du PIB de la zone euro.





GRAPHIQUE 20.3.

Évolution des taux d'intérêt directeurs et du bilan de la BCE entre 2008 et 2018

- Les **effets attendus des programmes de « quantitative easing »** : effet « liquidités », effet « devise » et effet « crédit ».

Selon les économistes, l'efficacité du programme de « quantitative easing » (QE) repose sur trois canaux principaux de transmission.

L'effet « **liquidités** » : la BCE, en rachetant massivement aux institutions financières des titres de dette publique, améliore les conditions de financement des États et maintient durablement les taux à des niveaux très bas. Elle incite également les investisseurs à réallouer leurs portefeuilles vers des actifs plus risqués, si bien que c'est l'ensemble des actifs financiers qui bénéficient au final de l'injection de liquidités.

L'effet « **devise** » : la monnaie additionnelle créée, rémunérée à des taux plus faibles, induit mécaniquement une dépréciation de l'euro sur le marché des changes, favorable à la croissance de la production et aux marges des exportateurs de la zone euro.

Enfin, la BCE favorise l'effet « **crédit** ». En éliminant du bilan des banques une quantité d'actifs divers, la banque centrale donne aux prêteurs les moyens d'augmenter l'offre de crédit, tout en en réduisant le coût.

– **Risques et incertitudes sur les effets du quantitative easing (QE) de la BCE**

Le QE est un programme d'achats, par la BCE, de titres de dette souveraine ou de titres de dette émis par des institutions publiques garanties par les 19 États de la zone euro. Ce QE pourrait ne pas avoir tous les effets escomptés, pour deux raisons principalement. De nombreux commentateurs et investisseurs ont évoqué le risque que la BCE puisse être confrontée à une possible insuffisance des titres de dette éligibles au rachat, de titres d'État à acheter. Mais la possibilité d'acheter des titres à rendement négatif élargit son champ d'action et les chances de succès du QE. De surcroît, l'appréciation assez positive portée sur son programme devrait crédibiliser son objectif. De plus, l'ampleur de l'impact du QE sur l'économie réelle (la consommation, l'investissement, la production) relève de la plus pure incertitude. Pourtant, un redémarrage soutenu et durable des crédits au secteur privé serait un soutien solide pour l'activité. L'effet sur les spreads de taux, écarts de taux par rapport au Bund (l'obligation d'État allemande qui est l'équivalent de l'OAT, l'obligation de l'État français), est également très incertain. En ce qui concerne la France, les achats de la BCE devraient, d'une part, soutenir le marché immobilier et, d'autre part, favoriser la poursuite du repli des taux d'emprunt de l'État et des agences (CADES, UNEDIC), permettant d'éviter de nouveaux gels des prestations sociales et de faciliter le redémarrage de la consommation.

– **Les effets attendus du programme d'achat d'actifs**

Cette stratégie fonctionnera d'autant mieux qu'elle est accompagnée de réformes pour améliorer le fonctionnement de l'économie. Une augmentation de la quantité de monnaie en circulation dans l'économie. Une baisse du coût d'emprunt

à long terme dans tous les États de la zone euro. Des banques qui redonnent la priorité aux prêts à l'économie. Plus généralement, un soutien à la confiance.

20.5.3 Les controverses théoriques sur les politiques monétaires non standard

D'un point de vue théorique, on peut distinguer trois types de politiques monétaires non conventionnelles : l'assouplissement quantitatif (*quantitative easing*), l'assouplissement qualitatif (*qualitative easing ou credit easing*) et la politique de guidage des taux d'intérêt (*forward guidance*).

20.5.3.1 L'assouplissement quantitatif (*quantitative easing*)

C'est un instrument qui agit sur la **taille du bilan** (actif avec les titres cédés par les banques et passif avec les liquidités accordées en contrepartie) de la banque centrale. L'objectif est d'assouplir toujours davantage l'accès à la liquidité en améliorant les conditions de refinancement pour les banques : hausse des montants et des durées de financement, élargissement de l'éventail des titres acceptés en collatéral... **La base monétaire augmente rapidement.** Les risques d'inflation sont potentiellement plus élevés, car selon la logique du multiplicateur monétaire, la hausse de la base monétaire pourrait entraîner une hausse de la masse monétaire et donc de l'inflation si les banques augmentent le montant des crédits accordés... Cette politique de QE fut déjà menée au Japon entre le 19 mars 2001 et le 9 mars 2006. La BCE a davantage utilisé cette modalité, car la structure des financements est plus orientée vers les financements indirects et bancaires. La BCE qualifie d'ailleurs son action de « politique de soutien renforcé au crédit. »

Bilan de la banque centrale		Bilan des banques commerciales	
Actif	Passif	Actif	Passif
Achat de titres :	Réserves : + 100	Vente de titres : –	M3 inchangé

+ 100 non toxiques	Liquidité bancaire	100 Réserves : – 100	

20.5.3.2 L'assouplissement qualitatif (*qualitative easing*)

C'est un instrument qui agit sur la **composition du bilan** de la banque centrale. La taille du bilan reste inchangée. La base monétaire reste stable. Les risques d'inflation sont inchangés, mais la qualité du bilan est dégradée. En effet, la BC achète des titres directement aux banques ou sur le marché pour améliorer leur liquidité et soutenir leur prix à la hausse. Cela conduit la BC à accepter de plus en plus de titres de qualité réduite et plus risqués. Cependant, ces interventions sont temporaires et la hausse des prix des actifs peut entraîner des baisses des taux longs et provoquer des effets de richesse favorables au soutien de l'économie via le canal du prix des actifs. La FED et la Banque d'Angleterre ont davantage utilisé cette modalité, car la structure des financements dans ces économies est plus orientée vers les financements directs sur les marchés.

Bilan de la banque centrale		Bilan des banques commerciales	
Actif	Passif	Actif	Passif
Vente de titres : – 100 de qualité Achat de titres : + 100 toxiques		Vente de titres : – 100 toxiques Achat de titres : + 100 de qualité	

20.5.3.3 La politique de guidage des taux d'intérêt (*forward guidance*)

Elle consiste en un pilotage des anticipations. *Le forward guidance* est une politique qui fournit aux ménages, entreprises et investisseurs, des détails concernant le déroulement à venir de la politique monétaire. Cela signifie que les autorités monétaires communiquent à l'avance le niveau des taux d'intérêt directs futurs. Ainsi, en situation de crise, les banquiers centraux s'engagent à conserver des taux directs très faibles, généralement proches de 0 %. En

guidant les anticipations des taux d'intérêt courts, les banques centrales peuvent agir sur les taux longs, seule issue lorsque les taux courts sont déjà à 0 %.

Selon la BCE, il s'agit de « *fournir des indications relatives à la trajectoire future des taux directeurs de la BCE conditionnée par les perspectives de stabilité des prix.* » (Bulletin mensuel, avril 2014) Quand le principal taux d'intérêt directeur est à son niveau plancher de 0 % et reste durablement fixe à ce niveau, l'objectif est d'ancrer les anticipations d'inflation et de taux d'intérêt pour mieux contrôler la stabilité des prix à court terme et les taux d'intérêt à long terme. Comme l'explique clairement Jezabel Couppey-Soubeyran, « *cette stratégie est fondée sur la théorie des anticipations de la courbe des taux (les taux longs dépendent des anticipations de taux courts) et sur le **canal de transmission des anticipations*** » (in Monnaie, Banques, Finance, PUF, 2018, p. 264).

Dans son discours du 4 juillet 2013, le président de la BCE Mario Draghi a entamé cette politique de *forward guidance* des taux d'intérêt en déclarant : « *Le Conseil des gouverneurs prévoit que les taux d'intérêt directeurs de la BCE resteront à leurs niveaux actuels ou à des niveaux plus bas sur une période prolongée. Cette anticipation se fonde sur le maintien, à moyen terme, de perspectives d'une inflation globalement modérée, compte tenu de la faiblesse généralisée de l'économie réelle et de l'atonie de la dynamique monétaire. Au cours de la période à venir, le Conseil des gouverneurs suivra toutes les informations relatives aux évolutions économiques et monétaires qui lui parviendront et évaluera leur incidence éventuelle sur les perspectives de stabilité des prix.* »

LES CRISES FINANCIÈRES ET LES RÉCESSIONS ASSOCIÉES PEUVENT ÊTRE LA CONSÉQUENCE DES POLITIQUES MONÉTAIRES

En se basant sur le cas des États-Unis, Patrick Artus montre que depuis la fin des années 1990, les crises financières et les récessions associées viennent systématiquement de la politique monétaire de la Banque centrale américaine (FED ou Réserve fédérale). Ainsi, à la fin des années 1990, la FED, en menant une politique monétaire très expansionniste, a facilité le gonflement d'une bulle sur le cours des actions aux États-Unis. En 2000-2001, cette bulle a explosé et provoqué une crise financière et

une récession globale. Entre 2002 et 2007, la Réserve Fédérale a participé, via une politique monétaire très expansionniste et au développement incontrôlé de la titrisation des crédits immobiliers (avec des primes de risque anormalement basses) au gonflement d'une bulle immobilière aux États-Unis. L'éclatement de cette bulle immobilière à partir de 2007 a déclenché la crise financière et économique mondiale, dite crise des *subprimes*. Le laxisme monétaire, en maintenant des taux d'intérêt durablement bas, facilite l'endettement excessif, la formation de bulles financières qui, en éclatant, provoquent des crises financières et économiques.

Source : « Les crises viennent toujours de la politique monétaire des États-Unis », Patrick Artus, Natixis, 18/01/2018

20.6 LES ÉVOLUTIONS INSTITUTIONNELLES DE LA ZONE EURO DEPUIS 2010 ET LES PERSPECTIVES

20.6.1 *Les mécanismes de solidarité budgétaire pour faciliter le financement des États*

D'un point de vue plus institutionnel, les États et la BCE ont mis en place en 2010 un mécanisme de solidarité entre États : le **FESF**. La crise grecque du printemps 2010 a conduit l'UE, confrontée pour la première fois au risque de défaut d'un des pays membres, à mettre en place le Fonds européen de stabilité financière (FESF), pour une durée de trois ans au-delà du prêt consenti à la Grèce. Le FESF est réservé aux pays membres de la zone euro. C'est une structure de financement ad hoc pouvant mobiliser jusqu'à 440 milliards d'euros. Les ressources du FESF sont des titres de dette émis sur les marchés qui sont garantis par les États, au prorata de leur participation au capital de la BCE.

L'UE avait, par ailleurs, mis à la disposition des 27 pays membres une facilité de prêt pouvant atteindre 60 milliards d'euros : le **MESF**. Le dispositif du FESF peut donc être « associé à des prêts accordés par le mécanisme européen de stabilité financière (MESF, dont le financement est assuré par des fonds levés par la Commission européenne et garantis par le budget de l'UE), à concurrence de 60 milliards d'euros, et par le FMI à hauteur de 250 milliards d'euros (soit

la moitié des fonds mis à disposition par l'Europe). » Le FESF a été pérennisé en devenant le **MES** (Mécanisme européen de stabilité) en 2013. Les réformes institutionnelles ont été nombreuses depuis 2010, même si tout n'a pas été fait. De nombreux textes ont été adoptés. Les objectifs étaient multiples : améliorer l'efficacité des instruments de contrôle et de coordination des politiques économiques dans l'optique d'un rééquilibrage du « policy mix » dans la zone euro ; repenser et redéfinir le cadre de surveillance macroéconomique pour le rendre plus pertinent et cohérent avec la nature hétérogène de la zone euro (indicateur de balance des paiements, compétitivité, stabilité financière.) Avec le **Pacte Europlus**, le but est de renforcer la gouvernance de l'Union économique et monétaire, en déplaçant le curseur d'une discipline des pairs (surveillance mutuelle) vers une discipline de marché et des règles institutionnelles. On peut encore retenir le **TSCG**, le cadre de surveillance fourni par le « semestre européen », le « two pack ».

20.6.2 *L'Union bancaire pour améliorer la solidarité entre États en cas de crise bancaire*

La crise actuelle de la zone euro est aussi une crise bancaire, pas seulement une crise de la dette souveraine. Dans l'ensemble, les banques européennes ont été insuffisamment assainies et recapitalisées après la crise financière de 2008. Cependant, la crise des dettes publiques a accentué leurs difficultés, en particulier dans les pays qui se sont enfoncés dans la récession ou la dépression économique (Grèce, Espagne, Portugal). Les capitaux fuient les pays les plus fragilisés. Les banques sont poussées à se défaire de leurs actifs. Cela risque de se traduire par une contraction de l'offre de crédit, qui pèsera davantage sur des pays déjà fragilisés par la crise. Le projet d'union bancaire a été adopté lors du sommet européen de mi-octobre 2012. Il prévoyait

une mise en place progressive jusqu'en mars 2014. En effet, la crise de la zone euro a mis en lumière une interdépendance croissante entre les systèmes bancaires nationaux d'une part, et entre risque bancaire et risque souverain d'autre part. Établir une véritable union bancaire et financière permettra de casser le lien entre risque bancaire et risque souverain et d'éviter le risque de fragmentation de la zone euro.

L'Union bancaire repose sur quatre piliers : un superviseur bancaire unique ; un système de garantie collective des dépôts ; un système commun de gestion et de résolution des défaillances bancaires ; un fonds commun de résolution des crises et de recapitalisation bancaire (jugé nécessaire par certains). Sur chacun des piliers comme pour l'ensemble, les difficultés de mise en œuvre sont importantes. Elles tiennent notamment aux réticences de l'Allemagne concernant les mécanismes de solidarité financière et de mutualisation des dettes bancaires puisque, sur ce point, des règles de gouvernance n'ont pas encore été suffisamment mises en place au niveau européen.

20.6.3 *Le fédéralisme budgétaire en question*

La réforme de la gouvernance de l'euro pourrait passer également par un fédéralisme budgétaire et une harmonisation fiscale (voire la création d'un impôt européen ou la mise en commun de certaines parties de la fiscalité, à l'instar de la TVA), qui permettrait de consolider la solidarité financière dans la zone euro. Mais cette perspective de gouvernement économique européen ne fait pas consensus sur le plan politique, dans la mesure où elle nécessiterait, au nom du projet d'union politique de l'Europe, qui reste controversé, l'acceptation de nouveaux transferts de souveraineté consentis par les États-nations (en particulier une union budgétaire). Certains évoquent même l'opportunité d'avancées institutionnelles vers le fédéralisme en matière de système de retraites ou de système d'indemnisation chômage, par exemple, pour éliminer certaines distorsions de

compétitivité. Le fédéralisme dans la zone euro pourrait passer par une émission en commun de dettes publiques (qu'on appelle *eurobonds*). En effet, si les pays émettaient leur dette publique en commun, il se produirait potentiellement un transfert des pays solvables vers les pays qui seraient en difficulté, par le biais de la solidarité entre les emprunteurs. La mutualisation des dettes publiques pourrait ainsi réduire les coûts d'emprunt dans le cadre d'un marché obligataire suffisamment profond et liquide libellé en euros (avec la possibilité de distinguer les titres de cette dette commune en différentes tranches de risque). La Commission européenne pourrait ainsi émettre des obligations garanties par le budget européen pour financer des investissements publics (infrastructures, réseaux de communication), permettant d'élever la croissance potentielle de la zone euro. En 2010, l'émission d'obligations communes par le Fonds européen de stabilité financière (FESF), un fond commun créé en réponse à la crise de l'euro devant lever directement des fonds sur les marchés financiers pour le compte des États membres de la zone euro, a constitué un embryon de mutualisation des dettes souveraines au sein de la zone euro. Depuis, Le Mécanisme européen de stabilité (MES²²) lui a succédé.

Aujourd'hui, la capacité de la zone euro à devenir une zone monétaire optimale (ZMO) reste incertaine. Comme l'a établi Robert Mundell dès 1961, plusieurs pays constituent une zone monétaire optimale si l'usage d'une monnaie unique ne dégrade pas leur bien-être. Pour cela, les chocs qu'ils subissent doivent être homogènes et les facteurs de production (d'abord le travail) mobiles entre les frontières. À ces deux critères initiaux ont été ajoutées la possibilité de transferts budgétaires et financiers (Kenen, Ingram) entre les pays et la diversification nécessaire des appareils productifs (Kenen, McKinnon). L'intégration économique et financière n'étant pas parfaite, il aurait fallu prévoir des mécanismes de transferts budgétaires puissants pour faire face à des chocs asymétriques, puisque la possibilité d'ajustement par les taux

de change avait disparu avec la création de l'euro. Ces transferts pouvaient être remplacés par un partage des risques au travers de marchés financiers intégrés. L'intégration financière incomplète faisait du fédéralisme budgétaire la meilleure assurance contre les écarts de cycle. La crise, en provoquant la fragmentation des systèmes bancaires et l'explosion des déficits et des dettes, a fait exploser l'homogénéité des préférences (scission nord-sud), qui aurait pu soutenir le volontarisme politique pour approfondir l'union politique. Sur le plan budgétaire, tout en décalant pour certains pays (dont la France) les échéances du retour à un déficit de 3 %, la Commission n'a pas opté pour autant pour des actions de relance, arguant du fait que l'important reste la réduction du déficit structurel (c'est-à-dire corrigé du cycle économique) et que celle-ci est incontournable pour contrôler l'endettement public. Mais le problème est que le PIB potentiel, qui permet de définir ce déficit structurel, est affecté par les politiques de rigueur : de fait, plus la rigueur est forte et moins les pays sont à même de satisfaire à la « règle d'or », car les investissements non réalisés et les pertes de capital humain liées au chômage pèsent sur la croissance potentielle.

Pour de nombreux analystes, l'absence de convergence des économies en matière de taux d'intérêt (les pays en difficultés ayant des taux plus élevés) et d'inflation (les mêmes ayant une inflation plus basse) ne pourrait être contrée que par les progrès rapides d'une union bancaire et budgétaire. Favorisant l'intégration financière et la solidarité budgétaire, celle-ci pourrait rendre l'Union plus solidaire, si ce n'est en faire une zone monétaire optimale. Or, le gouvernement allemand subordonne ces éventuels progrès à une modification des traités qui sera longue à obtenir, surtout face à l'ombre croissante de la déflation.

20.6.4 *La politique monétaire de la BCE au risque de déflation en zone euro*

Depuis 2010, la zone euro fut durablement menacée par le risque déflationniste, la stagnation de la croissance et la persistance du chômage de masse. En dépit des stimulations monétaires de la Banque Centrale européenne, l'inflation restait très faible dans la zone euro. Depuis l'entrée en fonction de Mario Draghi en 2010, la BCE a baissé son taux directeur à plusieurs reprises pour le ramener à un minimum historique. Cette baisse s'est imposée face à la persistance de la récession dans la zone. Cette situation n'est pas propre à l'Europe, car les États-Unis et le Japon connaissent des dynamiques comparables et les interdépendances entre les zones renforcent les pressions à la baisse des prix, à l'image de ce qui s'est passé dans les années 1930. Cependant, la zone euro est plus exposée à la déflation, car le taux d'inflation que cherche à stabiliser la BCE est un taux moyen, ce qui la dispense d'agir plus fortement face à des évolutions nationales plus négatives, mais aussi parce que la gestion de la crise de la zone passe d'abord par une politique de dévaluation interne (par baisse des prix et des salaires), ce qui fait de la déflation, pour un nombre de pays de plus en plus élevé, une des conditions paradoxales de leur sortie de crise. Toutefois, la « BCE ne peut pas tout » (pour reprendre la formule de l'économiste Charles Wyplosz) et une relance budgétaire (avec éventuellement un budget européen plus consistant) resterait une manière plus directe de stimuler l'investissement productif et de favoriser le retour de la croissance, en combinant celle-ci à des achats de la dette publique primaire si nécessaire. Mais il est évident que l'ensemble de ces mesures se heurte pour l'instant à des résistances très fortes dans le cœur de la zone, à commencer par l'Allemagne et sa banque centrale, la Bundesbank.

20.6.5 Vers davantage d'intégration par une politique macroéconomique d'ensemble

Dans la note du Conseil d'Analyse économique (CAE) publiée en mars 2015 et intitulée « Pour une politique macroéconomique d'ensemble en zone euro », Agnès Bénassy-Quéré (présidente déléguée du Conseil d'analyse économique) et Xavier Ragot (président de l'Observatoire français des conjonctures économiques) insistent sur les défaillances du pilotage macroéconomique propre à la zone euro. Ils examinent ce qui n'a pas fonctionné depuis 1999.

Ils montrent que ce pilotage est resté « *polarisé sur la surveillance budgétaire sans prêter suffisamment attention aux divergences nominales (différences de salaire, de compétitivité...), aux déséquilibres financiers nourris par la dette privée et à la nécessité d'une gestion d'ensemble de la demande en zone euro* ». Sur la base de ce constat, ils formulent « *huit recommandations pragmatiques pour corriger les faiblesses de la gouvernance macroéconomique actuelle de la zone euro, tout en plaidant pour davantage d'intégration à moyen terme* ».

Si la coordination des politiques économiques est par nature difficile entre États souverains, « *il reste possible d'améliorer significativement le dispositif de coordination actuel organisé autour du "semestre européen"* ». En effet, dans le cadre de ce dispositif trop complexe, le diagnostic d'ensemble ne permet pas de formuler des recommandations adaptées à des pays dans des situations économiques hétérogènes. La surveillance est, en outre, compliquée et asymétrique, et le calendrier est discutable. Les auteurs proposent deux « trimestres » européens. Un premier trimestre de surveillance de la zone euro dans son ensemble, au-delà des situations des pays constitutifs. Ils recommandent une surveillance combinée des politiques budgétaires, politique de compétitivité, politique de demande et politique macroprudentielle. Ils expliquent comment les États membres pourraient contribuer à améliorer les processus de

coordination de leurs décisions et actions. En résumé, les huit recommandations peuvent se rassembler autour de deux grands axes :

- simplifier et mieux intégrer la surveillance entre la zone euro et les États membres ;
- permettre une meilleure appropriation par les États membres et renforcer le diagnostic partagé.

Pour les auteurs, le doute n'est plus permis et les faux débats n'ont plus de raison d'être, car « *Les instruments et l'expertise existent. L'enjeu est de mieux les utiliser et de les rendre plus lisibles* ». De surcroît, « *Les recommandations formulées [...] doivent être lues comme une première étape vers une plus grande intégration, indispensable au succès de l'Union économique et monétaire.* »

Cependant, au-delà des réformes et priorités de court terme, il convient de garder à l'esprit les enjeux de long terme quand on réfléchit au fonctionnement de la zone euro. Dans le cadre de la stratégie « Europe 2020 », il semble impératif de renforcer la croissance potentielle de l'Union européenne, mais aussi de repenser les conditions d'entrée dans la zone euro (à la lumière de la crise grecque et de l'expérience des douze premières années de l'euro) et d'avancer sur le dossier épineux de la coordination des politiques fiscales et du fédéralisme budgétaire.

²² Le Fonds européen de stabilité financière (FESF) ainsi que le Mécanisme européen de stabilité financière (MESF) ont été créés dans l'urgence en 2010 pour faire face à la gravité de la crise économique et financière. Le MESF concernait l'ensemble des pays de l'Union européenne. Le FESF était réservé aux membres de la zone euro. Le FESF et le MESF étaient des instruments temporaires. Ils ont été fusionnés pour donner naissance au MES qui reprend leurs fonctions. Le Mécanisme européen de stabilité (MES) est un fonds commun de créances pour faire face à la défaillance potentielle d'un État membre de la zone euro. Il est entré en vigueur le 27 septembre 2012 et a commencé à fonctionner le 8 octobre 2012.

Ouvertures Économiques

Série LMD

Cette série, déclinaison de la réputée collection **Ouvertures économiques**, offre à l'étudiant en économie et en gestion des manuels pratiques, directement en phase avec le cursus universitaire, qu'ils soient de niveaux Licence (Baccalauréat), Master ou Doctorat. Ces manuels ont été conçus dans un esprit tout à la fois synthétique et pédagogique.

Abdelmalki Lahsen, Sandretto René, *Le commerce international. Analyses, institutions et politiques des États.*

Abdelmalki Lahsen, Sandretto René, *Politiques commerciales des grandes puissances. La tentation néoprotectionniste.*

Barbier-Gauchard Amélie, *Intégration budgétaire européenne.*

Enjeux et perspectives pour les finances publiques européennes.

Barbier-Gauchard Amélie, Sidiropoulos Moïse, Varoudakis Aristomène, *La gouvernance économique de la zone euro. Réalités et perspectives.*

Béal Sylvain, Gabuthy Yannick, *Théorie des jeux coopératifs et non coopératifs. Application aux sciences sociales.*

Beffy Pierre-Olivier, *Initiation à l'économie.*

Compléments en ligne : QCM-exercices.

Brillet Jean-Louis, Cette Gilbert, Gambini Ian, Lagoarde-Segot Thomas, *Politiques macroéconomiques. Un outil d'évaluation : MacSim2.*

Burgenmeier Beat, *Politiques économiques du développement durable.*

Caulier Jean-François, *Fondements mathématiques pour l'économie et la gestion.*

Cayatte Jean-Louis, *Microéconomie de l'incertitude.*

Dumas André, *Économie mondiale. Les règles du jeu commercial, monétaire et financier.*

Eber Nicolas, *Introduction à la microéconomie moderne. Une approche expérimentale.*

Faucheux S., Hue C., Nicolaï I., *T.I.C. et développement durable. Les conditions du succès.*

Goutte Stéphane, *Mathématiques financières. Théorie, exercices et simulations numériques.*

Lupton S., *Économie des déchets. Une approche institutionnaliste.*

Mussard Stéphane, Seyte Françoise, *Inférence statistique et probabilités.*

Poudon Jean-Christophe, Thomas Lionel, *Optimisation pour l'analyse économique et les sciences de gestion.*

Schwengler Bernard, *Les règles budgétaires. Un frein à l'endettement ?*

Tazdaït Tarik, *L'analyse économique de la confiance.*

Terraza Virginie, Toque Carole, *Analyse statistique pour la gestion bancaire et financière. Applications avec R.*

Pour toute information sur notre fonds et les nouveautés dans votre domaine de spécialisation, consultez notre site web : **www.deboecksuperieur.com**

Cette œuvre est protégée par le droit d'auteur et strictement réservée à l'usage privé du client. Toute reproduction ou diffusion au profit de tiers, à titre gratuit ou onéreux, de tout ou partie de cette œuvre, est strictement interdite et constitue une contrefaçon prévue par les articles L.335-2 et suivants du Code de la Propriété Intellectuelle. L'éditeur se réserve le droit de poursuivre toute atteinte à ses droits de propriété intellectuelle devant les juridictions civiles ou pénales.

© De Boeck Supérieur s.a, 2019
Rue du Bosquet, 7 – B-1348 Louvain-la-Neuve

« La seule manière d'apprendre utilement l'économie est bien celle proposée dans ce livre : l'interaction entre l'analyse historique des faits et des institutions, la réflexion théorique, l'utilisation des modèles, la réflexion sur les controverses entre économistes. »

Patrick Artus (extrait de la préface)

L'actualité économique est foisonnante et difficile à appréhender sans avoir les clés de lecture nécessaires. Les problèmes économiques sont toujours complexes, les analyses nombreuses et parfois contradictoires, et les décisions politiques se multiplient, selon des intérêts parfois divergents...

Dans ces conditions, **comprendre les mécanismes fondamentaux du fonctionnement de l'économie** se révèle plus que jamais indispensable. En tant que discipline scientifique, l'économie vise à proposer – au-delà des opinions, préférences idéologiques et préjugés – un ensemble de savoirs construits selon une démarche rigoureuse, afin d'échapper à tous les poncifs récurrents et argumentations trop pressées sur de nombreux sujets : croissance, emploi, pouvoir d'achat, entreprise, mondialisation, Europe, environnement, inégalités...

Afin de faciliter la tâche à tous ceux qui veulent mieux comprendre le monde dans lequel ils vivent et où l'économie occupe une si grande place, les auteurs proposent un ouvrage dont la principale qualité doit être l'**clarté pédagogique** et la **rigueur de l'analyse**, tout en suscitant l'intérêt du lecteur et en donnant du sens.

Cet **ouvrage introductif et généraliste**, cours complet sur les fondamentaux de l'économie, s'adresse plus particulièrement aux étudiants en sciences économiques, aux étudiants des classes préparatoires économiques et commerciales et des IEP. Il intéressera également les candidats aux concours administratifs (agrégations, CAPES, CAPET).

David Mourey est professeur de SES au lycée Charles Le Chauve à Roissy-en-Brie, professeur de TD de Politique économique à l'UFR d'économie-gestion de Marne-la-Vallée, professeur à l'ISMaPP pour un cours d'Introduction à l'économie, formateur en SES dans l'académie de Créteil, fondateur, organisateur et modérateur des « Rencontres économiques » de Pontault-Combault et de Paris.

Laurent Braquet est professeur de SES au lycée Gustave Flaubert de Rouen, formateur en SES dans l'académie de Rouen. Il enseigne également en classes préparatoires au concours commun des IEP.

deboeck
SUPÉRIEUR

www.deboecksuperieur.com

L Dans le cadre du Système
Européen de Transfert
de Crédits (E.C.T.S.), ce manuel
couvre le niveau Licence
(Baccalauréat/Bachelor).