

Économétrie appliquée

Méthodes – Applications – Corrigés

2^e édition

Isabelle Cadoret – Catherine Benjamin
Franck Martin – Nadine Herrard – Steven Tanguy


◀ ▶ ÉCONOMIQUES

OUVERTURES

Site
internet



de boeck



Digitized by the Internet Archive
in 2023 with funding from
Kahle/Austin Foundation

<https://archive.org/details/economtrieappliq0000isab>

Économétrie appliquée

OUVERTURES ◀▶ ÉCONOMIQUES

Économétrie
appliquée

Jacques-Henri Jeannerot

Économétrie appliquée

Méthodes – Applications – Corrigés

2^e édition

Isabelle Cadoret – Catherine Benjamin
Franck Martin – Nadine Herrard – Steven Tanguy

◀▶ ÉCONOMIQUES

OUVERTURES



de boeck

Pour toute information sur notre fonds et les nouveautés dans votre domaine de spécialisation, consultez notre site web : www.deboeck.com

© Groupe De Boeck s.a., 2009
Éditions De Boeck Université
Rue des Minimes 39, B-1000 Bruxelles

2^e édition

Tous droits réservés pour tous pays.

Il est interdit, sauf accord préalable et écrit de l'éditeur, de reproduire (notamment par photocopie) partiellement ou totalement le présent ouvrage, de le stocker dans une banque de données ou de le communiquer au public, sous quelque forme et de quelque manière que ce soit.

Imprimé en Belgique

Dépôt légal:
Bibliothèque Nationale, Paris : avril 2009
Bibliothèque royale de Belgique, Bruxelles : 2009/0074/010

ISSN 2030-501X
ISBN 978-2-8041-5976-4

AVANT-PROPOS

La réalisation de cet ouvrage d'économétrie appliquée résulte de l'expérience des auteurs dans l'enseignement de cette matière en deuxième et troisième cycle à la Faculté des Sciences Economiques de l'Université de Rennes 1 et de leur recherche en économie appliquée au sein du CREM, Unité Mixte de Recherche du CNRS, et de l'INRA.

Cette seconde édition intègre de manière plus détaillée l'économétrie des données de panel via un chapitre méthodologique spécifique (chapitre 12), deux applications sur données de panel sont proposées : la première dans un cadre statique, la seconde dans un cadre dynamique (chapitre 13 et chapitre 14).

1. Objectifs du livre

L'objectif de ce manuel est d'initier et d'accompagner le lecteur à la pratique de l'économétrie en lui proposant des cas d'applications concrètes fondées sur la théorie économique. L'ouvrage fait le lien direct entre l'économie et l'économétrie. Il réalise un rappel de cours sur les méthodes économétriques de base et avancées et propose des applications originales accompagnées de corrigés complets et détaillés. Les applications portent sur des thèmes économiques diversifiés et mettent en jeu des données statistiques réelles de natures différentes (série chronologique, coupe instantanée, panel).

Le lecteur peut réaliser lui-même toutes les applications développées dans le manuel à partir des données fournies sous format Excel, ASCII et TXT. Il peut également utiliser les programmes de traitement développés par les auteurs notamment

avec les logiciels économétriques WinRats, SAS, Eviews. Ces fichiers répertoriés selon le problème économétrique et économique traité, sont disponibles par téléchargement sur le serveur web :

<http://perso.univ-rennes1.fr/isabelle.cadoret-david/>

Les auteurs proposent ainsi de suivre pas à pas le lecteur dans son apprentissage. Ils lui montrent comment (i) construire un modèle économétrique pour répondre à une question réelle, (ii) choisir la méthode économétrique appropriée, (iii) créer une base de données, (iv) analyser et interpréter ses résultats.

2. À qui s'adresse l'ouvrage

Ce livre s'adresse aux étudiants du deuxième et troisième cycles, universitaires et non universitaires, de sciences économiques et de gestion, d'économétrie et d'économie quantitative mais également aux chercheurs et praticiens de l'économétrie dans les entreprises et les organismes publics. Les enseignants pourront y retrouver un rappel de cours et des exercices d'application en lien direct avec les cours d'économie.

3. Contenu et plan de l'ouvrage

L'ouvrage est organisé par ordre croissant de difficultés méthodologiques. Les chapitres de méthodes sont systématiquement suivis de un ou deux chapitres d'applications économiques :

1. Le chapitre 1 présente le modèle de régression linéaire simple. Il est suivi d'un chapitre traitant de la question de l'évaluation du degré de mobilité des capitaux dans les pays de l'OCDE. Il s'agit d'expliquer et d'utiliser les concepts de bases de l'économétrie (aléas, résidus, variables endogènes, variables exogènes, droite de régression, estimateur des Moindres Carrés Ordinaires, ...).
2. Le chapitre 3 élargit le modèle de régression linéaire simple au modèle de régression linéaire multiple. Deux chapitres d'applications accompagnent ce rappel de cours méthodologique : un chapitre sur la croissance et la convergence des économies européennes et un chapitre sur la politique monétaire européenne. L'accent est mis sur l'estimation des modèles avec la méthode des Moindres Carrés Ordinaires et sur les tests statistiques usuels.
3. Le chapitre 6 discute les hypothèses du modèle de régression linéaire multiple et présente les méthodes d'estimations des Moindres Carrés Généralisés et des Variables Instrumentales. Le chapitre 7 propose une application traitant les problèmes d'autocorrélation des aléas à partir de l'estimation du modèle d'évaluation des actifs financiers (MEDAF) sur des données boursières des sociétés de CAC 40. Le chapitre 8 aborde la question de l'endogénéité des variables dans un modèle dynamique portant sur la fonction d'offre de viande bovine française.
4. Le chapitre 9 étend le modèle de régression linéaire multiple vers les modèles à plusieurs équations apparemment non reliées et les modèles à équations

simultanées. Les méthodes d'estimation des modèles équations simultanées sont présentées : les méthodes des Doubles Moindres Carrés Ordinaires et des Triples Moindres Carrés Ordinaires. Les deux chapitres suivants constituent des applications directes. Le chapitre 10 présente une illustration basée sur l'estimation d'un système de demande de facteurs de production avec la méthode de Zellner. Le chapitre 11 présente un modèle d'équilibre partiel où les paramètres sont estimés avec des méthodes pour équations simultanées.

5. Le chapitre 12 présente les notions de base en économétrie des données de panels : les modèles à effets fixes et à effets aléatoires ainsi que les méthodes d'estimation associées. Deux applications sur données de panel sont proposées : la première dans un cadre statique, la seconde dans un cadre dynamique. Le chapitre 13 développe une application basée sur l'estimation d'une équation statique de salaire à partir de données sectorielles. Le chapitre 14 porte sur la recherche des déterminants de l'investissement des entreprises en examinant le lien entre investissement et variables financières à partir de données individuelles d'entreprises.
6. Le chapitre 15 aborde le vaste thème de l'économétrie des séries non stationnaires. Il passe en revue les tests standards de racine unité, les tests de cointégration et les modèles à correction d'erreurs. Le chapitre 16 conduit à pratiquer sur des séries financières les tests usuels de stationnarité (Dickey-Fuller) et de cointégration et à estimer un modèle à correction d'erreur dans un cadre univarié (Engel et Granger) afin de construire un indicateur de valorisation des marchés boursiers. Le chapitre 17 évalue les modes de formation des salaires et des prix en France dans un cadre de concurrence imparfaite, à partir d'une spécification des marchés du travail sous la forme d'un modèle appelé WS-PS. Les tests et l'estimation des relations de cointégration sont pratiqués avec la méthode de Johansen.
7. Le chapitre 17 introduit l'économétrie des variables qualitatives : les modèles de choix binaire et les modèles à variables dépendantes limitées. Pour terminer, le chapitre 18 examine les décisions d'offre de travail. L'accent est mis sur les techniques d'analyse des comportements dichotomiques (ici l'exercice ou non d'une activité professionnelle) à partir de modèles probit ou logit et sur l'analyse de variables censurées (modèle tobit).

À travers la lecture et la résolution des exercices d'applications, les auteurs souhaitent montrer l'intérêt et le vaste champ d'application des techniques économétriques. Ils cherchent à expliquer comment, à partir d'une question, un modèle économétrique est construit et évalué. L'économètre doit savoir utiliser les outils économétriques mais il doit aussi savoir interpréter ses résultats, revenir à la question économique posée et apporter des éléments de réponse.

4. Utilisation du manuel et niveaux de difficulté

Les chapitres de l'ouvrage sont classés selon un ordre sensiblement croissant de difficulté économétrique. Le niveau de difficulté économique des chapitres d'applications ne suit pas rigoureusement la même progression. Néanmoins on peut considérer trois niveaux d'entrée distincts pour l'utilisation du manuel.

- Les chapitres 1 à 8 constituent le niveau d'entrée des étudiants de troisième année de licence.
- Les chapitres 9 à 11 correspondent aux programmes traditionnels de Master première année.
- Enfin les chapitres 12 à 19, plus avancés à la fois du point de vue des méthodes économétriques et de la théorie économique mobilisée, intéresseront davantage les étudiants de dernière année de Master et un public plus large de jeunes chercheurs ou praticiens de l'économétrie.

5. Remerciements

Pour conclure cet avant-propos, les auteurs tiennent à remercier toutes les personnes qui ont participé au travail important de relecture du manuel et restent naturellement responsables des erreurs éventuelles qui subsisteraient. Les étudiants de la Faculté de Sciences Economiques de Rennes 1 ainsi que ceux de l'ENSAI ont eux aussi par leurs remarques et suggestions, à l'occasion des cours, grandement contribué à l'amélioration de cet ouvrage. Qu'ils soient également remerciés !

Nous tenons par ailleurs à remercier D. Delaunay pour avoir développé les programmes sur le logiciel SAS.

1

LE MODÈLE DE RÉGRESSION LINÉAIRE SIMPLE

1. Présentation du modèle de régression linéaire simple
2. Estimation avec la méthode des Moindres Carrés Ordinaires
3. Intervalles de confiance et test de significativité des paramètres
4. Réalisation de prédictions et intervalles de confiance des prédictions

OBJECTIFS DU CHAPITRE

Présenter des concepts de base à partir d'un modèle économétrique simple

- Variable endogène ou expliquée
- Variable exogène ou explicative
- Aléas
- Estimateur des Moindres Carrés Ordinaires
- Droite de régression
- Analyse de la variance
- Coefficient de détermination
- Test de significativité des paramètres
- Prédictions

NIVEAU

LICENCE, deuxième et troisième années

1. PRÉSENTATION DU MODÈLE DE RÉGRESSION LINÉAIRE SIMPLE

Le modèle de régression linéaire simple permet d'expliquer une variable endogène en fonction d'une variable explicative

$$y_i = \beta_1 + \beta_2 x_i + \varepsilon_i \quad i = 1, \dots, N \quad (1.1)$$

avec

y : la variable endogène (la variable expliquée)

x : la variable explicative supposée exogène

ε : un aléa

N : le nombre d'observations (ces observations sont des données individuelles ou temporelles, on parlera ici d'individus)

Les variables y et x sont des variables observées pour un nombre N d'individus. La variable y est une variable aléatoire dont les observations sont issues d'un tirage aléatoire d'un échantillon dans une population. On suppose, pour simplifier, que la variable x est non aléatoire. Pour alléger les écritures, les notations ne distinguent pas les variables aléatoires et leurs réalisations. Par exemple, la notation y désigne la variable aléatoire Y et sa réalisation y .

L'aléa ε est non observé. Il s'agit d'un terme d'erreur stochastique qui permet de prendre en compte le fait que la variable y est affectée par d'autres variables que la variable x , autrement dit le fait que x n'explique pas pleinement la variable y .

Le modèle est supposé linéaire par rapport aux paramètres β_1 et β_2 .

Exemple 1.1

Considérons l'analyse, pour l'année 2003, de la dépense de consommation des étudiants en communications téléphoniques en fonction du revenu des parents. On construit un modèle économétrique dans lequel la variable endogène est la dépense de consommation en communications téléphoniques et la variable exogène est le revenu des parents. On note que des étudiants dont les parents ont le même revenu ne dépendent pas un montant identique en communications téléphoniques, d'autres facteurs expliquent le comportement de l'étudiant (étendue du réseau d'amis, goût pour le bavardage, etc.). Dans le modèle économétrique, ces facteurs sont pris en compte par l'aléa.

Dans la pratique, on ne dispose pas des observations sur toute la population étudiante mais il est possible de constituer un échantillon de cette population afin d'évaluer la relation entre les deux variables. Pour une valeur fixée de la variable revenu des parents, on observe dans l'échantillon une dépense de consommation (dans un autre échantillon ce niveau de dépense serait différent).

À partir des données observées (les données de l'échantillon) il est possible d'estimer la relation (1.1). Soient $\widehat{\beta}_1$ et $\widehat{\beta}_2$ les valeurs estimées des paramètres β_1 et β_2 , la droite de régression de l'échantillon est donnée par

$$\widehat{y}_i = \widehat{\beta}_1 + \widehat{\beta}_2 x_i \quad i = 1, \dots, N \quad (1.2)$$

et e_i le résidu d'estimation, $e_i = y_i - \widehat{y}_i$, mesure pour chaque individu l'écart entre la valeur observée y_i et la valeur estimée \widehat{y}_i . La valeur estimée est aussi appelée la valeur ajustée ou prédiction.

La droite de régression de la population pourrait être évaluée si l'on disposait des observations sur toute la population, elle correspond à la valeur espérée de la variable aléatoire y pour une valeur fixée de x

$$E(y_i/x_i) = \beta_1 + \beta_2 x_i$$

on suppose que $E(\varepsilon_i/x_i) = 0$.

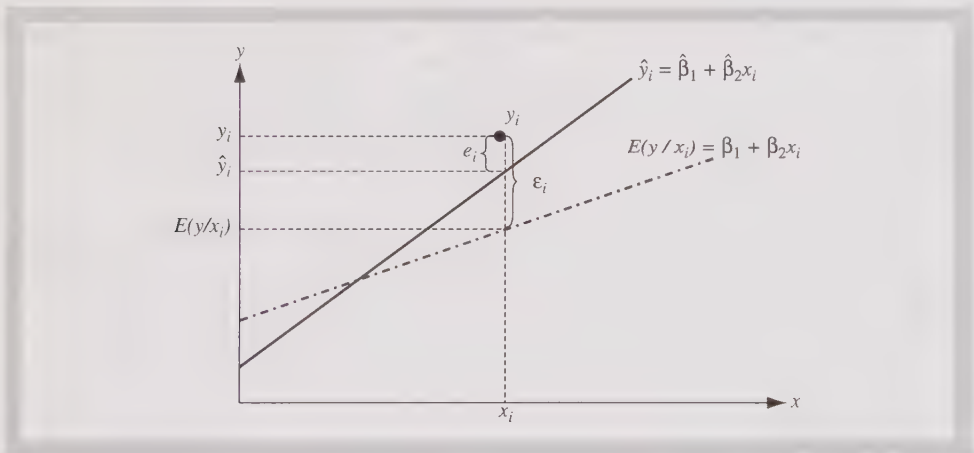


Figure 1.1 Droites de régression de la population et de l'échantillon

Dans la pratique, les données de toute la population ne sont jamais observées et donc on ne connaît pas les valeurs des paramètres de la population β_1 et β_2 . L'estimation économétrique vise à évaluer les paramètres de la population. On utilise pour cela les données d'un échantillon et, de ce fait, les résultats d'estimation sont liés à l'échantillon retenu.

L'objectif est de trouver des estimateurs non biaisés, efficaces et convergents des paramètres. Un estimateur non biaisé est un estimateur dont l'espérance est égale à la valeur des paramètres de la population, il est efficace si, parmi les estimateurs non biaisés des paramètres, il a la variance la plus faible. Enfin, il est convergent en probabilité si, lorsqu'on accroît la taille de l'échantillon, la probabilité que les valeurs estimées des paramètres soient différentes de β_1 et β_2 est nulle.

2. ESTIMATION AVEC LA MÉTHODE DES MOINDRES CARRÉS ORDINAIRES

La méthode d'estimation de base des paramètres du modèle (1.1) est la méthode des Moindres Carrés Ordinaires (MCO). Dans cette section sont présentés les hypothèses du modèle, l'estimateur des MCO, les caractéristiques et les propriétés de cet estimateur.

2.1 *Hypothèses sur le modèle de régression linéaire simple*

Les hypothèses de base sont les suivantes

Hypothèse 1

L'espérance des aléas est nulle

$$E(\varepsilon_i) = 0 \quad i = 1, \dots, N$$

De ce fait

$$E(y_i/x_i) = \beta_1 + \beta_2 x_i$$

Hypothèse 2

Les aléas sont homoscédastiques (leur variance est constante) et non auto-corrélés (leur covariance est nulle)

$$\begin{aligned} \text{var}(\varepsilon_i) &= E(\varepsilon_i - E(\varepsilon_i))^2 = E(\varepsilon_i)^2 = \sigma^2 & i = 1, \dots, N \\ \text{cov}(\varepsilon_i, \varepsilon_j) &= E(\varepsilon_i - E(\varepsilon_i))(\varepsilon_j - E(\varepsilon_j)) \\ &= E(\varepsilon_i \varepsilon_j) = 0 & i, j = 1, \dots, N, \quad i \neq j \end{aligned}$$

Pour une valeur de x fixée, cette hypothèse implique que

$$\text{var}(y_i) = E(y_i - E(y_i))^2 = E(\varepsilon_i)^2 = \sigma^2 \quad i = 1, \dots, N$$

et

$$\begin{aligned} \text{cov}(y_i, y_j) &= E(y_i - E(y_i))(y_j - E(y_j)) \\ &= E(\varepsilon_i \varepsilon_j) = 0 & i, j = 1, \dots, N, \quad i \neq j \end{aligned}$$

Si les hypothèses 1 et 2 sont vérifiées, l'aléa ε est un bruit blanc.

Hypothèse 3

La variable x est supposée certaine. Dans ce cas

$$E(y_i/x_i) = E(y_i) \quad i = 1, \dots, N$$

Cette hypothèse forte n'est pas nécessaire elle permet dans un premier temps de simplifier la présentation.

Hypothèse 4

La variance empirique de la variable explicative est non nulle et donc les observations x_i ne sont pas toutes identiques

$$S_X = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \neq 0 \quad \text{avec} \quad \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

et lorsque la taille de l'échantillon s'accroît, la variable explicative x_i conserve toujours une certaine variance et toute nouvelle observation apporte de l'information

$$\lim_{N \rightarrow +\infty} S_X = \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 = \sigma_x^2 \neq 0$$

2.2 Estimateur des Moindres Carrés Ordinaires (MCO)

L'estimateur des MCO est dérivé de la minimisation de la somme des carrés des aléas

$$\underset{\beta_1, \beta_2}{\text{Min}} \sum_{i=1}^N \varepsilon_i^2 = \underset{\beta_1, \beta_2}{\text{Min}} \sum_{i=1}^N (y_i - \beta_1 - \beta_2 x_i)^2 = \underset{\beta_1, \beta_2}{\text{Min}} S \quad (1.3)$$

D'après les conditions du premier ordre, $\frac{\partial S}{\partial \beta_1} = 0$ et $\frac{\partial S}{\partial \beta_2} = 0$, la solution est donnée par

$$-2 \sum_{i=1}^N (y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i) = -2 \sum_{i=1}^N e_i = 0 \quad (1.4)$$

$$-2 \sum_{i=1}^N (y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i) x_i = -2 \sum_{i=1}^N e_i x_i = 0 \quad (1.5)$$

La résolution des équations (1.4) et (1.5) donne

$$\hat{\beta}_2 = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad \text{et} \quad \hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \bar{x} \quad (1.6)$$

La droite de régression de l'échantillon est alors donnée par

$$\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_i \quad i = 1, \dots, N \tag{1.7}$$

et les caractéristiques de la solution sont les suivantes

- la droite de régression de l'échantillon passe par le point moyen (\bar{y}, \bar{x})

$$\bar{y} = \hat{\beta}_1 + \hat{\beta}_2 \bar{x}$$

- la somme des résidus est nulle, d'après l'équation (1.4)

$$\sum_{i=1}^N e_i = 0$$

par conséquent $\bar{y} = \bar{\hat{y}}$

- le vecteur des résidus et de la variable explicative sont orthogonaux, d'après l'équation (1.5)

$$\sum_{i=1}^N e_i x_i = 0$$

De ces caractéristiques, il est possible de déduire l'équation d'analyse de la variance

$$\boxed{\sum_{i=1}^N (y_i - \bar{y})^2 = \sum_{i=1}^N (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^N e_i^2} \tag{1.8}$$

Au facteur $(1/N)$ près, la variance totale de la variable endogène est la somme de la variance expliquée par le modèle et la variance résiduelle. Le coefficient de détermination R^2 qui mesure le rapport entre la variance expliquée par le modèle et la variance totale est un indicateur de la qualité de l'ajustement. Il donne le pourcentage de la variance totale de y expliqué par le modèle de régression

$$\boxed{R^2 = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^N (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^N e_i^2}{\sum_{i=1}^N (y_i - \bar{y})^2}} \tag{1.9}$$

dans le modèle de régression simple le coefficient de détermination est égal au coefficient de corrélation entre les variables y et x

$$R^2 = r^2 = \frac{\left(\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) \right)^2}{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2} \quad (1.10)$$

Dans le modèle de régression linéaire avec constante, le coefficient de détermination est compris entre 0 et 1, une valeur proche de 1 traduit une bonne qualité de l'ajustement. Le rôle de la constante est important car la somme des résidus est nulle seulement lorsque le modèle comporte une constante. L'équation d'analyse de la variance n'est vérifiée que dans un modèle avec constante.

2.3 Propriétés de l'estimateur des MCO

Théorème de Gauss-Markov

Sous les hypothèses 1, 2, 3 et 4, l'estimateur des MCO est Best Linear Unbiased Estimator (BLUE).

Autrement dit, l'estimateur des MCO est

- une fonction linéaire de y ,
- non biaisé : $E(\hat{\beta}_1) = \beta_1$ et $E(\hat{\beta}_2) = \beta_2$,
- efficace : parmi les estimateurs linéaires non biaisés des paramètres β_1 et β_2 du modèle, l'estimateur des MCO a la variance la plus faible. La matrice de variance-covariance des paramètres est donnée par

$$\begin{pmatrix} \text{var}(\hat{\beta}_1) & \text{cov}(\hat{\beta}_1, \hat{\beta}_2) \\ \text{cov}(\hat{\beta}_1, \hat{\beta}_2) & \text{var}(\hat{\beta}_2) \end{pmatrix}$$

avec

$$\begin{aligned} \text{var}(\hat{\beta}_1) &= \sigma_{\hat{\beta}_1}^2 = E(\hat{\beta}_1 - E(\hat{\beta}_1))^2 = E(\hat{\beta}_1 - \beta_1)^2 \\ &= \frac{\sigma^2}{N} \frac{\sum_{i=1}^N x_i^2}{\sum_{i=1}^N (x_i - \bar{x})^2} \end{aligned} \quad (1.11)$$

$$\begin{aligned} \text{var}(\hat{\beta}_2) &= \sigma_{\hat{\beta}_2}^2 = E(\hat{\beta}_2 - E(\hat{\beta}_2))^2 = E(\hat{\beta}_2 - \beta_2)^2 \\ &= \frac{\sigma^2}{N \sum_{i=1}^N (x_i - \bar{x})^2} \end{aligned} \quad (1.12)$$

$$\begin{aligned} \text{cov}(\widehat{\beta}_1, \widehat{\beta}_2) &= E(\widehat{\beta}_1 - E(\widehat{\beta}_1))(\widehat{\beta}_2 - E(\widehat{\beta}_2)) = E(\widehat{\beta}_1 - \beta_1)(\widehat{\beta}_2 - \beta_2) \\ &= -\bar{x} \frac{\sigma^2}{\sum_{i=1}^N (x_i - \bar{x})^2} \end{aligned} \quad (1.13)$$

La matrice de variance-covariance des paramètres ne peut pas être calculée car la valeur de σ^2 n'est pas connue. Pour déterminer cette matrice de variance-covariance, il faut un estimateur $\widehat{\sigma}^2$ non biaisé de la variance des aléas. On montre que

$$E\left(\sum_{i=1}^N e_i^2\right) = (N-2)\sigma^2 \quad (1.14)$$

d'où

$$\widehat{\sigma}^2 = \frac{1}{N-2} \sum_{i=1}^N e_i^2 \quad (1.15)$$

est un estimateur non biaisé de σ^2

$$E(\widehat{\sigma}^2) = \frac{1}{N-2} E\left(\sum_{i=1}^N e_i^2\right) = \sigma^2 \quad (1.16)$$

On peut alors en déduire une valeur estimée de la matrice de variance-covariance des paramètres

$$\begin{aligned} \widehat{\sigma}_{\widehat{\beta}_1}^2 &= \frac{\widehat{\sigma}^2}{N} \frac{\sum_{i=1}^N x_i^2}{\sum_{i=1}^N (x_i - \bar{x})^2}, & \widehat{\sigma}_{\widehat{\beta}_2}^2 &= \frac{\widehat{\sigma}^2}{\sum_{i=1}^N (x_i - \bar{x})^2}, \\ \text{cov}(\widehat{\beta}_1, \widehat{\beta}_2) &= -\bar{x} \frac{\widehat{\sigma}^2}{\sum_{i=1}^N (x_i - \bar{x})^2} \end{aligned} \quad (1.17)$$

3. INTERVALLES DE CONFIANCE ET TEST DE SIGNIFICATIVITÉ DES PARAMÈTRES

Afin de déterminer les intervalles de confiance des paramètres du modèle, il faut déterminer la distribution statistique de $\widehat{\beta}_1$ et $\widehat{\beta}_2$.

Les principales distributions statistiques utilisées en économétrie appliquée sont la distribution Normale (N), la distribution du Chi-deux (χ^2), la distribution de Fisher-Snédecor (F) et la distribution de Student (τ).

Soient z_1, z_2, \dots, z_r , une suite de variables normales centrées réduites indépendantes, la variable

$$\chi^2 = \sum_{i=1}^r z_i^2$$

est distribuée suivant une loi du Chi-deux à r degrés de liberté ($\chi^2(r)$). Si ces variables ne sont pas indépendantes mais liées entre elles par p ($p < r$), relations linéaires distinctes, alors la variable χ^2 est distribuée suivant une loi à $r - p$ degrés de liberté ($\chi^2(r - p)$).

Soient z_1 et z_2 2 variables aléatoires indépendantes distribuées selon une loi du χ^2 respectivement à r et m degrés de liberté,

$$F(r, m) = \frac{\chi^2(r)/r}{\chi^2(m)/m}$$

est distribuée suivant un loi de Fisher Snédecor à r et m degrés de liberté.

Soient z_1 et z_2 2 variables aléatoires indépendantes respectivement distribuées selon une loi normale centrée réduite et un $\chi^2(r)$

$$\tau(r) = \frac{z_1}{\sqrt{\chi^2(r)/r}}$$

est distribué selon une loi de Student à r degrés de liberté.

Pour déterminer la distribution statistique puis les intervalles de confiance des paramètres on pose une hypothèse supplémentaire

Hypothèse 5

Les aléas sont distribués de manière identique et indépendante selon une loi normale

$$\varepsilon_i \sim N(0, \sigma^2) \quad i = 1, \dots, N \quad (1.18)$$

la variable endogène y , étant une fonction linéaire des aléas, est aussi distribuée selon une loi normale

$$y_i \sim N(\beta_1 + \beta_2 x_i, \sigma^2) \quad i = 1, \dots, N \quad (1.19)$$

les y_i sont distribués de manière indépendante mais ils ne sont pas distribués de manière identique car ils n'ont pas la même espérance.

Les estimations des paramètres dépendent de l'échantillon de données, ainsi, les estimateurs $\hat{\beta}_1$ et $\hat{\beta}_2$ sont des variables aléatoires. On en déduit que

$$\hat{\beta}_1 \sim N(\beta_1, \sigma_{\hat{\beta}_1}^2), \quad \hat{\beta}_2 \sim N(\beta_2, \sigma_{\hat{\beta}_2}^2) \tag{1.20}$$

d'où

$$\frac{\hat{\beta}_1 - \beta_1}{\sigma_{\hat{\beta}_1}} \sim N(0,1), \quad \frac{\hat{\beta}_2 - \beta_2}{\sigma_{\hat{\beta}_2}} \sim N(0,1) \tag{1.21}$$

On ne peut pas en déduire directement les intervalles de confiance des paramètres car on ne peut pas calculer les écart type $\sigma_{\hat{\beta}_1}$ et $\sigma_{\hat{\beta}_2}$. On utilise pour cela $\hat{\sigma}^2$ l'estimateur non biaisé de σ^2 dont la distribution est un χ^2 à $N - 2$ degrés de liberté

$$\frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2(N - 2) \tag{1.22}$$

Ainsi on obtient

$$\frac{\frac{\hat{\beta}_k - \beta_k}{\sigma_{\hat{\beta}_k}}}{\frac{\hat{\sigma}}{\sigma}} = \frac{\hat{\beta}_k - \beta_k}{\hat{\sigma}_{\hat{\beta}_k}} \sim t(N - 2) \quad k = 1,2 \tag{1.23}$$

et les intervalles de confiance des paramètres

$$\Pr \left[\hat{\beta}_k - \hat{\sigma}_{\hat{\beta}_k} \tau_{\alpha/2}(N - 2) \leq \beta_k \leq \hat{\beta}_k + \hat{\sigma}_{\hat{\beta}_k} \tau_{\alpha/2}(N - 2) \right] = 1 - \alpha \quad k = 1,2 \tag{1.24}$$

avec α le risque de première espèce généralement fixé à 5 % ou 10 %.

À partir de (1.23) on peut réaliser le test de significativité des paramètres

$$\begin{aligned} H_0 : \beta_k &= 0 \\ H_1 : \beta_k &\neq 0 \end{aligned} \quad k = 1,2 \tag{1.25}$$

on calcule la statistique

$$t_{cal} = \frac{\hat{\beta}_k}{\hat{\sigma}_{\hat{\beta}_k}} \tag{1.26}$$

distribuée selon une distribution de Student à $N - 2$ degrés de liberté.

Si $|t_{cal}| > \tau_{\alpha/2}(N - 2)$ alors l'hypothèse H_0 est refusée au seuil α et le coefficient β_k peut être considéré significativement différent de 0. Ce test est un

test de base qui a pour but de vérifier que la variable explicative x influence la variable y . Le non rejet de l'hypothèse nulle $H_0 : \beta_2 = 0$ signifie que la variable x n'explique pas la variable y . Une manière alternative de réaliser ce test dans un modèle de régression simple est d'utiliser le coefficient de détermination, c'est-à-dire le coefficient de corrélation entre x et y . En effet on peut montrer que

$$\begin{aligned} t_{cal}^2 = F &= \left[\frac{\widehat{\beta}_2}{\widehat{\sigma}_{\beta_2}} \right]^2 \\ &= \frac{R^2/1}{(1 - R^2)/N - 2} \sim > F(1, N - 2) \end{aligned} \quad (1.27)$$

Si la statistique de Fisher F est supérieure à la valeur tabulée $F_\alpha(1, N - 2)$ alors l'hypothèse $H_0 : \beta_2 = 0$ est rejetée au seuil α .

4. RÉALISATION DE PRÉDICTIONS ET INTERVALLE DE CONFIANCE DES PRÉDICTIONS

À partir de l'équation de la droite de régression de l'échantillon, il est possible de fournir des prédictions. Pour une valeur de la variable x donnée, notée x_0 , on peut calculer la valeur estimée \widehat{y}_0

$$\widehat{y}_0 = \widehat{\beta}_1 + \widehat{\beta}_2 x_0 \quad (1.28)$$

Il ne s'agit pas ici forcément de prévision du futur. On détermine la valeur prédite de y par le modèle lorsque la variable explicative exogène x se modifie. La prédiction consiste à calculer la valeur estimée de y à l'intérieur ou en dehors de l'échantillon des observations.

La valeur de y , lorsque $x = x_0$, serait $y_0 = \beta_1 + \beta_2 x_0 + \varepsilon_0$. D'après le théorème de Gauss-Markov \widehat{y}_0 est un estimateur BLUE de $E(y_0/x_0)$.

L'erreur de prévision se définit comme l'écart entre y_0 et \widehat{y}_0 .

$$e_0 = y_0 - \widehat{y}_0 = (\beta_1 - \widehat{\beta}_1) + (\beta_2 - \widehat{\beta}_2) x_0 + \varepsilon_0 \quad (1.29)$$

En utilisant les hypothèses 1, 2, 3 on montre que l'espérance et la variance de l'erreur de prévision sont données par

$$E(e_0) = \left(\beta_1 - \underbrace{E(\widehat{\beta}_1)}_{\beta_1} \right) + \left(\beta_2 - \underbrace{E(\widehat{\beta}_2)}_{\beta_2} \right) x_0 + E(\varepsilon_0) = 0 \quad (1.30)$$

$$\begin{aligned} \text{var}(e_0) &= \text{var}(\hat{\beta}_1) + x_0^2 \text{var}(\hat{\beta}_2) + 2x_0 \text{cov}(\hat{\beta}_1, \hat{\beta}_2) + \text{var}(\varepsilon_0) \\ &= \sigma^2 \left[\frac{1}{N} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^N (x_i - \bar{x})^2} + 1 \right] = \sigma_{e_0}^2 \end{aligned} \quad (1.31)$$

L'espérance étant nulle, le prédicteur \hat{y}_0 est sans biais. L'équation (1.31) mesure la variance de l'erreur de prévision et donc l'imprécision de la prédiction. Elle est d'autant plus faible que la valeur de N est élevée et que la variance des paramètres estimés est faible. Le terme σ^2 étant constant, on ne peut jamais réaliser de prévisions parfaites.

La distribution de l'erreur de prévision est une loi normale sous l'hypothèse 5

$$\frac{e_0}{\sigma_{e_0}} = \frac{y_0 - \hat{y}_0}{\sigma_{e_0}} \sim N(0,1) \quad (1.32)$$

La variance de l'erreur de prévision peut être estimée en utilisant $\hat{\sigma}^2$

$$\hat{\sigma}_{e_0}^2 = \hat{\sigma}^2 \left[\frac{1}{N} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^N (x_i - \bar{x})^2} + 1 \right] \quad (1.33)$$

d'où

$$\frac{y_0 - \hat{y}_0}{\hat{\sigma}_{e_0}} \sim \tau(N-2) \quad (1.34)$$

l'intervalle de confiance de y_0 s'écrit alors

$$\Pr \left[\hat{y}_0 - \hat{\sigma}_{e_0} \tau_{\alpha/2}(N-2) \leq y_0 \leq \hat{y}_0 + \hat{\sigma}_{e_0} \tau_{\alpha/2}(N-2) \right] = 1 - \alpha \quad (1.35)$$

2

ÉPARGNE, INVESTISSEMENT ET MOBILITÉ DU CAPITAL

1. Introduction
2. Cadre d'analyse
3. Évaluation de la relation épargne-investissement
4. Éléments de corrigé

OBJECTIFS DU CHAPITRE

D'un point de vue méthodologique

- Estimer un modèle de régression linéaire simple
- Pratiquer des tests statistiques
- Élargir le modèle de régression linéaire simple

D'un point de vue économique

- Évaluer le degré de mobilité des capitaux

NIVEAU

LICENCE, deuxième et troisième années

CHAMPS ÉCONOMIQUES TRAITÉS

Macroéconomie monétaire et financière

1. INTRODUCTION

La construction de l'Europe monétaire s'est accompagnée d'un mouvement vers l'abandon total du contrôle des capitaux. Le contrôle des mouvements de capitaux permet à un pays donné de conserver l'épargne nationale afin de financer son investissement et de créer des emplois. Il permet également d'éviter toute évasion fiscale. Par ailleurs, en régime de change fixe, il permet de conserver une relative indépendance monétaire. Mais, en contrepartie, il conduit à une segmentation des marchés financiers intérieurs et extérieurs et invalide le principe de parité des taux d'intérêt non couverte, c'est-à-dire l'égalisation internationale des taux d'intérêt tenant compte de la variation anticipée des taux de change.

Depuis la fin des années 80, on observe une libéralisation accrue du marché international des capitaux et cette libéralisation a abouti à la globalisation financière, c'est-à-dire à un marché planétaire des capitaux. Les marchés internationaux des capitaux sont le théâtre des opérations de prêts et d'emprunts dans les différentes monnaies et les prix qui se forment sur ces marchés sont les taux d'intérêt. Les principaux acteurs sont : les banques, les entreprises multinationales mais aussi les États. Ces marchés comportent trois pôles importants : les prêts bancaires internationaux, les obligations internationales et les instruments du marché monétaire.

La libéralisation du marché international des capitaux s'est matérialisée par une croissance importante, en volume et en volatilité, des transactions financières internationales et des flux de capitaux et a induit des crises financières très coûteuses. Les effets positifs attendus de cette libéralisation sont une hausse de l'investissement, une croissance plus rapide et une augmentation du niveau de vie dans beaucoup de pays.

Dans ce contexte de libéralisation du marché des capitaux, Feldstein et Horioka (1980)¹ ont cherché à évaluer le degré de mobilité des capitaux internationaux à long terme sur la base d'une analyse du lien entre l'investissement et l'épargne des pays de l'OCDE. Cette question est importante car, selon le degré de mobilité du capital, l'impact des politiques économiques peut être faible ou fort. Par exemple, une politique fiscale en faveur de l'épargne nationale aura un effet d'autant plus important sur l'investissement du pays que le degré de mobilité du capital est faible. En effet, en cas d'absence de mobilité des capitaux, on retrouve les résultats de l'économie fermée, l'épargne nationale est égale à l'investissement national et tout accroissement de l'épargne induit un accroissement identique de l'investissement. En revanche, si le degré de mobilité du capital est élevé, cette même politique en faveur de l'épargne aura moins d'impact sur l'investissement du pays.

Feldstein et Horioka ont donc proposé d'analyser la corrélation entre le taux d'épargne et le taux d'investissement afin de déterminer si les économies ont un comportement proche de celui d'une économie fermée ou, au contraire, si les marchés des capitaux peuvent être considérés comme parfaitement intégrés. Ce chapitre reprend leur analyse. Les première et seconde sections présentent respectivement la spécification retenue par Feldstein et Horioka et leurs résultats. Les troisième et

¹ Feldstein M., Horioka C., 1980, « Domestic saving and international capital flows ». *Economic Journal*, 90, 314-329.

quatrième sections présentent une correction succincte des problèmes posés par leur analyse.

2. CADRE D'ANALYSE

2.1 *Les relations macroéconomiques en économie ouverte*

Le cadre retenu est un modèle macroéconomique d'économie ouverte. Il repose sur les identités comptables

$$Y = C + I + G + X - M \quad (2.1)$$

$$S_{privée} = Y - T - C \quad (2.2)$$

$$S_{publique} = T - G \quad (2.3)$$

L'équation (2.1) représente l'identité comptable de répartition du produit national brut (Y) entre les dépenses de consommation (C), les dépenses d'investissement (I), les dépenses publiques (G) et le solde de la balance commerciale donné par l'écart entre les exportations vers le reste du monde (X) et les importations en provenance de l'étranger (M). L'équation (2.2) indique que l'épargne privée est donnée par l'écart entre le produit national brut net des taxes (T) et les dépenses de consommation. Enfin, l'équation (2.3) définit l'épargne publique comme l'écart entre les recettes et les dépenses courantes de l'État. On retrouve l'égalité traditionnelle selon laquelle l'investissement national est égal à l'épargne nationale moins le compte courant

$$I = (Y - C - T) + (T - G) - X + M \quad (2.4)$$

$$I = S_{privée} + S_{publique} - (X - M) \quad (2.5)$$

Par la suite, on note l'épargne nationale $S_{nat} = S_{privée} + S_{publique}$. En économie fermée $(X - M) = 0$ et l'investissement est égal à l'épargne nationale. La corrélation entre l'épargne et l'investissement est alors égale à 1, ce qui signifie qu'une variation de l'épargne induit une variation équivalente de l'investissement.

En revanche, en économie ouverte avec une parfaite mobilité du capital, il n'y a théoriquement plus de corrélation entre l'épargne nationale et l'investissement national. D'une part, l'épargne nationale répond aux opportunités d'investissement du marché mondial et, d'autre part, l'investissement national est financé par l'épargne mondiale. Si l'investissement est déterminé par les taux d'intérêt réels, selon le principe de décroissance des productivités marginales on devrait, jusqu'à ce que la productivité marginale du capital soit égalisée partout dans le monde, observer d'importants flux de capitaux des pays industrialisés vers les pays en développement. En effet, théoriquement, les pays en développement moins biens dotés en capital,

out des productivités marginales du capital plus élevées que les pays industrialisés et donc des taux d'intérêt réels également plus importants². Ce schéma théorique est connu sous le nom de condition de parité des taux d'intérêt selon laquelle les rendements du capital ne peuvent pas être systématiquement différents d'un pays à l'autre s'il n'existe aucun obstacle au mouvement des capitaux. En revanche, en cas d'imparfaite mobilité du capital, des écarts entre les taux d'intérêt pratiqués dans les différents pays subsistent.

2.2 Spécification du modèle

Feldstein et Horioka ont proposé de mesurer le degré de mobilité du capital sur la base d'une analyse de la corrélation entre l'épargne et l'investissement. Ils ont estimé l'équation

$$\frac{I}{Y} = \beta_1 + \beta_2 \frac{S}{Y} + \varepsilon \quad (2.6)$$

où $\frac{I}{Y}$ et $\frac{S}{Y}$ représentent respectivement la part de l'investissement brut dans le produit national et la part de l'épargne nationale dans le produit national. ε est un aléa.

En cas de parfaite mobilité des capitaux, un accroissement du taux d'épargne dans un pays doit induire un accroissement de l'investissement. La productivité marginale du capital étant une fonction décroissante du stock de capital, l'accroissement de l'investissement bénéficiera d'abord aux pays initialement les moins dotés en capital. Par conséquent, en cas de forte mobilité du capital, un accroissement de l'épargne nationale n'induit pas un accroissement identique de l'investissement national. Une estimation de β_2 proche de zéro conduit donc à considérer qu'il y a une forte mobilité des capitaux dans la mesure où cela signifie que la corrélation entre l'épargne nationale et l'investissement national est nulle. En revanche, une estimation de β_2 proche de 1 indique qu'un accroissement de l'épargne nationale induit un accroissement identique de l'investissement national. L'épargne supplémentaire reste dans son pays d'origine et la mobilité des capitaux est faible.

La relation (2.6) peut également être interprétée en termes de flux de capitaux. L'équation (2.5) exprimée en part du produit national s'écrit

$$\frac{I}{Y} = \frac{S}{Y} - \frac{(X - M)}{Y} \quad (2.7)$$

L'équation (2.6) devient

$$\frac{S}{Y} - \frac{(X - M)}{Y} = \beta_1 + \beta_2 \frac{S}{Y} + \varepsilon \quad (2.8)$$

D'où

$$-\frac{(X - M)}{Y} = \beta_1 + (\beta_2 - 1) \frac{S}{Y} + \varepsilon \quad (2.9)$$

² En effet, théoriquement, le stock optimal de capital égalise la productivité marginale du capital et le taux d'intérêt réel.

Le solde du compte courant ($X - M$) traduit l'excédent du revenu national sur la dépense nationale

$$X - M = Y - C - I - G$$

Mais il peut également s'interpréter en termes de flux de capitaux. Ainsi, une valeur positive du compte courant indique que le pays est un prêteur net vis-à-vis du reste du monde. Une valeur négative indique que le pays dépense plus qu'il ne gagne et qu'il emprunte à l'étranger pour combler cet écart. Ainsi, un excédent du compte courant se traduit par une sortie nette de capitaux tandis qu'un déficit induit une entrée nette de capitaux sur le territoire national. Cela reflète la contrainte budgétaire de la nation : la partie de la dépense qui ne correspond pas à un revenu ($X - M$) doit être empruntée et inversement.

La balance des paiements doit toujours être nulle : le compte courant additionné du compte de capital et des interventions officielles des autorités monétaires sur le marché des changes doit être nul. Un coefficient estimé de β_2 égal à 1 implique donc que l'épargne nationale finance l'investissement national (2.6), mais également que les flux nets de capitaux ne dépendent pas du taux d'épargne nationale (équation (2.9)).

2.3 Principaux résultats de Feldstein et Horioka

L'objectif de Feldstein et Horioka était de tester l'hypothèse de mobilité forte du capital à long terme. Afin d'éliminer les effets des chocs conjoncturels, ils ont estimé leur modèle sur une coupe transversale à partir d'un échantillon de 16 pays de l'OCDE. Plusieurs périodes ont été examinées : 1960-1974, 1960-1964, 1965-1969 et 1970-1974. Dans le modèle estimé, la variable endogène est le taux d'investissement moyen sur la période d'analyse et la variable explicative est le taux d'épargne moyen sur la même période. La relation entre l'épargne et l'investissement est interprétée comme une relation de long terme dans laquelle les différences de taux d'épargne entre les pays reflètent des différences structurelles. Le paramètre β_2 représente l'impact d'un choc soutenu de l'épargne sur l'investissement. Il s'interprète comme un multiplicateur de long terme. L'équation suivante est estimée sur l'ensemble de la période 1960-1974, puis sur les sous-périodes de 5 ans 1960-1964, 1965-1969 et 1970-1974.

$$\overline{\left(\frac{I}{Y}\right)}_i = \beta_1 + \beta_2 \overline{\left(\frac{S}{Y}\right)}_i + \varepsilon_i \quad i = 1, \dots, 16 \quad (2.10)$$

avec

$$\overline{\left(\frac{I}{Y}\right)}_i = \frac{1}{T} \sum_{t=1}^T \left(\frac{I}{Y}\right)_{i,t} \quad \text{et} \quad \overline{\left(\frac{S}{Y}\right)}_i = \frac{1}{T} \sum_{t=1}^T \left(\frac{S}{Y}\right)_{i,t}$$

T est le nombre d'années de la période d'analyse et i caractérise les pays.

Les résultats d'estimation, avec la méthode des MCO, de l'équation (2.10) donnent

Tableau 2.1 Résultats de Feldstein et Horioka

Période	1960-1974	1960-1964	1965-1969	1970-1974
β_1	0.035 (0.018)	0.029 (0.015)	0.039 (0.025)	0.039 (0.024)
β_2	0.887 (0.074)	0.909 (0.060)	0.872 (0.101)	0.871 (0.092)
R^2	0.91	0.94	0.83	0.85

entre parenthèses figure l'écart type des coefficients estimés

À partir de ces estimations, les auteurs testent les deux hypothèses suivantes

$$\begin{array}{ll} H_0 : \beta_2 = 0 & \text{puis} \\ H_1 : \beta_2 \neq 0 & H_0 : \beta_2 = 1 \\ & H_1 : \beta_2 \neq 1 \end{array}$$

La table de Student pour un degré de liberté ³ égal à 14 indique une valeur du t de Student égale à 2.14, pour un seuil de risque de première espèce égal à 5 %. Par conséquent, l'intervalle de confiance de β_2 est donné pour les 4 périodes respectivement par

$$[0.887 \pm 2.14 \times 0.074] = [0.73 ; 1.04]$$

$$[0.909 \pm 2.14 \times 0.060] = [0.78 ; 1.03]$$

$$[0.872 \pm 2.14 \times 0.101] = [0.66 ; 1.08]$$

$$[0.871 \pm 2.14 \times 0.092] = [0.67 ; 1.06]$$

Quelle que soit la période d'analyse, la valeur 0 n'appartient pas à l'intervalle de confiance tandis que la valeur 1 y figure. Le paramètre β_2 est donc significativement différent de zéro (l'hypothèse $H_0 : \beta_2 = 0$ est refusée) tandis que le paramètre β_2 n'est pas significativement différent de un (l'hypothèse $H_0 : \beta_2 = 1$ n'est pas refusée). Ces résultats contredisent l'hypothèse de parfaite mobilité des capitaux, un accroissement du taux d'épargne induisant un accroissement quasi équivalent du taux d'investissement. Le supplément d'épargne est donc investi dans le pays d'origine. Les différences internationales de taux d'épargne n'expliquent pas les flux de capitaux observés.

Ces résultats, surprenants compte tenu des mouvements de capitaux internationaux observés, ont fait l'objet de critiques diverses. Feldstein et Horioka ont approfondi leur analyse dans plusieurs directions. Ils ont notamment suggéré que le paramètre β_2 pouvait varier avec le degré d'ouverture de l'économie et ont estimé l'équation d'investissement en supposant que $\beta_2 = \alpha_1 + \alpha_2 \frac{X+M}{Y}$, le rapport $\frac{X+M}{Y}$ représentant le degré d'ouverture de l'économie. Ils ont souligné également

³ Le degré de liberté est égal au nombre d'observations moins le nombre de paramètres estimés.

le problème d'endogénéité du taux d'épargne et ont estimé un modèle à équations simultanées dans lequel le taux d'épargne est endogénéisé. Les résultats d'estimation n'ont cependant pas été significativement modifiés.

Dans la section suivante on propose de reprendre le modèle de Feldstein et Horioka afin d'analyser la relation épargne-investissement des pays de l'OCDE (on dispose d'un échantillon de 19 pays) sur la période 1970-1998.

3. ÉVALUATION DE LA RELATION ÉPARGNE-INVESTISSEMENT

3.1 *Présentation de la base de données*

Les données sont issues de la publication de l'OCDE «Perspectives économique de l'OCDE», 2000. Elles concernent 19 pays de l'OCDE sur la période 1970-1998. Les séries PIB, épargne nationale, épargne privée, épargne publique et investissement national sont exprimées en monnaie nationale courante tandis que les séries exportations et importations en biens et services sont disponibles en part du PIB.

Ces données ont permis de calculer les taux d'investissement, d'épargne et de déficit commercial moyens sur différentes sous-périodes. Dans le fichier de données chapitre2.xls, on a dans l'ordre les 19 pays suivants : USA, Japon, Allemagne, France, Italie, Canada, Autriche, Belgique, Danemark, Finlande, Irlande, Pays-Bas, Norvège, Portugal, Espagne, Suède, Suisse, Australie et Royaume-Uni. Les séries sont des moyennes annuelles

- $txi = I/Y$ la part de l'investissement dans le PIB.
- $txs = S/Y$ est la part de l'épargne nationale dans le PIB.
- $txsprive = S_{privée}/Y$ est la part de l'épargne privée dans le PIB.
- $txspub = S_{publique}/Y$ est la part de l'épargne publique dans le PIB.
- $dc = -(X - M)/Y$ est le déficit commercial en part du PIB.

Par exemple, $txi7098$ est le taux d'investissement moyen annuel sur la période 1970-1998.

3.2 *Évaluation du degré de mobilité du capital à long terme*

Dans les problèmes suivants, la spécification initiale proposée par Feldstein et Horioka est estimée avec nos données. Ensuite, l'analyse est approfondie par la décomposition de la variable épargne nationale en deux composantes, l'épargne publique et l'épargne privée. On propose enfin de tester l'hypothèse de mobilité du capital à partir de l'équation (2.9) spécifiée en termes de déficit commercial.

Problème 2.1

Estimation du modèle de Feldstein et Horioka

Ce premier problème se base sur l'estimation de l'équation (2.10) avec les 19 pays de l'échantillon et les moyennes sont calculées sur la période 1970-1998.

1. Estimer avec la méthode des MCO l'équation

$$txi7098_i = \beta_1 + \beta_2 txs7098_i + \varepsilon_i \quad (2.11)$$

Puis, représenter sur une même figure les valeurs observées du taux d'investissement moyen des différents pays en fonction du taux d'épargne moyen ainsi que la droite de régression de l'échantillon. Commenter la figure et indiquer la différence entre l'aléa et le résidu du modèle.

2. Estimer le modèle sur un échantillon de taille plus faible (utiliser par exemple les 15 premiers pays de l'échantillon) et comparer les résultats obtenus avec l'échantillon de 19 pays.
3. Commenter les résultats statistiques, notamment la qualité de l'ajustement et la significativité des paramètres (utiliser les résultats de la première question). Donner l'intervalle de confiance du paramètre β_2 et tester l'hypothèse : $H_0 : \beta_2 = 1$ contre $H_1 : \beta_2 \neq 1$.
4. Estimer avec la méthode des MCO le modèle sur les sous-périodes 1970-1979, 1980-1989 et 1990-1998. Comparer les résultats.
5. Estimer avec la méthode des MCO le modèle sur la période 1970-1974 puis comparer les résultats à ceux de Feldstein et Horioka.

Problème 2.2

Extension du modèle initial

Ce problème propose un élargissement vers un modèle de régression linéaire multiple (cf. chapitre 3, Le modèle de régression linéaire multiple).

Le taux d'épargne nationale (txs) est la somme du taux d'épargne privée ($txsprive$) et du taux d'épargne publique ($txspub$). L'estimation de l'équation (2.11) conduit donc à imposer une contrainte selon laquelle l'impact d'une variation d'une unité du taux d'épargne privée et l'impact d'une variation d'une unité du taux d'épargne publique sur le taux d'investissement sont identiques et égaux à β_2 unités.

$$txi7098_i = \beta_1 + \beta_2 (txsprive7098_i + txspub7098_i) + \varepsilon_i \quad (2.12)$$

1. Estimer l'équation (2.13) sur l'ensemble de la période et sur les différentes sous-périodes 1970-1979, 1980-1989 et 1990-1998

$$txi_{ti} = \beta_1 + \beta_{21} txsprive_{ti} + \beta_{22} txspub_{ti} + \varepsilon_i \quad (2.13)$$

avec : $t = 7098, 7079, 8089, 9098$.

Commenter les résultats statistiques, notamment la qualité de l'ajustement, la significativité des paramètres et la significativité de la régression. Tester ensuite l'hypothèse $H_0 : \beta_{21} = \beta_{22} = \beta_2$ contre $H_1 : \beta_{21} \neq \beta_{22}$.

2. Procéder aux étapes suivantes avec les données calculées en moyenne sur la période 1990-1998.
 - a) Estimer avec la méthode des MCO le modèle

$$txi9098_i = \gamma_1 + \gamma_2 txsprive9098_i + \varepsilon_{1i}$$

on note $res1$ les résidus MCO de cette estimation

- b) Estimer avec la méthode des MCO le modèle

$$txspub9098_i = \gamma'_1 + \gamma'_2 txsprive9098_i + \varepsilon_{2i}$$

on note $res2$ les résidus MCO de cette estimation.

- c) Estimer avec la méthode des MCO le modèle

$$res1_i = \delta_1 + \delta_2 res2_i + \varepsilon_i \quad (2.14)$$

- d) Comparer les résultats d'estimation des équations (2.13) et (2.14).

Problème 2.3

Test du degré de mobilité des capitaux et déficit commercial

Dans la section précédente, on montre qu'il est équivalent de tester l'hypothèse $\beta_2 = 1$ dans le modèle (2.11) ou de tester $\beta_2 - 1 = 0$ dans le modèle

$$dc7098_i = PIMM_i - PEXM_i = \beta_1 + (\beta_2 - 1) (txs7098_i) + \varepsilon_i \quad (2.15)$$

où $PIMM$ et $PEXM$ indiquent respectivement la part des importations et la part des exportations dans le PIB en moyenne sur la période 1970-1998.

1. Estimer l'équation (2.15) sur la période 1970-1998 et commenter vos résultats.
2. Comme dans le problème 2.2, décomposer le taux d'épargne nationale en taux d'épargne privée et taux d'épargne publique, puis estimer le modèle

$$dc_i = \beta_1 + \gamma_1 txsprive_i + \gamma_2 txspub_i + \varepsilon_i \quad (2.16)$$

Tester ensuite l'hypothèse

$$\begin{aligned} H_0 : \gamma_1 &= \gamma_2 \\ H_1 : \gamma_1 &\neq \gamma_2 \end{aligned}$$

4. ÉLÉMENTS DE CORRIGÉ

Problème 2.1

Estimation du modèle de Feldstein et Horioka

1. L'estimation avec la méthode des MCO du modèle permet de déterminer l'équation de la droite de régression

Tableau 2.2 Estimation avec les MCO — Modèle de Feldstein-Horioka 1970-1998

Linear Regression — Estimation by Least Squares					
Dependent Variable TXI7098					
Usable Observations	19	Degrees of Freedom	17		
Centered R**2	0.580464	R Bar **2	0.555786		
Uncentered R**2	0.991993	T x R**2	18.848		
Mean of Dependent Variable	0.2302731641				
Std Error of Dependent Variable	0.0330007947				
Standard Error of Estimate	0.0219948324				
Sum of Squared Residuals	0.0082241351				
Regression F(1,17)	23.5210				
Significance Level of F	0.00015020				
Durbin-Watson Statistic	1.948668				

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	0.0854663615	0.0302814127	2.82240	0.01173824
2.	TXS7098	0.6230565192	0.1284694008	4.84984	0.00015020

L'équation de la droite s'écrit

$$esti7098_i = 0.085 + 0.623 \times txs7098_i \quad i = 1, \dots, 19$$

$esti7098$ correspond à la valeur estimée du taux d'investissement moyen sur la période 1970-1998. La série des résidus est donnée par l'écart entre $txi7098_i$ et $esti7098_i$, i allant de 1 à 19. Le résidu e_i donne une estimation de l'aléa ε_i qui dans le modèle théorique représente la partie non expliquée du ~~taux~~ taux moyen d'investissement par le taux moyen d'épargne du pays i .

$$\varepsilon_i = txi7098_i - \beta_1 - \beta_2 \times txs7098_i \quad i = 1, \dots, 19$$

$$e_i = txi7098_i - (0.085 + 0.623 \times txs7098_i) \quad i = 1, \dots, 19$$

La figure 2.1 représente la relation entre le taux d'épargne et le taux d'investissement moyen annuel sur la période 1970-1998, pour les pays de l'OCDE.

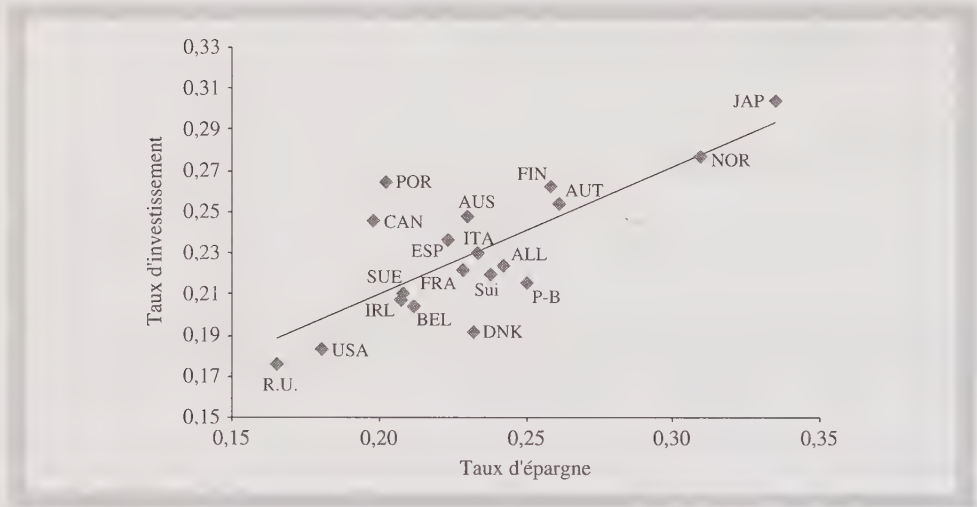


Figure 2.1 Relation Épargne-Investissement des Pays de l'OCDE (moyenne 1970-1998)

Les losanges correspondent aux valeurs observées du taux d'épargne annuel moyen (en ordonnée) et du taux d'investissement annuel moyen (en abscisse) pour les 19 pays considérés dans l'échantillon. La droite tracée est la droite de régression de l'échantillon. Pour un pays, l'écart entre le point observé et le point correspondant de la droite est le résidu d'estimation.

2. Les paramètres estimés $\hat{\beta}_1$ et $\hat{\beta}_2$ de l'équation (2.11) prennent des valeurs différentes selon le nombre de pays retenus dans l'échantillon. Ainsi, dans le cas où la régression porte sur 15 pays on obtient⁴

$$esti15_i = 0.095 + 0.585 \times txs7098_i \quad i = 1, \dots, 15$$

La figure 2.2 représente la droite de régression selon l'échantillon choisi. Elle illustre le fait que les estimateurs $\hat{\beta}_1$ et $\hat{\beta}_2$ sont des variables aléatoires. Ils suivent la même loi de probabilité que les aléas car ils sont des fonctions linéaires de ces derniers. Ainsi, si les aléas suivent une loi normale, alors les estimateurs $\hat{\beta}_1$ et $\hat{\beta}_2$ suivent aussi une loi normale.

⁴ On note $esti15$, la série des valeurs estimées du taux d'investissement, lorsque l'échantillon comporte 15 pays.

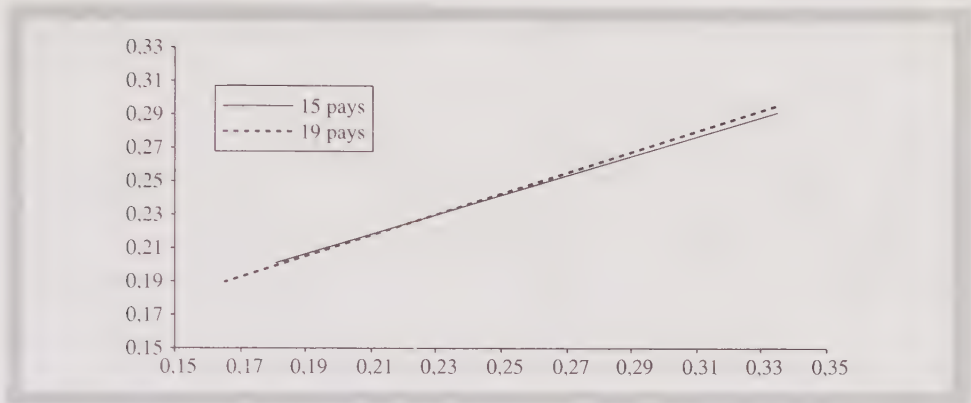


Figure 2.2 Droite de régression de l'échantillon

3. Les résultats d'estimation présentés dans le tableau 2.2, indiquent que 58 % des variations du taux moyen d'investissement sont expliquées par les variations du taux moyen d'épargne ($R^2 = 0.58$). Dans le cas d'une régression simple, il est équivalent de tester la significativité de la régression et la significativité du paramètre β_2 . Le test de Student et le test de Fisher concluent au rejet de l'hypothèse H_0 selon laquelle β_2 est égal à 0

$$Tstat = \frac{0.623}{0.128} = 4.84 > \tau(17) = 2.11$$

$$Fstat = \left(\underbrace{4.84}_{Tstat} \right)^2 = \frac{R^2/1}{(1-R^2)/(19-2)} = 23.52 > F(1,17) = 4.45$$

On note ici trois manières alternatives pour présenter les résultats du test de significativité du paramètre β_2 ($H_0 : \beta_2 = 0$) : l'intervalle de confiance, le calcul de la statistique $Tstat$ ou le calcul du niveau de significativité.

- a) Un intervalle de confiance donne un ensemble de valeurs possibles de β_2 , avec usuellement un niveau de confiance de 95 %, c'est-à-dire avec un risque de première espèce de $\alpha = 5$ % qui correspond à la probabilité de refuser l'hypothèse H_0 alors qu'elle est vraie

$$\Pr \left[\hat{\beta}_2 - \tau_{\alpha/2, (N-K)} \hat{\sigma}_{\hat{\beta}_2} \leq \beta_2 \leq \hat{\beta}_2 + \tau_{\alpha/2, (N-K)} \hat{\sigma}_{\hat{\beta}_2} \right] = 95 \%$$

N est le nombre d'observations, K le nombre de paramètres à estimer et $\hat{\sigma}_{\hat{\beta}_2}$ l'écart type estimé de $\hat{\beta}_2$.

$$\hat{\sigma}_{\hat{\beta}_2} = \sqrt{\frac{\sum_{i=1}^{19} e_i^2}{19-2}} \sqrt{\left(\sum_{i=1}^{19} (txs7098_i - \overline{txs7098})^2 \right)^{-1}} = 0.128$$

$\tau_\alpha(17)$ étant égal à 2.11, l'intervalle de confiance du paramètre β_2 est donné par

$$[0.623 \pm 2.11(0.128)] = [0.352 ; 0.893]$$

β_2 a 95 % de chance de se trouver dans l'intervalle $[0.352 ; 0.893]$. La valeur 0 n'appartenant pas à cet intervalle, l'hypothèse H_0 est refusée. On dit également que β_2 est significativement différent de 0.

b) Par ailleurs, l'intervalle de confiance de β_2 peut se réécrire sous la forme

$$\Pr \left[|Tstat| = \left| \frac{\widehat{\beta}_2 - \beta_2}{\widehat{\sigma}_{\widehat{\beta}_2}} \right| \leq \tau_{\alpha/2}(N - K) \right] = 95 \%$$

Pour tester $H_0 : \beta_2 = 0$, on calcule $Tstat = \frac{0.623}{0.128} = 4.84$. La valeur de la statistique $Tstat$ étant supérieure à 2.11, on rejette l'hypothèse H_0 .

c) Enfin, on peut également calculer le niveau de significativité du test. Il indique le niveau de risque de première espèce α' que l'on doit choisir pour que la valeur testée de β_2 appartienne à l'intervalle de confiance qui lui, sera défini pour un niveau de confiance égal à $1 - \alpha'$. On calcule le risque de première espèce α' tel que

$$\left[\left| \frac{\widehat{\beta}_2}{\widehat{\sigma}_{\widehat{\beta}_2}} \right| \leq \tau_{\alpha'/2}(17) \right]$$

Dans l'exercice, pour accepter l'hypothèse $H_0 : \beta_2 = 0$, on doit choisir un risque de première espèce égal à 0.01 %⁵. Ce risque étant inférieur à 5 %, on rejette l'hypothèse H_0 , car réduire le risque de première espèce conduit à accroître le risque de seconde espèce, c'est-à-dire la probabilité d'accepter l'hypothèse H_0 alors qu'elle est fautive.

Pour terminer l'analyse des résultats du tableau 2.2, on note que la valeur 1 n'est pas incluse dans l'intervalle de confiance. Par conséquent on refuse l'hypothèse $H_0 : \beta_2 = 1$ ou de manière équivalente

$$Tstat = \left| \frac{0.6236 - 1}{0.128} \right| = 2.94 > \tau(17) = 2.11$$

Sur l'ensemble de la période, selon ces estimations, une variation soutenue du taux d'épargne de 1 point de pourcentage induit une variation du taux d'investissement de 0.62 point de pourcentage, et non pas d'1 point. On ne peut cependant pas conclure à une mobilité parfaite des capitaux à long terme puisque le coefficient β_2 est significativement différent de 0.

4. L'estimation du modèle sur les trois sous-périodes 1970-1979, 1980-1989 et 1990-1998 donne les résultats suivants

⁵ Ce niveau de significativité du test apparaît dans la colonne « Signif » du tableau 2.2.

Tableau 2.3 Estimation de l'équation sur les 3 sous-périodes

	1970-1979	1980-1989	1990-1998
β_2	0.716 (0.138)	0.652 (0.141)	0.397 (0.148)
T-stat	5.18	4.61	2.67
F-stat	26.85	21.33	7.13
R^2	0.612	0.556	0.295

entre parenthèses figure l'écart type estimé des paramètres

Les résultats restent relativement proches lorsqu'on compare les deux premières sous-périodes, les intervalles de confiance du paramètre β_2 étant respectivement donnés par

$$[0.716 \pm 2.11 \times 0.138] = [0.424 ; 1.00]$$

$$[0.652 \pm 2.11 \times 0.141] = [0.354 ; 0.949]$$

Sur la dernière période, qui correspond à une libéralisation accrue des marchés des capitaux en Europe, on obtient l'intervalle de confiance suivant

$$[0.397 \pm 2.11 \times 0.148] = [0.084 ; 0.709]$$

L'hypothèse $H_0 : \beta_2 = 0$ est rejetée sur les 3 sous-périodes. En revanche l'hypothèse $H_0 : \beta_2 = 1$ n'est pas refusée pour la première période. Aucune période n'est donc caractérisée par une mobilité parfaite des capitaux car β_2 est toujours significativement différent de 0. On ne peut pas non plus conclure que les 19 pays ont un comportement correspondant à celui d'une économie fermée car β_2 est significativement différent de 1 dans les deux dernières sous-périodes, notamment dans la dernière. Cependant, lors des deux premières sous-périodes, le taux d'épargne a un impact relativement élevé (c'est-à-dire proche de 1) sur le taux d'investissement : le comportement des économies considérées est plus proche de celui d'une économie fermée en termes de relation épargne investissement. Il est intéressant de noter que sur la dernière sous-période le coefficient estimé de β_2 diminue de manière significative sans toutefois devenir non significativement différent de 0.

5. Les résultats sont très proches de ceux de Feldstein et Horioka avec notre échantillon de 19 pays sur la période 1970-1974 car on obtient une valeur de β_2 égale à 0.71 et la valeur de son écart type est de 0.11.

Problème 2.2

Extension du modèle initial avec prise en compte de la structure de l'épargne

Les résultats sont commentés pour la période 1970-1998.

1. Lorsque le modèle est élargi et estimé sans contrainte sur les paramètres avec les données calculées en moyenne sur l'ensemble de la période 1970-1998, on obtient les résultats présentés dans le tableau 2.4.

Tableau 2.4 Estimation du modèle élargi de Feldstein-Horioka

Linear Regression — Estimation by Least Squares					
Dependent Variable TXI7098					
Usable Observations	19	Degrees of Freedom	16		
Centered R**2	0.600606	R Bar **2	0.550682		
Uncentered R**2	0.992377	T x R**2	18.855		
Mean of Dependent Variable	0.2302731641				
Std Error of Dependent Variable	0.0330007947				
Standard Error of Estimate	0.0221208254				
Sum of Squared Residuals	0.0078292947				
Regression F(2,16)	12.0304				
Significance Level of F	0.00064746				
Durbin-Watson Statistic	1.914128				

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	0.0672940792	0.0365617283	1.84056	0.08430837
2.	TXSPRIVE7098	0.7077431175	0.1599441873	4.42494	0.00042468
3.	TXSPUB7098	0.5418280893	0.1577056503	3.43569	0.00339457

$F(1,16) = 0.80690$ with Significance Level 0.38235777

Le modèle permet d'expliquer 60 % de la variation du taux d'investissement annuel moyen ($R^2 = 0.60$).

Les tests de significativité des paramètres indiquent que les coefficients β_{21} et β_{22} sont significativement différents de 0 car les $Tstat$, respectivement égaux à 4.42 et 3.43, sont plus élevés que la valeur lue dans la table de Student égale à 2.12 pour un degré de liberté égal au nombre d'observations moins le nombre de paramètres à estimer, c'est-à-dire 16. On obtient la même conclusion avec le niveau de significativité des tests indiqué dans le tableau 2.4 car ces niveaux sont inférieurs à 5 %.

De même, le test de Fisher, qui permet de tester l'hypothèse $H_0 : \beta_{21} = \beta_{22} = 0$, donne une valeur de 12.03 et un niveau de significativité inférieur à 5 %. Le test de significativité de la régression conduit donc à rejeter l'hypothèse H_0 .

$$F = \frac{R^2/K - 1}{(1 - R^2)/(N - K)} = \frac{0.6000606/(3 - 1)}{(1 - 0.6000606)/(19 - 3)} = 12.03 \quad (2.17)$$

Le modèle estimé étant un modèle de régression linéaire multiple, les tests de Student et de Fisher sont différents. Le premier s'interroge sur la significativité d'un paramètre du modèle tandis que le test de Fisher teste l'égalité à 0 de tous les paramètres du modèle, sauf la constante (cf. chapitre 3 Le modèle de régression linéaire multiple).

La dernière ligne du tableau 2.4 indique également le résultat du test de contrainte d'égalité des paramètres β_{21} et β_{22} . L'objectif est de déterminer si une variation du taux d'épargne privée a le même impact sur le taux d'investissement qu'une variation du taux d'épargne publique. La statistique de Fisher est ici égale à 0.80. Elle est mesurée de la manière suivante

$$F = \frac{(0.0082241351 - 0.007829947) / 1}{0.007829947 / 16} = 0.80$$

Les valeurs 0.0082241351 et 0.007829947 correspondent respectivement à la somme des carrés des résidus du modèle estimé avec la contrainte $\beta_{21} = \beta_{22}$ (tableau 2.2) et à celle du modèle estimé sans cette contrainte (tableau 2.4). Au numérateur l'écart entre les deux sommes de carrés est divisé par 1, ce qui correspond au nombre de contrainte. Au dénominateur la somme des carrés des résidus du modèle non contraint est divisée par son degré de liberté égal ici à 16. Le niveau de significativité du test étant égal à 0.38, cette hypothèse n'est pas rejetée car le niveau de significativité est plus élevé que 0.05⁶. Ainsi, une variation du taux d'épargne privée ou publique semble avoir la même incidence sur le taux d'investissement.

2. La dernière question de ce problème, relative à la période 1990-1998, a pour but d'expliquer que, dans un modèle de régression linéaire multiple, les paramètres estimés sont des coefficients de régression partiels.

En pratiquant les 3 régressions proposées successivement on constate que $\hat{\delta}_2 = \hat{\beta}_{22}$. Les paramètres d'un modèle de régression multiple sont calculés de manière à évaluer le véritable impact de la variable exogène sur la variable endogène, dans la relation entre l'investissement et l'épargne

$$\frac{\partial txi}{\partial txpub} = \beta_{22} \quad (2.18)$$

le paramètre β_{22} mesure l'impact de l'épargne publique sur l'investissement. Il est estimé compte tenu de l'effet que peut avoir sur cette relation l'épargne privée. La série *res1* représente la partie du taux d'investissement non expliquée par le taux d'épargne privée et la série *res2* représente la partie du taux d'épargne publique non affectée par le taux d'épargne privée.

Le même calcul peut-être effectué pour le paramètre β_{21} .

Problème 2.3

Test du degré de mobilité des capitaux et déficit commercial

1. Les résultats de l'estimation du modèle dans lequel la variable endogène est le déficit commercial moyen annuel sur la période 1970-1998 par unité de PIB, et la variable exogène, le taux d'épargne nationale moyen annuel, sont présentés dans le tableau 2.5.

⁶ De manière équivalente, on constate que la valeur calculée du Fisher, égale à 0.68, est inférieure à sa valeur tabulée car $F_{(1,16)}=4$. Donc l'hypothèse H_0 n'est pas refusée.

Tableau 2.5 Part du déficit commercial dans le PIB en fonction du taux d'épargne

Linear Regression — Estimation by Least Squares					
Dependent Variable DC7098					
Usable Observations	19			Degrees of Freedom	17
Centered R**2	0.180960			R Bar **2	0.132781
Uncentered R**2	0.191355			T x R**2	3.636
Mean of Dependent Variable	-0.002781194				
Std Error of Dependent Variable	0.025202096				
Standard Error of Estimate	0.023469353				
Sum of Squared Residuals	0.0093637787				
Regression F(1,17)	3.7560				
Significance Level of F	0.06940965				
Durbin-Watson Statistic	2.139158				

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	0.058964163	0.032311460	1.82487	0.08564458
2.	TXS7098	-0.265670168	0.137081911	-1.93804	0.06940965

La qualité de l'ajustement est moins bonne que précédemment car le R^2 n'est que de 18 %. Les tests de Fisher et de Student conduisent à ne pas rejeter l'hypothèse $H_0 : (\beta_2 - 1) = 0$ car le niveau de significativité du test est de 6.9 % (supérieur à 5 %). Ce résultat est différent de celui obtenu dans le problème 2.1 puisque l'hypothèse nulle était alors rejetée. Cependant, on note que l'estimation du modèle (2.15) donne un niveau de significativité du test proche de 5 %, tandis que la borne supérieure de l'intervalle de confiance de β_2 , obtenue par l'estimation du modèle (2.11), est proche de 1. Les résultats d'estimation des deux modèles conduisent donc à considérer qu'il y a une faible mobilité des capitaux.

2. Lorsqu'on distingue l'épargne publique et l'épargne privée dans le modèle, le résultat est identique au résultat obtenu dans le problème 2.2. Le test de contrainte conduit à ne pas rejeter l'hypothèse $H_0 : \gamma_1 = \gamma_2$, la statistique de Fisher est égale à 0.046 et le niveau de significativité du test est égal à 83 %.

3

LE MODÈLE DE RÉGRESSION LINÉAIRE MULTIPLE

1. Présentation du modèle
2. Tests statistiques usuels
3. Prédications et qualité prédictive d'un modèle estimé
4. Les variables indicatrices dans les modèles économétriques

OBJECTIFS DU CHAPITRE

Présenter le modèle linéaire multiple avec une écriture matricielle

- Estimateur des Moindres Carrés Ordinaires
- Analyse de la variance
- Coefficient de détermination, coefficient de détermination ajusté
- Tests de contraintes (une et plusieurs contraintes)
- Tests de significativité (des coefficients et de la régression)
- Test de stabilité (Chow, CUSUM et CUSUM of Square)
- Prédications
- Variables indicatrices

NIVEAU

LICENCE, troisième année

Le modèle de régression linéaire multiple est un modèle composé d'une équation dans lequel une variable endogène est expliquée par plusieurs variables. Les paramètres de ce modèle sont estimés par la méthode des Moindres Carrés Ordinaires (MCO) lorsque certaines hypothèses sur les aléas sont satisfaites. Par ailleurs, dans toutes les études économétriques, des tests statistiques sont pratiqués afin d'établir si les résultats obtenus sont satisfaisants d'un point de vue statistique et économique. Dans une première section, les hypothèses du modèle de régression linéaire multiple et la méthode d'estimation des MCO sont présentées. Dans la seconde section, les tests statistiques usuels sont examinés. Dans la troisième section les prédictions et la qualité prédictive du modèle sont étudiées. Enfin, dans la dernière section, le cas particulier des modèles composés de variables explicatives discrètes est examiné.

1. PRÉSENTATION DU MODÈLE

De manière générale, le modèle de régression linéaire multiple s'écrit

$$\boxed{Y = X\beta + \varepsilon} \quad (3.1)$$

avec

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_N \end{pmatrix}_{(N,1)}, \quad X = \begin{pmatrix} 1 & x_{21} & \cdot & \cdot & x_{K1} \\ 1 & x_{22} & \cdot & \cdot & x_{K2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{2N} & \cdot & \cdot & x_{KN} \end{pmatrix}_{(N,K)}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \beta_K \end{pmatrix}_{(K,1)}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \varepsilon_N \end{pmatrix}_{(N,1)}$$

N indique le nombre d'observations contenues dans le modèle ($i = 1, \dots, N$). Le nombre de paramètres à estimer est égal à K ($k = 1, \dots, K$).

$$y_i = \beta_1 + \beta_2 x_{2i} + \dots + \beta_K x_{Ki} + \varepsilon_i \quad i = 1, \dots, N \quad (3.2)$$

$$= [\beta_1 \quad \beta_2 \dots \beta_K] \begin{bmatrix} 1 \\ x_{2i} \\ \cdot \\ \cdot \\ x_{Ki} \end{bmatrix} + \varepsilon_i = \beta' x_i + \varepsilon_i$$

$$= [1 \quad x_{2i} \dots x_{Ki}] \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \beta_K \end{bmatrix} + \varepsilon_i = x'_i \beta + \varepsilon_i$$

On note

Y : le vecteur des N observations y_i sur la variable endogène du modèle (ou variable expliquée par le modèle).

X : la matrice des variables exogènes du modèle. Chaque colonne de la matrice est une variable explicative. Il s'agit d'un modèle avec constante car le vecteur composé uniquement de la valeur 1 apparaît dans la matrice.

β : le vecteur des paramètres à estimer. Il y a K paramètres à estimer puisqu'il y a K variables explicatives. β' est la transposée du vecteur β .

ε : le vecteur des aléas. Les aléas prennent en compte le fait que la relation entre la variable endogène et les différentes variables explicatives n'est pas exacte.

Hypothèses sur le modèle de régression linéaire

Les hypothèses sur le modèle sont

Hypothèse 1

L'espérance des aléas est nulle

$$E(\varepsilon) = \begin{bmatrix} E(\varepsilon_1) \\ E(\varepsilon_2) \\ \vdots \\ E(\varepsilon_N) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (3.3)$$

Sous cette hypothèse, l'espérance de Y conditionnellement à X est donnée par l'expression

$$E(Y/X) = X\beta$$

Pour une observation i $E(y_i/x_i) = \beta' x_i$

Hypothèse 2

Les aléas sont homoscedastiques (leur variance est constante) et non auto-corrélés (leur covariance est nulle). Ainsi, la matrice de variance-covariance des aléas est donnée par l'expression

$$Var(\varepsilon) = E \left(\underbrace{(\varepsilon - E(\varepsilon))}_0 \underbrace{(\varepsilon - E(\varepsilon))'}_0 \right) = E(\varepsilon\varepsilon') = \sigma^2 I \quad (3.4)$$

avec

$$Var(\varepsilon) = \begin{pmatrix} var(\varepsilon_1) & cov(\varepsilon_1\varepsilon_2) & \cdot & \cdot & cov(\varepsilon_1\varepsilon_N) \\ cov(\varepsilon_2\varepsilon_1) & var(\varepsilon_2) & \cdot & \cdot & cov(\varepsilon_2\varepsilon_N) \\ \vdots & \vdots & \cdot & \cdot & \vdots \\ cov(\varepsilon_N\varepsilon_1) & cov(\varepsilon_N\varepsilon_2) & \cdot & \cdot & var(\varepsilon_N) \end{pmatrix} = \begin{pmatrix} \sigma^2 & 0 & \cdot & \cdot & 0 \\ 0 & \sigma^2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \sigma^2 \end{pmatrix}$$

$$\text{var}(\varepsilon_i) = E(\varepsilon_i - \underbrace{E(\varepsilon_i)}_0)^2 = E(\varepsilon_i^2) - \sigma^2 \quad i = 1, \dots, N$$

$$\text{cov}(\varepsilon_i, \varepsilon_j) = E((\varepsilon_i - \underbrace{E(\varepsilon_i)}_0)(\varepsilon_j - \underbrace{E(\varepsilon_j)}_0)) = E(\varepsilon_i \varepsilon_j) = 0 \quad i, j = 1, \dots, N, \quad i \neq j$$

Pour x_i fixé, cette hypothèse implique que

$$\text{var}(y_i) = E(y_i - E(y_i))^2 = \text{var}(\varepsilon_i) = \sigma^2 \quad i = 1, \dots, N$$

et

$$\text{cov}(y_i, y_j) = E((y_i - E(y_i))(y_j - E(y_j))) = \text{cov}(\varepsilon_i \varepsilon_j) = 0 \quad i, j = 1, \dots, N, \quad i \neq j.$$

Par conséquent

$$\text{Var}(Y) = \begin{pmatrix} \text{var}(y_1) & \text{cov}(y_1 y_2) & \cdot & \cdot & \text{cov}(y_1 y_N) \\ \text{cov}(y_2 y_1) & \text{var}(y_2) & \cdot & \cdot & \text{cov}(y_2 y_N) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \text{cov}(y_N y_1) & \text{cov}(y_N y_2) & \cdot & \cdot & \text{var}(y_N) \end{pmatrix} = \sigma^2 I \quad (3.5)$$

Si les hypothèses 1 et 2 sont vérifiées, l'aléa ε est un bruit blanc.

Hypothèse 3

La matrice des variables explicatives est composée de variables certaines c'est-à-dire non aléatoires; dans ce cas

$$E(Y/X) = E(Y)$$

Hypothèse 4

Le rang de la matrice des variables explicatives est égal à K , le nombre de paramètres à estimer. Autrement dit, les variables explicatives sont linéairement indépendantes. Dans ce cas, la matrice $X'X$ est inversible, X' désignant la transposée de la matrice X .

Hypothèse 5

La matrice de variance-covariance empirique des variables explicatives S_X tend vers V_X une matrice finie, définie positive quand N tend vers l'infini

$$\lim_{N \rightarrow +\infty} S_X = \lim_{N \rightarrow +\infty} \frac{X'X}{N} = V_X \quad (K, K)$$

Lorsque la taille de l'échantillon s'accroît, les variables explicatives x_{ki} conservent toujours une certaine variance et toute nouvelle observation apporte de l'information.

1.1 Estimateur des Moindres Carrés Ordinaires (MCO)

L'estimateur des MCO est dérivé de la minimisation de la somme des carrés des aléas ¹

$$\underset{\beta}{\text{Min}} \varepsilon' \varepsilon = \underset{\beta}{\text{Min}} (Y - X\beta)' (Y - X\beta)$$

Soit

$$S = (Y - X\beta)' (Y - X\beta)$$

D'après les conditions du premier ordre, $\frac{\partial S}{\partial \beta} = 0$, la solution est donnée par

$$\boxed{\widehat{\beta}_{MCO} = (X'X)^{-1}X'Y} \quad (3.6)$$

avec

$$\widehat{\beta}_{MCO} = \begin{bmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \\ \vdots \\ \widehat{\beta}_K \end{bmatrix}$$

1.1.1 Propriétés finies de l'estimateur des MCO

L'estimateur $\widehat{\beta}_{MCO}$ est calculé avec les observations sur les x_{ki} et y_i ($k = 1, \dots, K$, $i = 1, \dots, N$), sa valeur dépend de l'échantillon de données retenu. Chaque estimateur $\widehat{\beta}_k$ est une variable aléatoire.

$$\overset{1}{\underbrace{[\varepsilon_1 \quad \varepsilon_2 \quad \dots \quad \varepsilon_N]}_{\varepsilon'}} \underbrace{\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix}}_{\varepsilon} = \varepsilon' \varepsilon = \sum_{i=1}^N \varepsilon_i^2$$

Théorème de Gauss-Markov

Sous les hypothèses 1, 2, 3 et 4, l'estimateur des MCO est *Best Linear Unbiased Estimator* (BLUE). C'est-à-dire que, parmi les estimateurs linéaires non biaisés de β , l'estimateur des MCO est à variance minimale. On dit aussi que c'est un estimateur efficace. Son espérance et sa matrice de variance-covariance sont données par (3.7) et (3.8)

$$\begin{aligned}\widehat{\beta}_{MCO} &= (X'X)^{-1}X'[X\beta + \varepsilon] \\ &= \beta + (X'X)^{-1}X'\varepsilon\end{aligned}$$

sous les hypothèses 1 et 3

$$\boxed{E(\widehat{\beta}_{MCO}) = \beta} \quad (3.7)$$

sous l'hypothèse 2

$$\begin{aligned}Var(\widehat{\beta}_{MCO}) &= E \left[\left(\widehat{\beta}_{MCO} - E(\widehat{\beta}_{MCO}) \right) \left(\widehat{\beta}_{MCO} - E(\widehat{\beta}_{MCO}) \right)' \right] \\ &= E \left[\left(\widehat{\beta}_{MCO} - \beta \right) \left(\widehat{\beta}_{MCO} - \beta \right)' \right] \\ &= E \left[\left((X'X)^{-1}X'\varepsilon \right) \left((X'X)^{-1}X'\varepsilon \right)' \right] \\ &= (X'X)^{-1}X'E(\varepsilon\varepsilon')X(X'X)^{-1}\end{aligned}$$

$$\boxed{Var(\widehat{\beta}_{MCO}) = \sigma^2(X'X)^{-1}} \quad (3.8)$$

avec $Var(\widehat{\beta}_{MCO})$ la matrice de variance-covariance de $\widehat{\beta}_{MCO}$

$$Var(\widehat{\beta}_{MCO}) = \begin{pmatrix} var(\widehat{\beta}_1) & cov(\widehat{\beta}_1, \widehat{\beta}_2) & \cdot & \cdot & cov(\widehat{\beta}_1, \widehat{\beta}_K) \\ cov(\widehat{\beta}_2, \widehat{\beta}_1) & var(\widehat{\beta}_2) & \cdot & \cdot & cov(\widehat{\beta}_2, \widehat{\beta}_K) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ cov(\widehat{\beta}_K, \widehat{\beta}_1) & cov(\widehat{\beta}_K, \widehat{\beta}_2) & \cdot & \cdot & var(\widehat{\beta}_K) \end{pmatrix}$$

(K,K)

1.1.2 Propriétés asymptotiques de l'estimateur des MCO

De plus, sous l'hypothèse 5, lorsque $N \rightarrow \infty$, la matrice de variance-covariance de l'estimateur des MCO peut s'écrire sous la forme

$$Var(\widehat{\beta}_{MCO}) = \frac{\sigma^2}{N}V_X$$

et

$$\lim_{N \rightarrow +\infty} \text{Var}(\widehat{\beta}_{MCO}) = 0$$

L'estimateur des MCO converge alors en probabilité vers β , c'est-à-dire que pour η une quantité positive arbitrairement petite

$$\lim_{N \rightarrow +\infty} \text{Pr} \left\{ \left| \widehat{\beta}_{MCO} - \beta \right| < \eta \right\} = 1 \quad \eta > 0$$

où Pr désigne la probabilité.

Autrement dit, $\widehat{\beta}_{MCO}$ s'approche de sa vraie valeur β lorsque la taille de l'échantillon s'accroît. Cette propriété s'exprime également sous la forme

$$\boxed{P \lim \widehat{\beta}_{MCO} = \beta}$$

où $P \lim$ désigne la probabilité limite.

Par ailleurs, on note que si les aléas sont indépendants entre eux et identiquement distribués, quand $N \rightarrow +\infty$, l'estimateur centré dilaté $\sqrt{N}(\widehat{\beta}_{MCO,N} - \beta)$ suit asymptotiquement une loi normale

$$\sqrt{N}(\widehat{\beta}_{MCO,N} - \beta) \sim N(0, \sigma^2 V_X^{-1})$$

où $\widehat{\beta}_{MCO,N}$ indique l'estimateur des MCO lorsque la taille de l'échantillon est égale à N .

En conclusion sur les propriétés de l'estimateur des MCO : $\widehat{\beta}_{MCO}$ est BLUE, c'est l'estimateur le plus précis² de β et cette précision s'améliore avec la taille de l'échantillon puisque sa variance devient nulle lorsque la taille de l'échantillon tend vers l'infini.

1.2 Équation d'analyse de la variance et qualité de l'ajustement

Ayant évalué par la méthode des MCO les paramètres du modèle on cherche à déterminer si le modèle permet de bien expliquer la variable endogène.

Soit e le vecteur des résidus et \widehat{Y} le vecteur des valeurs estimées de Y

$$\boxed{Y = \widehat{Y} + e \text{ avec } \widehat{Y} = X \widehat{\beta}_{MCO}} \quad (3.9)$$

² Il a la variance la plus faible parmi les estimateurs non biaisés de β .

Selon la condition du premier ordre du problème de minimisation de la somme des carrés des aléas

$$\begin{aligned}\widehat{\beta}_{MCO} &= (X'X)^{-1}X'Y \iff X'Y - (X'X)\widehat{\beta}_{MCO} = 0 \\ &\iff X' \underbrace{[Y - X\widehat{\beta}_{MCO}]}_{\epsilon} = 0\end{aligned}$$

Par conséquent

$$X'\epsilon = \begin{pmatrix} 1 & 1 & \cdot & \cdot & 1 \\ x_{21} & x_{22} & \cdot & \cdot & x_{2N} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{K1} & x_{K2} & \cdot & \cdot & x_{KN} \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ e_N \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^N e_i \\ \sum_{i=1}^N x_{2i}e_i \\ \cdot \\ \cdot \\ \sum_{i=1}^N x_{Ki}e_i \end{pmatrix} = 0$$

La somme des résidus étant nulle

$$\sum_{i=1}^N y_i = \sum_{i=1}^N (\widehat{y}_i + e_i) = \sum_{i=1}^N \widehat{y}_i$$

D'où

$$\bar{y} = \overline{\widehat{y}}$$

On montre alors que

$$\boxed{\sum_{i=1}^N (y_i - \bar{y})^2 = \sum_{i=1}^N (\widehat{y}_i - \bar{y})^2 + \sum_{i=1}^N e_i^2} \quad (3.10)$$

Cette équation est l'équation d'analyse de la variance. Au facteur $1/N$ près, la variance empirique de y , appelée la variance totale, est la somme de la variance expliquée par le modèle et de la variance résiduelle. À partir de cette équation, la qualité de l'ajustement peut être mesurée par le coefficient de détermination noté R^2

$$R^2 = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^N (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^N e_i^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (3.11)$$

Cette statistique est comprise dans l'intervalle $[0,1]$. Une valeur proche de un indique que la qualité de l'ajustement est bonne dans la mesure où la part de la variance de Y expliquée par le modèle est élevée. L'estimateur des MCO minimise la somme des carrés des résidus et par conséquent maximise le R^2 . Cependant, cette statistique croît avec le nombre de variables exogènes, c'est pourquoi, on calcule également le coefficient de détermination ajusté, \bar{R}^2 , qui tient compte du nombre de variables explicatives présentes dans le modèle

$$\bar{R}^2 = 1 - \frac{\frac{\sum_{i=1}^N e_i^2}{N - K}}{\frac{\sum_{i=1}^N (y_i - \bar{y})^2}{N - 1}} \quad (3.12)$$

Cette statistique est utile pour comparer des modèles emboîtés.

2. TESTS STATISTIQUES USUELS

Une étude économétrique consiste non seulement à estimer les paramètres d'un modèle, mais aussi, à tester des hypothèses afin de valider le modèle économique théorique. On doit effectuer des tests statistiques afin de compléter les résultats des estimations. Ces tests se fondent sur la distribution statistique des paramètres estimés. On examine ici les tests de contraintes sur les paramètres et les tests de changements structurels et de stabilité des paramètres.

2.1 *Distribution de l'estimateur des MCO*

Afin de déterminer la loi de l'estimateur des MCO, une hypothèse supplémentaire sur la distribution des aléas du modèle doit être posée.

Hypothèse 6

Les aléas sont supposés indépendants et identiquement distribués selon une loi normale.

Sous l'hypothèse 6 la variable endogène Y est distribuée selon une loi normale

$$\varepsilon \sim N(0, \sigma^2 I) \implies Y \sim N(X\beta, \sigma^2 I) \tag{3.13}$$

$\hat{\beta}_{MCO}$ étant un estimateur BLUE et sa matrice de variance-covariance étant égale à $\sigma^2(X'X)^{-1}$

$$\boxed{\hat{\beta}_{MCO} \sim N(\beta, \sigma^2(X'X)^{-1})} \tag{3.14}$$

Ce résultat implique que pour un paramètre $\beta_k, k = 1, \dots, K$

$$\frac{\hat{\beta}_{kMCO} - \beta_k}{\sqrt{\text{var}(\hat{\beta}_{kMCO})}} \sim N(0, 1) \tag{3.15}$$

avec $\text{var}(\hat{\beta}_{kMCO}) = \sigma^2 s_{kk}$ où s_{kk} est l'élément (k, k) de la matrice $(X'X)^{-1}$. Cette variance ne peut pas être calculée car la variance des aléas n'est pas connue.

Or, $\hat{\sigma}^2 = \frac{1}{N-K} \sum_{i=1}^N e_i^2 = \frac{e'e}{N-K}$ est un estimateur non biaisé de σ^2 ($E(\hat{\sigma}^2) = \sigma^2$) et

$$\frac{\hat{\sigma}^2}{\sigma^2} \sim \frac{\chi^2(N-K)}{N-K} \tag{3.16}$$

Ainsi, on peut obtenir une estimation de la matrice de variance-covariance des paramètres estimés du modèle $\widehat{\text{Var}}(\hat{\beta}) = \hat{\sigma}^2(X'X)^{-1}$ et la statistique

$$\boxed{\frac{\frac{\hat{\beta}_{kMCO} - \beta_k}{\sqrt{\sigma^2 s_{kk}}}}{\sqrt{\frac{\hat{\sigma}^2}{\sigma^2}}} = \frac{\hat{\beta}_{kMCO} - \beta_k}{\sqrt{\hat{\sigma}^2 s_{kk}}}} \tag{3.17}$$

est distribuée suivant une loi de Student à $(N - K)$ degrés de liberté ($\tau(N - K)$). On note que $\sqrt{\hat{\sigma}^2 s_{kk}} = \sqrt{\widehat{\text{var}}(\hat{\beta}_{kMCO})}$ est l'écart type estimé de l'estimateur $\hat{\beta}_{kMCO}$.

2.2 Test de contraintes linéaires sur les paramètres du modèle

Les tests de contraintes permettent de tester des hypothèses économiques comme par exemple l'hypothèse de rendements d'échelle constants dans une fonction de production. Ils peuvent avoir pour but de simplifier le modèle en excluant par exemple des variables. On distingue les tests d'une seule contrainte et les tests de plusieurs contraintes. Les premiers utilisent une statistique distribuée selon une loi de Student, les seconds une statistique distribuée selon une loi de Fisher.

2.2.1 Test d'une contrainte

Le modèle ayant été estimé, on peut tester une contrainte sur les paramètres. Par exemple dans le modèle $y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i$ on peut tester l'hypothèse $\beta_2 + 0.5\beta_3 = 1$, qui peut s'écrire sous la forme $r' \beta = q$

$$\underbrace{(0 \quad 1 \quad 0.5)}_{r'} \underbrace{\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}}_{\beta} = \underbrace{1}_q$$

Plus généralement, une contrainte linéaire sur les paramètres d'un modèle de régression linéaire multiple peut s'écrire sous la forme

$$\begin{array}{l} H_0 : r_1\beta_1 + r_2\beta_2 + \dots + r_K\beta_K = r' \beta = q \\ H_1 : r_1\beta_1 + r_2\beta_2 + \dots + r_K\beta_K = r' \beta \neq q \end{array}$$

avec q une constante et $r' = (r_1, r_2, \dots, r_K)$.

Pour tester cette hypothèse, on estime le modèle par la méthode des MCO et on calcule la statistique $r' \hat{\beta}_{MCO}$. Sous l'hypothèse H_0 , son espérance et sa variance sont respectivement

$$\begin{aligned} E(r' \hat{\beta}_{MCO}) &= r' E(\hat{\beta}_{MCO}) = r' \beta = q \\ \text{Var}(r' \hat{\beta}_{MCO}) &= r' \text{Var}(\hat{\beta}_{MCO}) r = \sigma^2 r' (X' X)^{-1} r \end{aligned}$$

La statistique $r' \hat{\beta}_{MCO}$ étant une fonction linéaire de $\hat{\beta}$, elle est distribuée suivant une loi normale et, sous l'hypothèse H_0

$$\frac{r' \hat{\beta}_{MCO} - q}{\sqrt{\sigma^2 r' (X' X)^{-1} r}} \sim N(0,1)$$

Par ailleurs, la variance σ^2 étant inconnue, on peut diviser cette statistique par la racine carrée de $\frac{\hat{\sigma}^2}{\sigma^2}$ afin d'obtenir la valeur de la statistique de Student

$$t_{cal} = \frac{r' \widehat{\beta}_{MCO} - q}{\sqrt{\widehat{\sigma}^2 r' (X' X)^{-1} r}} = \frac{r' \widehat{\beta}_{MCO} - q}{\sqrt{\widehat{\sigma}^2}} \quad (3.18)$$

La statistique t_{cal} est distribuée selon une loi de Student à $N - K$ degrés de liberté sous l'hypothèse H_0 ($\tau(N - K)$).

La région critique du test est définie par $\Pr\{|t_{cal}| > \tau_{\alpha/2}(N - K)\} = \alpha$, avec α le risque de première espèce. Ce seuil de risque est généralement fixé à 5 % ou 10 %.

Ce test peut être utilisé pour tester la significativité de chacun des paramètres du modèle. Pour tester l'hypothèse $H_0 : \beta_k = 0$ contre $H_1 : \beta_k \neq 0$, on calcule la statistique

$$t_{cal} = \frac{\widehat{\beta}_{kMCO}}{\sqrt{\widehat{\sigma}^2 s_{kk}}}$$

Cette statistique est souvent notée *t-stat* dans les logiciels. Si $|t_{cal}| > \tau_{\alpha/2}(N - K)$ alors l'hypothèse H_0 est refusée au seuil α et le coefficient β_k peut être considéré significativement différent de 0.

2.2.2 Test de plusieurs contraintes

Le test précédent peut être généralisé à plusieurs contraintes. Il faut dans ce cas calculer une statistique de Fisher. Par exemple dans un modèle de régression linéaire multiple

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_u x_{ui} + \varepsilon_i$$

on peut souhaiter tester simultanément les deux hypothèses suivantes : $\beta_3 + \beta_4 = 1$ et $\beta_2 = 0$, qui peut s'écrire $R\beta = q$ avec

$$\underbrace{\begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}}_R \underbrace{\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}}_\beta = \underbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_q$$

Plus généralement, soit R une matrice d'ordre (J, K) , où J définit le nombre de contraintes, le problème consiste à réaliser le test formé par l'hypothèse

$$\begin{aligned} H_0 : R\beta &= q \\ H_1 : R\beta &\neq q \end{aligned}$$

avec q un vecteur composé de J lignes. Pour réaliser ce test on estime avec la méthode des MCO le modèle afin de calculer $R\hat{\beta}_{MCO}$. Cette statistique est une fonction linéaire de l'estimateur $\hat{\beta}_{MCO}$, elle est distribuée suivant une loi normale. Son espérance et sa variance sont données par les expressions suivantes sous l'hypothèse H_0

$$E(R\hat{\beta}_{MCO}) = RE(\hat{\beta}_{MCO}) = R\beta = q$$

et

$$Var(R\hat{\beta}_{MCO}) = RVar(\hat{\beta}_{MCO})R' = \sigma^2 R(X'X)^{-1}R'$$

Comme précédemment

$$R\hat{\beta}_{MCO} \sim N(q, \sigma^2 R(X'X)^{-1}R')$$

Ainsi, sous l'hypothèse H_0

$$\left(R\hat{\beta}_{MCO} - q\right)' \left(\sigma^2 R(X'X)^{-1}R'\right)^{-1} \left(R\hat{\beta}_{MCO} - q\right) \sim \chi^2(J)$$

Cette dernière statistique ne pouvant pas être calculée car la variance des aléas est inconnue, on détermine sous H_0 la statistique F_{cal} distribuée selon une loi de Fisher Snedecor

$$F_{cal} = \frac{\left(R\hat{\beta}_{MCO} - q\right)' \left(\sigma^2 R(X'X)^{-1}R'\right)^{-1} \left(R\hat{\beta}_{MCO} - q\right)}{\frac{J}{\hat{\sigma}^2}}$$

$$F_{cal} = \frac{\left(R\hat{\beta}_{MCO} - q\right)' \left(\hat{\sigma}^2 R(X'X)^{-1}R'\right)^{-1} \left(R\hat{\beta}_{MCO} - q\right)}{J} \quad (3.19)$$

$$\sim F(J, N - K)$$

On montre que cette statistique est équivalente à la suivante

$$F_{cal} = \frac{\frac{e'^* e^* - e' e}{J}}{\frac{e' e}{N - K}} \quad (3.20)$$

avec

- $(e'^* e^*)$ la somme des carrés des résidus du modèle estimé sous les contraintes données par $R\beta = q$,
- $(e' e)$ la somme des carrés des résidus du modèle estimé sans les contraintes.

En effet, soit $\hat{\beta}_{MCO}^*$ l'estimateur des MCO obtenu par la résolution du programme

$$\underset{\beta}{\text{Min}} (Y - X\beta)' (Y - X\beta) + 2\lambda' (R\beta - q)$$

avec λ le multiplicateur de Lagrange.

À partir des conditions du premier ordre on montre que

$$\hat{\beta}_{MCO}^* = \hat{\beta}_{MCO} - (X'X)^{-1}R' \left[R(X'X)^{-1}R' \right]^{-1} (R\hat{\beta}_{MCO} - q) \quad (3.21)$$

Cet estimateur est conditionnellement sans biais, il est non biaisé sous l'hypothèse H_0 et sa matrice de variance-covariance est donnée par l'expression

$$\begin{aligned} \text{Var}(\hat{\beta}_{MCO}^*) &= \sigma^2(X'X)^{-1} - \sigma^2(X'X)^{-1}R' \left[R(X'X)^{-1}R' \right]^{-1} R(X'X)^{-1} \\ &= \text{Var}(\hat{\beta}_{MCO}) - \text{matrice définie non négative} \end{aligned}$$

Les résidus du modèle estimé sous les contraintes ($R\beta = q$) s'écrivent

$$\begin{aligned} e^* &= Y - X\hat{\beta}_{MCO}^* = Y - X\hat{\beta}_{MCO} - X\hat{\beta}_{MCO}^* + X\hat{\beta}_{MCO} \\ &= e - X(\hat{\beta}_{MCO}^* - \hat{\beta}_{MCO}) \end{aligned}$$

Par conséquent

$$\begin{aligned} e^{*'}e^* &= \left[e - X(\hat{\beta}_{MCO}^* - \hat{\beta}_{MCO}) \right]' \left[e - X(\hat{\beta}_{MCO}^* - \hat{\beta}_{MCO}) \right] \\ &= e'e + (\hat{\beta}_{MCO}^* - \hat{\beta}_{MCO})' X'X(\hat{\beta}_{MCO}^* - \hat{\beta}_{MCO}) \end{aligned}$$

L'équation (3.21) donne $\hat{\beta}_{MCO}^* - \hat{\beta}_{MCO}$, on en déduit

$$e^{*'}e^* - e'e = \left(R\hat{\beta}_{MCO} - q \right)' \left(R(X'X)^{-1}R' \right)^{-1} \left(R\hat{\beta}_{MCO} - q \right)$$

Cette dernière expression est quasi identique au numérateur de l'équation (3.19). Par ailleurs, $\hat{\sigma}^2$ étant égal à $\frac{e'e}{N-K}$, les expressions (3.19) et (3.20) sont égales.

La statistique de Fisher (3.19) et (3.20) peut s'écrire également en termes de coefficient de détermination

$$F_{cal} = \frac{\frac{R^2 - R^{*2}}{1 - R^2} \frac{J}{N - K}}{\quad} \quad (3.22)$$

$$\text{avec } R^{*2} = 1 - \frac{e^{*'} e^*}{\sum (y_i - \bar{y})^2} \quad \text{et} \quad R^2 = 1 - \frac{e' e}{\sum (y_i - \bar{y})^2}$$

Par conséquent pour tester plusieurs contraintes simultanément les trois statistiques suivantes sont équivalentes

$$F_{cal} = \frac{(R\hat{\beta}_{MCO} - q)' (\hat{\sigma}^2 R(X'X)^{-1}R')^{-1} (R\hat{\beta}_{MCO} - q)}{J}$$

$$F_{cal} = \frac{\frac{e^{*'} e^* - e' e}{J}}{\frac{e' e}{N - K}}$$

$$F_{cal} = \frac{\frac{R^2 - R^{*2}}{1 - R^2} \frac{J}{N - K}}{\quad}$$

La région critique du test est donnée par $\Pr\{F_{cal} > F_{\alpha}(J, N - K)\} = \alpha$, avec α le risque de première espèce.

Ce test peut-être utilisé pour tester la significativité globale de la régression :

$$H_0 : \beta_2 = \beta_3 = \dots \beta_K = 0$$

H_1 : les coefficients ne sont pas tous simultanément nuls

Le coefficient de détermination du modèle estimé avec les $K - 1$ contraintes ($\beta_2 = 0, \dots, \beta_K = 0$) est nul, par conséquent sous H_0

$$F_{cal} = \frac{\frac{R^2}{1 - R^2} \frac{K - 1}{N - K}}{\quad}$$

et si F_{cal} est supérieur au $F_{\alpha}(K - 1, N - K)$, alors l'hypothèse H_0 est refusée.

2.3 Test de changement structurel : test de Chow

Il s'agit ici de tester une forme différente de contrainte sur les paramètres. Le test de Chow permet de déterminer si les paramètres du modèle se sont modifiés au cours du temps. Lorsqu'il est appliqué sur des données temporelles, il est nécessaire de connaître a priori une date de rupture. Le test indique alors si les paramètres sont identiques avant et après la date de rupture. Sur des données en coupe transversale, c'est-à-dire des données individuelles, ce test est utilisé afin de déterminer si des groupes d'individus sont homogènes ou hétérogènes.

On suppose que les aléas sont homoscedastiques³, c'est-à-dire qu'ils ont une variance identique pour chaque observation. Sur données temporelles, soient n_1 la date de rupture, β^1 le vecteur des paramètres du modèle sur la période $1, \dots, n_1$, β^2 le vecteur des paramètres sur la période $n_1 + 1, \dots, N$, on teste

$$\begin{aligned} H_0 : \beta^1 = \beta^2 = \beta \\ H_1 : \beta^1 \neq \beta^2 \end{aligned} \quad (3.23)$$

Sous l'hypothèse H_0 le modèle s'écrit

$$y_i = \beta_1 + \beta_2 x_{2i} + \dots + \beta_K x_{Ki} + \varepsilon_i = \beta' x_i + \varepsilon_i \quad i = 1, \dots, N$$

Sous l'hypothèse H_1 les paramètres sont différents selon la période considérée

$$y_i = \beta_1^1 + \beta_2^1 x_{2i} + \dots + \beta_K^1 x_{Ki} + \varepsilon_i = \beta^{1'} x_i + \varepsilon_i \quad i = 1, \dots, n_1$$

et

$$y_i = \beta_1^2 + \beta_2^2 x_{2i} + \dots + \beta_K^2 x_{Ki} + \varepsilon_i = \beta^{2'} x_i + \varepsilon_i \quad i = n_1 + 1, \dots, N$$

Le modèle est estimé par les MCO d'une part sur l'ensemble de la période et d'autre part sur les deux sous-périodes $1, \dots, n_1$ et $n_1 + 1, \dots, N$.

On note

- $e^* e^*$ la somme des carrés des résidus du modèle estimé avec les N observations (le degré de liberté correspondant est égal à $N - K$),
- $e^1 e^1$ la somme des carrés des résidus de l'estimation sur la période $1, \dots, n_1$ (le degré de liberté correspondant est égal à $n_1 - K$),
- $e^2 e^2$ la somme des carrés des résidus de l'estimation sur la période $n_1 + 1, \dots, N$ (le degré de liberté correspondant est égal à $N - n_1 - K$)
- $e' e$ la somme de $e^1 e^1$ et $e^2 e^2$ (le degré de liberté correspondant est égal à $N - 2K$).

³ Pour plus de détails lorsque les aléas sont hétéroscedastiques, le lecteur peut se reporter à B. Dormont, 1999, *Introduction à l'Économétrie*, Montchrestien.

Ce test correspond à un test de K contraintes, sous l'hypothèse nulle

$$\begin{aligned}\beta_1^1 &= \beta_1^2 = \beta_1, \\ \beta_2^1 &= \beta_2^2 = \beta_2, \\ &\vdots \\ &\vdots \\ &\vdots \\ \beta_K^1 &= \beta_K^2 = \beta_K\end{aligned}$$

et la statistique de Fisher-Snédecor s'écrit

$$F_{cal} = \frac{\frac{e^{*'} e^* - e' e}{(N - K) - [(n_1 - K) + (N - n_1 - K)]}}{\frac{e' e}{(n_1 - K) + (N - n_1 - K)}}$$

$$F_{cal} = \frac{\frac{e^{*'} e^* - e' e}{K}}{\frac{e' e}{N - 2K}} \quad (3.24)$$

Cette statistique est distribuée suivant une loi de Fisher-Snédecor à K et $N - 2K$ degrés de liberté ($F(K, N - 2K)$).

Si $F_{cal} < F_\alpha(K, N - 2K)$ alors l'hypothèse H_0 n'est pas rejetée et on retient une estimation par les MCO avec l'ensemble des observations. Le test s'applique de manière identique sur données individuelles pour tester l'homogénéité du comportement de groupes d'individus.

2.4 Tests de stabilité temporelle basés sur les résidus récurrents

Le test précédent suppose la date de rupture n_1 connue a priori. Mais il existe également des tests de stabilité temporelle qui permettent de déterminer les dates de rupture. Ces tests sont basés sur l'utilisation des résidus récurrents. Dans le modèle de régression linéaire on suppose que les paramètres (β, σ^2) ne varient pas au cours du temps

$$\begin{aligned}y_i &= x_i' \beta + \varepsilon_i \quad i = 1, \dots, N \\ \text{var}(\varepsilon_i) &= \sigma^2\end{aligned}$$

on suppose

$$\begin{aligned}\beta^1 &= \dots = \beta^t \dots = \beta^N = \beta \\ \sigma^{21} &= \dots = \sigma^{2t} = \dots = \sigma^{2N} = \sigma^2\end{aligned}$$

avec (β^t, σ^{2t}) les paramètres du modèle à la période t . Les tests de stabilité temporelle développés par Brown, Durbin et Evans (1975) testent l'hypothèse nulle que le vecteur β est identique à chaque période conditionnellement à l'hypothèse $\sigma^{2_1} = \dots = \sigma^{2_t} = \dots = \sigma^{2_N} = \sigma^2$

$$\begin{aligned} H_0 : \beta^1 = \dots = \beta^t \dots = \beta^N = \beta \\ H_1 : \beta^t \neq \beta \quad \text{ou} \quad \sigma^{2_t} = \sigma^2 \end{aligned} \tag{3.25}$$

Il existe notamment deux tests très souvent utilisés, le test du CUSUM et le test du CUSUM of Square. Ces deux tests utilisent les résidus récurrents.

2.4.1 Les résidus récurrents

Pour calculer les résidus récurrents du modèle, on doit estimer par la méthode des MCO les paramètres β avec un nombre d'observations qui varie de K à N . Soient

$$Y^r = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_r \end{pmatrix} \quad X^r = \begin{pmatrix} 1 & x_{21} & \cdot & \cdot & x_{K1} \\ 1 & x_{22} & \cdot & \cdot & x_{K2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{2r} & \cdot & \cdot & x_{Kr} \end{pmatrix} \quad \varepsilon^r = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \varepsilon_r \end{pmatrix}$$

avec r le nombre d'observations, $r = K+1, \dots, N$. Soit e_r l'erreur de prévision ex-post sur l'observation y_r

$$e_r = y_r - \hat{y}_r = y_r - x_r' \hat{\beta}_{MCO}^{r-1}$$

avec $\hat{\beta}_{MCO}^{r-1}$ l'estimateur des MCO lorsque la taille de l'échantillon est égale à $r - 1$.

Sous l'hypothèse de stabilité $y_r = x_r' \beta + \varepsilon_r$ et l'erreur de prévision e_r s'écrit

$$\begin{aligned} e_r &= x_r' \beta + \varepsilon_r - x_r' \hat{\beta}_{MCO}^{r-1} \\ &= \varepsilon_r - x_r' \left(\hat{\beta}_{MCO}^{r-1} - \beta \right) \end{aligned}$$

Or l'espérance et la matrice de variance-covariance de $\hat{\beta}_{MCO}^{r-1}$ sont données par

$$\begin{aligned} E \left(\hat{\beta}_{MCO}^{r-1} \right) &= \beta \\ Var \left(\hat{\beta}_{MCO}^{r-1} \right) &= E \left(\left(\hat{\beta}_{MCO}^{r-1} - \beta \right) \left(\hat{\beta}_{MCO}^{r-1} - \beta \right)' \right) = \sigma^2 \left(X_{r-1}' X_{r-1} \right)^{-1} \end{aligned}$$

d'où

$$\begin{aligned} E(e_r) &= 0 \\ var(e_r) &= \sigma_{e_r}^2 = \sigma^2 \left(1 + x_r' \left(X_{r-1}' X_{r-1} \right)^{-1} x_r \right) \end{aligned}$$

Les résidus récurrents notés w_r sont donnés par

$$w_r = \frac{e_r}{\sqrt{1 + x_r' (X_{r-1}' X_{r-1})^{-1} x_r}} \quad (3.26)$$

Sous l'hypothèse de stabilité, les résidus récurrents sont distribués suivant une loi normale

$$w_r \sim N(0, \sigma^2)$$

2.4.2 Test du CUSUM

Le test du CUSUM consiste à calculer la série cumulée des w_r

$$W_r = \sum_{j=K+1}^r \frac{w_j}{\hat{\sigma}} \quad r = K + 1, \dots, N \quad (3.27)$$

avec

$$\hat{\sigma}^2 = \frac{\sum_{r=K+1}^N (w_r - \bar{w})^2}{N - K - 1} \quad \text{et} \quad \bar{w} = \frac{\sum_{r=K+1}^N w_r}{N - K}$$

Sous l'hypothèse de stabilité, W_r a une moyenne nulle et la région de confiance du test est donnée par

$$\Pr[-C_\alpha \leq W_r \leq C_\alpha] = 1 - \alpha$$

avec $C_\alpha = a(N - K)^{1/2} + 2a(r - K)(N - K)^{-1/2}$. La valeur de a dépend du risque de première espèce α du test. Pour $\alpha = 0.01, 0.05, 0.10$, a est égal à 1.143, 0.948, 0.850 respectivement.

Afin de faciliter la lecture des résultats du test effectué pour chaque statistique W_r ($r = K + 1, \dots, N$), on représente graphiquement la série W_r et la région de confiance donnée par l'intervalle $[-C_\alpha, C_\alpha]$.

Pour $r = K$, $C_\alpha = a(N - K)^{1/2}$ et pour $r = N$, $C_\alpha = 3a(N - K)^{1/2}$. La région de confiance du test pour les points $r = K + 1, \dots, N$ est représentée par la droite reliant ces deux points.

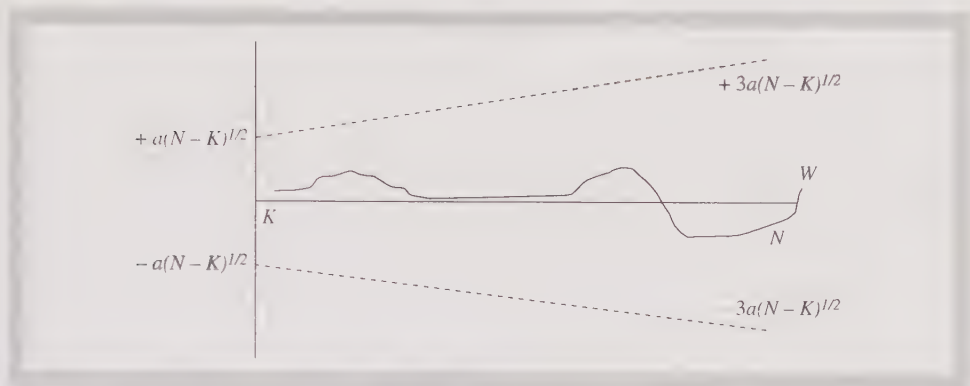


Figure 3.1 Test du CUSUM

Dans la figure 3.1, l'espace entre les deux droites donne la région de confiance du test du CUSUM. Si W_r reste dans cet intervalle quel que soit $r = K + 1, \dots, N$, comme sur le graphique, alors l'hypothèse de stabilité est retenue. Sinon, si à une date donnée, W_r va au delà d'une des droites, alors on considère qu'à cette date il y a une rupture et que les paramètres ne sont pas stables.

2.4.3 Test du CUSUM of Square

Le test du CUSUM of Square est similaire au précédent. Il consiste à représenter graphiquement la série S_r

$$S_r = \frac{\sum_{j=K+1}^r w_j^2}{\sum_{j=K+1}^N w_j^2} \quad r = K + 1, \dots, N \tag{3.28}$$

On peut montrer que sous l'hypothèse de stabilité, S_r a une distribution Bêta d'espérance

$$E(S_r) = \frac{r - K}{N - K}$$

Comme dans le test du CUSUM, la région critique du test peut-être représentée graphiquement. Les deux bornes sont données par $E(S_r) \pm c_0$, où c_0 ⁴ est déterminé pour un niveau de significativité en fonction de K et N .

⁴ Les valeurs de c_0 sont données dans la table de Durbin et peuvent être trouvées dans Harvey, 1999, *The Econometric Analysis of Time Series*, 2nd edition, Cambridge, Mass., MIT Press.

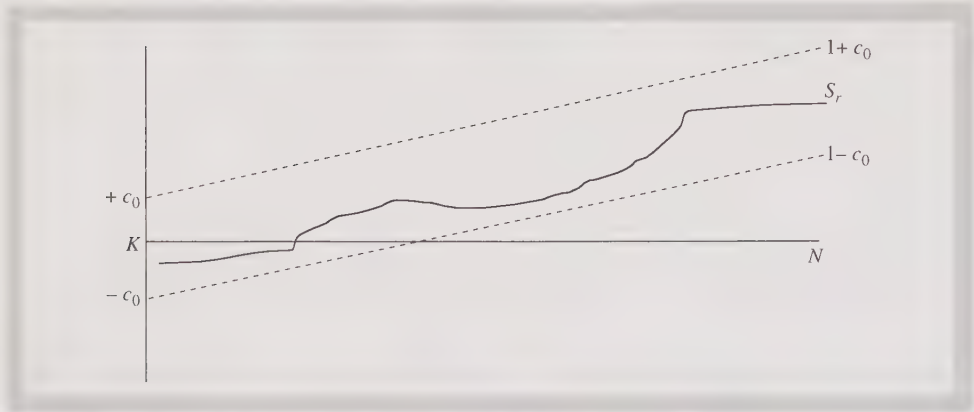


Figure 3.2 Test du CUSUM of Square

Pour $r = K$, la région d'acceptation du test est donnée par $\pm c_0$ et pour $r = N$, elle est donnée par $1 \pm c_0$.

Si quel que soit $r = K + 1, \dots, N$, S_r se situe entre les droites $E(S_r) \pm c_0$ alors les paramètres du modèle sont supposés stables.

3. PRÉDICTIONS ET QUALITÉ PRÉDICTIVE D'UN MODÈLE ESTIMÉ

Un modèle dont les paramètres ont été estimés peut être utilisé afin de réaliser des prédictions. On se pose alors la question de savoir si ces prédictions sont de bonnes qualités. Plusieurs méthodes, non exclusives, peuvent être appliquées pour apprécier la qualité prédictive d'un modèle estimé.

1. À partir de l'équation de la droite de régression de l'échantillon, il est possible de fournir des prédictions. Pour une valeur du vecteur x donnée, notée x_0 , on peut calculer la valeur estimée \hat{y}_0

$$\begin{aligned}\hat{y}_0 &= \hat{\beta}_{1MCO} + \hat{\beta}_{2MCO}x_{2,0} + \dots + \hat{\beta}_{KMCO}x_{K,0} \\ &= x_0' \hat{\beta}_{MCO}\end{aligned}\quad (3.29)$$

L'erreur de prévision est donnée par l'écart entre la valeur observée y_0 et \hat{y}_0

$$e_0 = y_0 - \hat{y}_0 = \varepsilon_0 + x_0' (\beta - \hat{\beta}_{MCO}) \quad (3.30)$$

e_0 est une variable aléatoire car la réalisation y_0 n'est pas connue au moment de la prévision. Les moments de cette variable aléatoire sont

$$E(e_0) = 0$$

et

$$\text{Var}(e_0) = E(e_0^2) = \sigma^2 \left(1 + x_0'(X'X)^{-1}x_0\right)$$

e_0 est une variable aléatoire distribuée selon une loi normale, par conséquent

$$\frac{e_0}{\sqrt{\sigma^2 \left(1 + x_0'(X'X)^{-1}x_0\right)}} \sim N(0,1)$$

et

$$\frac{e_0}{\sqrt{\hat{\sigma}^2 \left(1 + x_0'(X'X)^{-1}x_0\right)}} = \frac{y_0 - \hat{y}_0}{\hat{\sigma}_{e_0}} \sim \tau(N - K)$$

On peut ainsi établir l'intervalle de confiance d'une prédiction de y_0 , pour un vecteur d'observations x_0 donné

$$\Pr \left[\hat{y}_0 - \hat{\sigma}_{e_0} \tau_{\alpha/2}(N - K) \leq y_0 \leq \hat{y}_0 + \hat{\sigma}_{e_0} \tau_{\alpha/2}(N - K) \right] = 1 - \alpha \quad (3.31)$$

2. Pour évaluer la qualité prédictive d'un modèle, on peut aussi comparer les valeurs observées et les valeurs estimées, c'est-à-dire s'intéresser à la série des résidus. La série des valeurs estimées doit bien reproduire l'évolution de la série observée et, en particulier, les points de retournement (les changements dans l'évolution de la série).
3. Pour tester la qualité prédictive du modèle on peut également utiliser la statistique de Chow. Pour mettre en œuvre ce test on procède de la manière suivante
 - a) On estime le modèle sur l'ensemble de la période, puis on extrait de cette estimation la somme des carrés des résidus notée $e^* e^*$ associée à un nombre de degrés de liberté $N - K$.
 - b) On estime le modèle sur n_1 observations avec $n_1 < N$, on extrait de cette estimation la somme des carrés des résidus notée $e' e$ associée à un nombre de degrés de liberté $n_1 - K$.
 - c) On construit un test de Fisher

$$F_{cal} = \frac{\frac{e^* e^* - e' e}{(N - K) - (n_1 - K)}}{\frac{e' e}{n_1 - K}}$$

$$F_{cal} = \frac{\frac{e^* e^* - e' e}{N - n_1}}{\frac{e' e}{n_1 - K}} \sim F(N - n_1, n_1 - K) \quad (3.32)$$

si $F_{cal} < F(N - n_1, n_1 - K)$ alors la capacité prédictive du modèle est considérée bonne.

4. Enfin on peut calculer différentes statistiques permettant de juger de la qualité des prévisions. On estime le modèle sur un nombre d'observations $n_1 < N$ et on effectue des prévisions sur n observations ($n = N - n_1$), puis on calcule les statistiques

Mean Error

$$\text{ME} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i) \quad (3.33)$$

Mean Absolute Error

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3.34)$$

Root Mean Square Error

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3.35)$$

Lorsqu'on compare deux modèles, le modèle pour lequel les trois statistiques sont les plus faibles est considéré avoir une meilleure capacité prédictive. Si la statistique ME est négative alors la variable endogène est globalement surestimée. Lorsqu'elle est surestimée, à chaque observation, les statistiques MAE et ME sont identiques. Enfin si les erreurs sont du même ordre de grandeur, les statistiques MAE et RMSE prennent la même valeur.

On peut également calculer la statistique de Theil qui permet de comparer les prévisions du modèle à des prévisions naïves.

$$\text{Theil} = \sqrt{\frac{\sum_{i=1}^n (\Delta y_i - \Delta \hat{y}_i)^2}{\sum_{i=1}^n (\Delta y_i)^2}} \quad (3.36)$$

Si on réalise des prévisions naïves $\hat{y}_i = y_{i-1}$ et $\Delta \hat{y}_i = 0$, par conséquent la statistique de Theil est égale à 1. Ainsi, une valeur de la statistique de Theil inférieure à 1 indique que le modèle permet de réaliser des prévisions meilleures que des prévisions naïves.

4. LES VARIABLES INDICATRICES DANS LES MODÈLES ÉCONOMÉTRIQUES

Pour terminer ce chapitre, on examine des modèles dont certaines variables explicatives sont des variables indicatrices. Ces variables sont très fréquemment introduites dans les modèles économétriques. L'estimation des paramètres de ces modèles peut être effectuée par les MCO et les tests présentés dans la section précédente sont valides. L'objectif de cette section est d'indiquer l'utilité des variables indicatrices. Elles prennent usuellement deux valeurs 0 ou 1, elles permettent de modifier, pour certaines observations, la constante β_1 et/ou les paramètres $\beta_k, k = 2, \dots, K$. Ici, trois exemples d'utilisation sont présentés.

1. Dans un modèle estimé en coupe longitudinale, ce type de variable permet de prendre explicitement en compte des changements temporels. Par exemple, si on observe la relation suivante entre des dépenses de consommation (C) et le revenu (R) au cours d'une période donnée ($t = 1, \dots, T$)

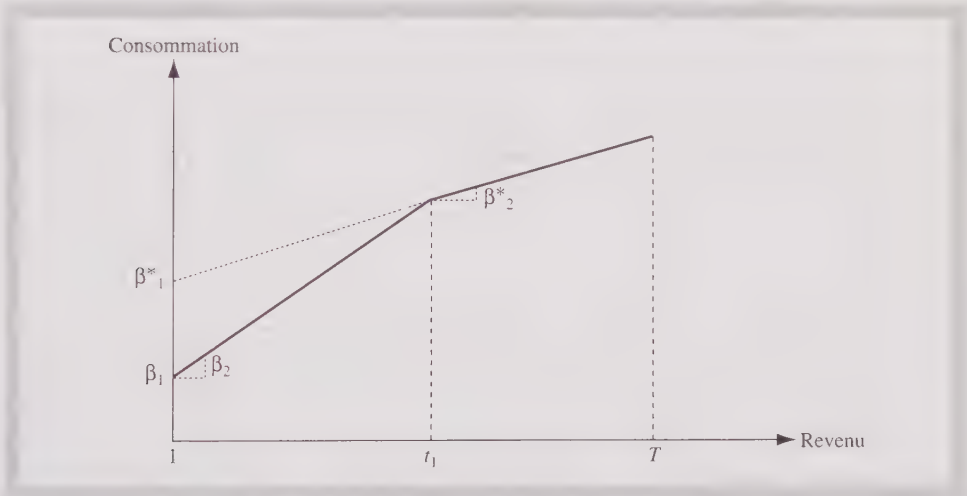


Figure 3.3 Changement structurel dans la relation consommation-revenu

Le modèle s'écrit

$$\begin{aligned} C_t &= \beta_1 + \beta_2 R_t + \varepsilon_t, & t = 1, \dots, t_1 \\ C_t &= \beta_1^* + \beta_2^* R_t + \varepsilon_t, & t = t_1 + 1, \dots, T \end{aligned} \quad (3.37)$$

Soit on estime le modèle sur chaque sous-période (3.37), soit on l'estime sur l'ensemble de la période en introduisant une variable indicatrice D qui prend la valeur 1 pour $t = t_1 + 1, \dots, T$ et 0 sinon (3.38).

Le modèle estimé avec l'ensemble des observations est donné par

$$C_t = \beta_1 + \psi_1 D_t + \beta_2 R_t + \psi_2 (D_t \times R_t) + \varepsilon_t, \quad t = 1, \dots, T \quad (3.38)$$

avec $\beta_1^* = \beta_1 + \psi_1$ et $\beta_2^* = \beta_2 + \psi_2$. Sous forme matricielle le modèle s'écrit

$$\begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_{t_1} \\ C_{t_1+1} \\ \vdots \\ C_T \end{bmatrix} = \begin{bmatrix} 1 & 0 & R_1 & 0 \\ 1 & 0 & R_2 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & R_{t_1} & 0 \\ 1 & 1 & R_{t_1+1} & R_{t_1+1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & R_T & R_T \end{bmatrix} \begin{bmatrix} \beta_1 \\ \psi_1 \\ \beta_2 \\ \psi_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{t_1} \\ \varepsilon_{t_1+1} \\ \vdots \\ \varepsilon_T \end{bmatrix}$$

2. Les variables indicatrices peuvent être utilisées pour tester l'homogénéité du comportement d'un groupe d'individus par rapport à l'ensemble des paramètres ou à un sous-ensemble de paramètres. Par exemple, soit C la consommation d'essence des ménages, R le revenu des ménages et Z un autre facteur explicatif. Ces variables sont observées pour un ensemble de N individus et le modèle estimé s'écrit

$$\log C_i = \beta_1 + \beta_2 \log R_i + \beta_3 \log Z_i + \varepsilon_i \quad i = 1, \dots, N$$

Pour répondre à la question, est-ce que l'élasticité revenu est identique lorsque les ménages habitent en zone rurale et urbaine, on peut créer une variable D qui prend la valeur 1 lorsque le ménage habite dans une zone rurale et 0 sinon, puis estimer la relation

$$\log C_i = \beta_1 + \beta_2 \log R_i + \beta_3 \log Z_i + \beta_4 (D_i \times \log R_i) + \varepsilon_i \quad i = 1, \dots, N$$

Si β_4 est significativement non nul alors l'élasticité revenu des ménages habitant en zone rurale est égale à $\hat{\beta}_2 + \hat{\beta}_4$ et l'élasticité revenu des ménages habitant en zone urbaine est donnée par $\hat{\beta}_2$. Sinon, les ménages ont une élasticité revenu égale à $\hat{\beta}_2$ quelle que soit la zone d'habitation.

3. Pour tenir compte de la saisonnalité des données on peut également utiliser des variables indicatrices. Par exemple, sur des données trimestrielles, afin de distinguer chacun des trimestres, on peut créer 4 variables indicatrices D_1, D_2, D_3, D_4 qui prennent respectivement la valeur 1 le premier, le second, le troisième et le quatrième trimestre et 0 sinon. Ainsi, pour chaque observation la somme de ces quatre variables observées $D_1 + D_2 + D_3 + D_4$ sera égale à 1. Cependant, si le modèle inclut une constante, la matrice X des variables explicatives ne sera pas de plein rang et la matrice $X'X$ sera singulière. Pour estimer les paramètres du modèle, il faut alors enlever dans la matrice X , soit la constante, soit une variable indicatrice. Les résultats s'interprètent relativement à la variable omise dans le modèle estimé.

4

CONVERGENCE DU PIB PAR TÊTE DES ÉCONOMIES EUROPÉENNES

1. Introduction
2. Le modèle de Solow
3. Analyse empirique
4. Éléments de corrigé
5. Annexes

OBJECTIFS DU CHAPITRE

D'un point de vue méthodologique

- Estimer un modèle de régression linéaire multiple
- Pratiquer des tests statistiques

D'un point de vue économique

- Présenter le modèle de croissance de Solow
- Tester la convergence du revenu par tête des pays européens

NIVEAU

LICENCE, troisième année

CHAMP ÉCONOMIQUE TRAITÉ

Macroéconomie réelle

1. INTRODUCTION

Le 1^{er} janvier 1999, onze pays (l'Allemagne, l'Autriche, la Belgique, l'Espagne, la Finlande, la France, l'Irlande, l'Italie, le Luxembourg, les Pays-Bas et le Portugal) ont adhéré à l'Union Economique et Monétaire (UEM). L'adhésion de ces pays a été conditionnée au respect des critères de convergence définis dans le cadre du traité de Maastricht.

Les 11 pays ont constitué une zone monétaire et la politique monétaire de l'ensemble est dirigée par la Banque Centrale Européenne. Le but de l'UEM est de constituer une zone monétaire optimale. Dans une zone monétaire, les pays ne sont plus maîtres de leur politique monétaire et ne disposent plus de l'instrument du taux de change pour amortir l'impact de chocs macroéconomiques. En contrepartie, les coûts liés aux transactions sont réduits, cette réduction de coûts étant d'autant plus forte que les échanges intra zone sont élevés. Une zone monétaire est qualifiée d'optimale si le recours à la monnaie commune ne provoque aucune perte de bien-être économique global.

Les travaux empiriques centrés sur la constitution de l'UEM et plus généralement, sur l'Union Européenne (UE), cherchent à déterminer le degré de ressemblance des pays de l'UE car, des pays à structure économique proche auront des réactions de même ampleur face à des chocs exogènes. La création de l'UE a ainsi suscité de nombreuses études sur la convergence des variables macroéconomiques telles que l'inflation, le taux de change, le revenu par tête, etc. Ce chapitre traite de la convergence du revenu par tête des pays de l'UE. Il s'agit de déterminer si, à long terme, les économies européennes convergent vers un revenu par tête identique. Autrement dit, parmi les pays européens, les pays initialement les plus pauvres rattrapent-ils les pays les plus riches.

Cette question est usuellement traitée à partir de l'estimation d'une équation de convergence dérivée du modèle de croissance de Solow. Dans un premier temps, on présente ce modèle dans lequel la productivité marginale des facteurs est supposée décroissante¹. Cette hypothèse implique qu'une économie atteint à long terme un niveau de production par tête stable. En d'autres termes, elle converge vers un état stationnaire, c'est-à-dire vers un état caractérisé par un taux de croissance de la production par tête constant. En outre, des économies identiques convergent vers le même état stationnaire. Dans un second temps, les résultats théoriques du modèle de Solow sont testés dans le cadre des pays membres de l'UE en 2000 (l'analyse empirique porte sur les années 1970-2000).

2. LE MODÈLE DE SOLOW

Le modèle de croissance de Solow constitue un des modèles de base de l'analyse de la croissance économique de long terme². Il correspond au cas d'une

¹ Autrement dit, un accroissement de l'utilisation d'un facteur, toutes choses égales par ailleurs, induit une hausse de la production mais cette augmentation devient de plus en plus faible au fur et à mesure que la quantité de facteur s'accroît.

² Pour une présentation détaillée des modèles de croissance, le lecteur peut se reporter à : Barro R.J., Sala-I-Martin X., 1996, *La croissance économique*, McGraw-Hill/Ediscience.

économie fermée (la production est égale au revenu et l'investissement à l'épargne) avec un taux d'épargne exogène.

2.1 Structure du modèle

Les variables endogènes sont : la production réelle (Y), l'investissement (I), la consommation (C), le stock de capital (K). Les variables exogènes sont l'emploi total (N) et l'indice d'efficacité de l'input travail (A).

La structure théorique de la version du modèle de Solow retenue dans ce chapitre est la suivante

$$Y(t) = C(t) + I(t) \quad (4.1)$$

$$I(t) = sY(t) \quad (4.2)$$

$$\dot{K}(t) = I(t) - \delta K(t) \quad (4.3)$$

$$Y(t) = F(K(t), A(t)N(t)) \quad (4.4)$$

$$\frac{\dot{N}(t)}{N(t)} = n \iff N(t) = N(0)e^{nt} \quad (4.5)$$

$$\frac{\dot{A}(t)}{A(t)} = a \iff A(t) = A(0)e^{at} \quad (4.6)$$

L'équation (4.1) représente l'équilibre sur le marché des biens. L'équation (4.2) indique que l'investissement est une fraction constante de la production, s étant la propension à épargner. La dynamique d'accumulation du stock de capital est donnée par (4.3), qui montre que la variation du stock de capital au cours du temps³ est égale à l'investissement net de la dépréciation du capital, δ représentant le taux de dépréciation du stock de capital. La fonction de production est définie par l'équation (4.4). Elle vérifie les hypothèses suivantes :

- Lorsque la quantité d'un input augmente la production s'accroît mais cet accroissement diminue au fur et à mesure que la quantité du facteur s'élève. Autrement dit, la productivité marginale des facteurs de production est positive⁴ et décroissante⁵.
- La productivité marginale des facteurs de production décroît de $+\infty$ à 0.

$$\lim_{K \rightarrow 0} \left(\frac{\partial F}{\partial K} \right) = \lim_{N \rightarrow 0} \left(\frac{\partial F}{\partial N} \right) = +\infty \quad \text{et} \quad \lim_{K \rightarrow +\infty} \left(\frac{\partial F}{\partial K} \right) = \lim_{N \rightarrow +\infty} \left(\frac{\partial F}{\partial N} \right) = 0$$

³ $\dot{K}(t) = \frac{\partial K(t)}{\partial t}$

⁴ $\frac{\partial F}{\partial K} > 0, \frac{\partial F}{\partial N} > 0$

⁵ $\frac{\partial^2 F(K, AN)}{\partial K^2} < 0, \frac{\partial^2 F(K, AN)}{\partial N^2} < 0$

- Lorsque les facteurs de production augmentent dans une même proportion, la production s'accroît de cette même proportion. Ainsi si les quantités des deux facteurs sont doublées, la production est également doublée. La fonction de production est donc à rendements d'échelle constants.

$$F(\lambda K, \lambda AN) = \lambda F(K, AN), \lambda > 0.$$

Enfin, les équations (4.5) et (4.6) indiquent que l'input travail et le progrès technologique augmentent respectivement aux taux exogènes n et a^6 . $N(0)$ et $A(0)$ définissent respectivement les valeurs initiales du niveau de l'emploi et de l'indice d'efficacité du travail.

Ce modèle est dynamique. Il décrit le comportement d'une économie qui converge vers un équilibre macroéconomique de long terme stable. En d'autres termes l'économie considérée retourne toujours vers son équilibre à une certaine vitesse lorsqu'un choc exogène l'en éloigne. Dans ce cadre d'analyse, les concepts d'équilibre de long terme et de convergence sont étroitement liés. On assiste au cours du temps à une succession d'équilibres de court terme et la dynamique d'accumulation du capital donnée par l'équation (4.3) conduit l'économie à converger vers son équilibre de long terme.

2.2 Analyse de l'équilibre de long terme du modèle

Dans ce paragraphe, on analyse l'évolution de l'économie au cours du temps afin de caractériser son équilibre de long terme. Pour simplifier la présentation, les variables sont exprimées en unité de travail effectif. Chaque variable est donc divisée par la quantité AN . On note $\hat{k} = \frac{K}{AN}$ le stock de capital par unité de travail effectif, et $\hat{y} = \frac{Y}{AN}$ la production par unité de travail effectif. La fonction de production $Y = F(K, AN)$ étant à rendements d'échelle constants, la fonction de production par unité de travail effectif s'écrit

$$\hat{y} = F\left(\frac{K}{AN}, 1\right) = f(\hat{k}) \quad (4.7)$$

Les propriétés de la fonction $f(\hat{k})$ se déduisent de celles de la fonction $F(K, AN)$: $f(\hat{k})$ est croissante⁷ et concave en \hat{k} ⁸.

En divisant par la quantité AN l'équation (4.3) et en utilisant l'équation (4.2) on détermine l'évolution au cours du temps du stock de capital par unité de travail effectif

$$\frac{\dot{K}}{AN} = \frac{I}{AN} - \delta \frac{K}{AN} = sf(\hat{k}) - \delta \hat{k} \quad (4.8)$$

⁶ $\dot{N}(t) = \frac{\partial N(t)}{\partial t}$ et $\dot{A}(t) = \frac{\partial A(t)}{\partial t}$

⁷ $\frac{\partial f(\hat{k})}{\partial \hat{k}} > 0$

⁸ $\frac{\partial^2 f(\hat{k})}{\partial \hat{k}^2} < 0$

Or

$$\dot{\widehat{k}} = \left(\frac{\dot{K}}{AN} \right) = \frac{\dot{K} AN - K(\dot{A} N + \dot{N} A)}{(AN)^2} \quad (4.9)$$

$$= \frac{\dot{K}}{AN} - \frac{K}{AN} \left[\frac{\dot{A}}{A} + \frac{\dot{N}}{N} \right] = \frac{\dot{K}}{AN} - \widehat{k}(a+n) \quad (4.10)$$

L'évolution au cours du temps du stock de capital par unité de travail effectif peut finalement être décrit par l'équation (4.11)

$$\dot{\widehat{k}} = sf(\widehat{k}) - \widehat{k}(a+n+\delta) \quad (4.11)$$

Cette équation est l'équation différentielle de base du modèle de Solow. Le terme $sf(\widehat{k})$ représente l'investissement par unité de travail effectif, et le terme $(a+n+\delta)$ peut s'interpréter comme le taux effectif de dépréciation du stock de capital par unité de travail effectif.

La figure 4.1 donne une représentation du modèle de Solow. La courbe d'investissement par unité de travail effectif $sf(\widehat{k})$ est proportionnelle à la courbe de production $f(\widehat{k})$. La consommation effective par tête C/AN est donnée par la distance entre les deux courbes. La droite d'équation $\widehat{k}(a+n+\delta)$, passant par l'origine, donne la dépréciation de \widehat{k} . La distance verticale entre les courbes $sf(\widehat{k})$ et $\widehat{k}(a+n+\delta)$ correspond à la variation du stock de capital par unité de travail effectif. Par exemple, une économie qui se situe au point \widehat{k}_0 , où $sf(\widehat{k}_0)$ est plus élevée que $\widehat{k}_0(a+n+\delta)$, c'est-à-dire où la variation instantanée de \widehat{k} est positive ($\dot{\widehat{k}} > 0$), va voir son stock de capital \widehat{k} s'accroître jusqu'au niveau de stock de capital par unité de travail effectif \widehat{k}^* . En ce point, $sf(\widehat{k}^*) = \widehat{k}^*(a+n+\delta)$ et $\dot{\widehat{k}}^* = 0$. Le stock de capital par unité de travail effectif ne varie plus lorsqu'il a atteint le niveau d'équilibre \widehat{k}^* . Ce point d'équilibre correspond à l'état stationnaire de l'économie⁹. De la même manière une économie disposant à une période donnée d'un stock de capital par unité de travail effectif plus élevé que \widehat{k}^* va voir son stock de capital \widehat{k} diminuer jusqu'au point d'équilibre \widehat{k}^* . Par conséquent, quel que soit le niveau initial du stock de capital par unité de travail effectif, l'économie décrite par le modèle de Solow converge à long terme vers un équilibre stable \widehat{k}^* .

⁹ Un état stationnaire correspond à une situation dans laquelle l'économie croît à un taux constant. Dans le modèle de Solow, cet état est atteint lorsque le stock de capital par unité de travail effectif ne varie plus.

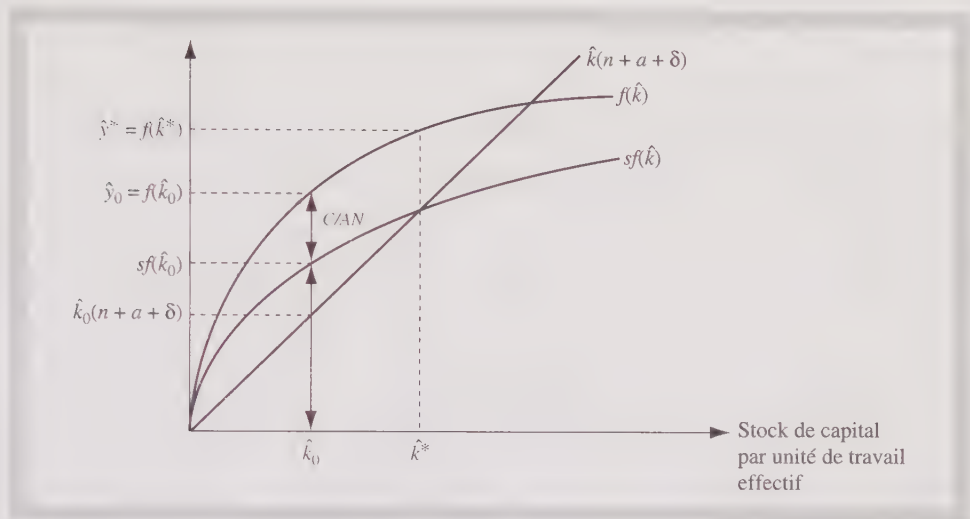


Figure 4.1 Modèle de Solow

Lorsque le taux d'épargne augmente, la courbe $sf(\widehat{k})$ se déplace vers le haut permettant au niveau d'équilibre d'augmenter. Par contre, une augmentation de la dépréciation du stock de capital par unité de travail effectif induit un déplacement vers le haut de la droite $\widehat{k}(a+n+\delta)$, et donc une baisse de \widehat{k}^* .

Au stock de capital par unité de travail effectif d'équilibre \widehat{k}^* , correspond un niveau de production par unité de travail effectif $\widehat{y}^* = f(\widehat{k}^*)$, un niveau d'investissement par unité de travail effectif $sf(\widehat{k}^*)$, et un niveau de consommation par unité de travail effectif correspondant à l'écart entre $f(\widehat{k}^*)$ et $sf(\widehat{k}^*)$.

Cet équilibre est caractérisé par un taux de croissance du stock de capital par unité de travail effectif nul, $\gamma_{\widehat{k}^*} = \dot{\widehat{k}^*} / \widehat{k}^* = 0$ car $\dot{\widehat{k}^*} = 0$, mais par un taux de croissance positif et constant du stock de capital d'équilibre égal à $a+n$ (cf. annexe 1 à la fin du chapitre).

Deux résultats fondamentaux sont déduits de cette analyse du modèle de Solow. D'une part, une économie, quel que soit sa situation initiale, converge à long terme vers un équilibre stable. D'autre part, une économie à l'équilibre est caractérisée par un stock de capital K^* qui croît à un taux constant égal à $(a+n)$ somme du taux de croissance de progrès technique et du taux de croissance de la population. Dans le modèle de Solow, ces deux taux sont supposés exogènes. L'économie à l'équilibre croît donc de manière constante et exogène.

2.3 Dynamique d'ajustement vers l'équilibre

Dans cette section, on examine d'abord le processus d'ajustement de l'économie vers son équilibre de revenu par tête et, ensuite, les conclusions du modèle de

Solow en termes de convergence de plusieurs économies vers un équilibre stable identique.

2.3.1 L'ajustement vers l'équilibre

Le taux de croissance du stock de capital par unité de travail effectif se déduit de (4.11)

$$\gamma_{\hat{k}} = \frac{\dot{\hat{k}}}{\hat{k}} = \frac{sf(\hat{k})}{\hat{k}} - (a + n + \delta) \quad (4.12)$$

Compte tenu des hypothèses adoptées pour la fonction de production, la fonction de production moyenne $f(\hat{k})/\hat{k}$ est une fonction décroissante du stock de capital par unité de travail effectif \hat{k} et elle varie de 0 (pour $\hat{k} \rightarrow +\infty$) à $+\infty$ (pour $\hat{k} \rightarrow 0$)

$$\left[\frac{\partial (f(\hat{k})/\hat{k})}{\partial \hat{k}} \right] - \left[\frac{f(\hat{k}) - \hat{k}f'(\hat{k})}{\hat{k}^2} \right] < 0, \quad \lim_{\hat{k} \rightarrow +\infty} \frac{f(\hat{k})}{\hat{k}} = 0, \quad \text{et} \quad \lim_{\hat{k} \rightarrow 0} \frac{f(\hat{k})}{\hat{k}} = +\infty$$

D'après (4.12), le taux de croissance du stock de capital par unité de travail effectif est donné par l'écart entre la fonction $sf(\hat{k})/\hat{k}$ et $(a + n + \delta)$. Cet écart est représenté sur la figure 4.2.

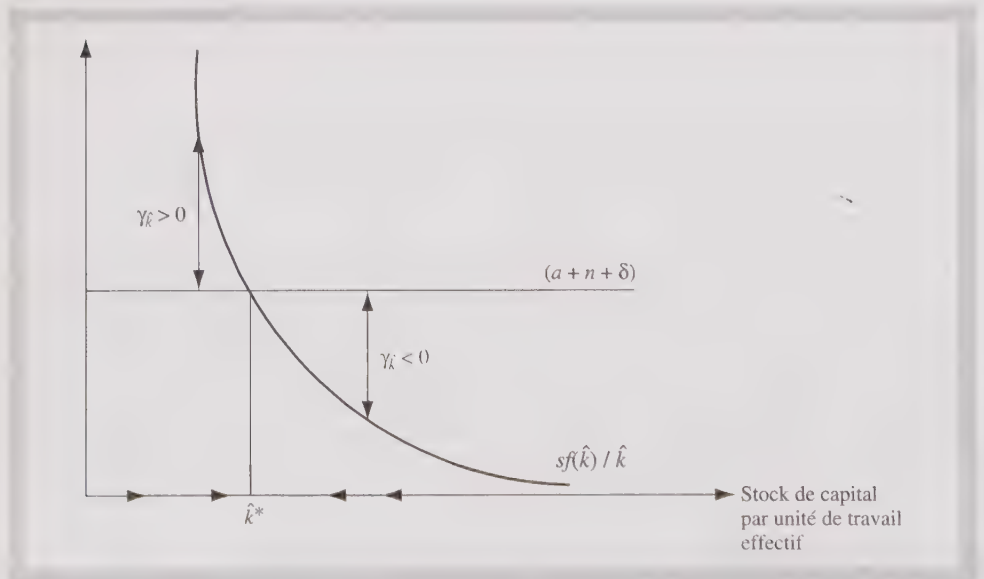


Figure 4.2 Dynamique d'ajustement vers l'équilibre

Le taux de croissance de \widehat{k} correspond à la distance entre la courbe $sf(\widehat{k})/\widehat{k}$ et la droite de dépréciation effective $(a+n+\delta)$. Si $\widehat{k} < \widehat{k}^*$ alors le taux de croissance de \widehat{k} est positif et \widehat{k} s'accroît jusqu'en \widehat{k}^* . Par contre, si $\widehat{k} > \widehat{k}^*$ alors le taux de croissance de \widehat{k} est négatif et \widehat{k} diminue jusqu'en \widehat{k}^* .

L'économie décrite par le modèle de Solow converge donc vers un équilibre stable et de nombreuses études empiriques cherchent à évaluer le délai de retour à l'équilibre lorsque, suite à un choc, l'économie s'en est écartée. Une estimation de la vitesse de convergence permet d'évaluer ce délai. Si on retient la définition de Barro-Sala-I-Martin (1996), la vitesse de convergence (β) vers l'équilibre est donnée par la dérivée du taux de croissance $\gamma_{\widehat{k}}$ par rapport à $\text{Log}\widehat{k}$

$$\beta = -\frac{\partial \gamma_{\widehat{k}}}{\partial \text{Log}\widehat{k}} \tag{4.13}$$

Une valeur positive de β indique que l'économie converge vers son équilibre.

2.3.2 Définition de la convergence d'un groupe d'économies : convergence absolue versus convergence conditionnelle

La figure 4.3 montre qu'une économie avec un stock de capital par unité de travail effectif initial \widehat{k}_p croît plus rapidement qu'une économie disposant d'un stock de capital par unité de travail effectif \widehat{k}_r supérieur à \widehat{k}_p . Par conséquent, les deux économies vont converger vers le même équilibre caractérisé par \widehat{k}^* . L'économie initialement la moins dotée en stock de capital par unité de travail effectif rattrape l'économie la mieux dotée.

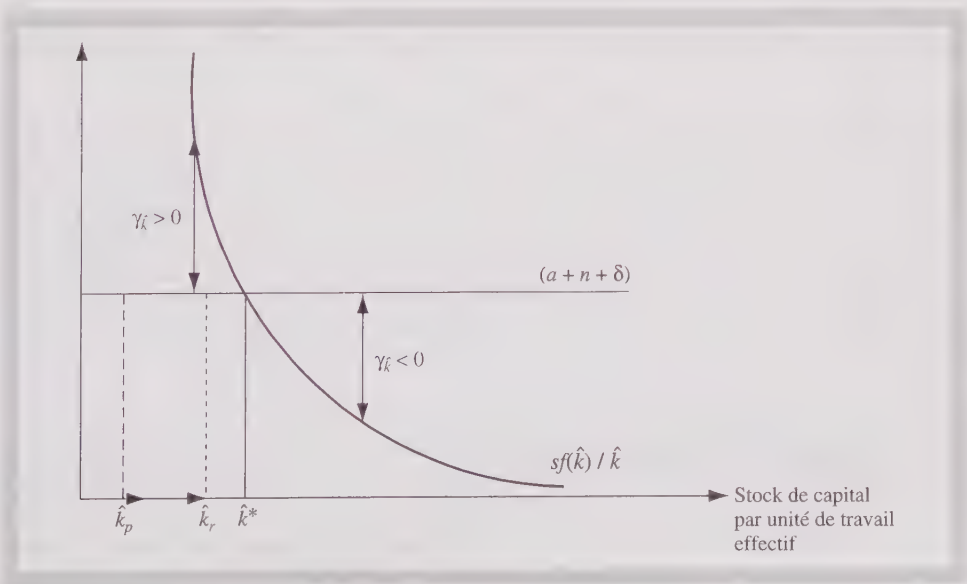


Figure 4.3 Convergence absolue

Par contre, les deux économies représentées sur la figure 4.4 convergent vers deux équilibres différents, et l'économie initialement la plus pauvre ne croît pas plus rapidement que la plus riche. L'économie dont la dotation initiale est \hat{k}_p converge vers l'équilibre \hat{k}_p^* tandis que l'économie avec la dotation initiale \hat{k}_r , converge vers l'équilibre \hat{k}_r^* . Les deux économies ne convergent pas vers le même équilibre car leur taux d'épargne est différent.

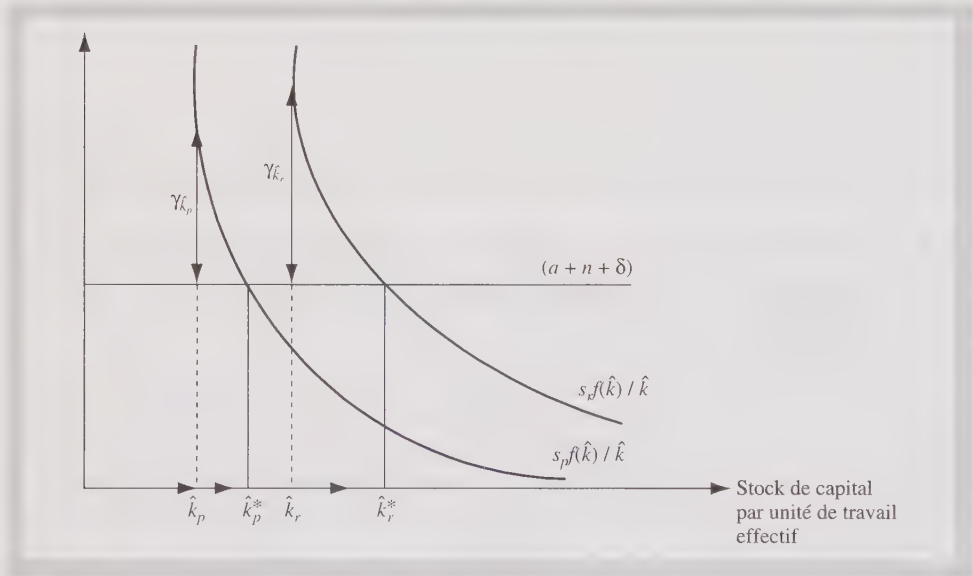


Figure 4.4 Convergence conditionnelle

Sur un ensemble de pays et une période longue, la convergence est absolue lorsqu'on observe une relation décroissante entre le taux de croissance des économies sur la période et leur dotation initiale, indépendamment des variables de conditionnement de l'équilibre c'est-à-dire, le taux d'épargne, le taux de croissance de la population, le taux de croissance du progrès technique et le taux de dépréciation du capital. Par contre, la convergence est conditionnelle lorsque cette relation décroissante est observée seulement si on prend en compte les variables de conditionnement de l'équilibre. Lorsque la convergence est absolue, les économies convergent vers le même équilibre à long terme. Lorsque la convergence est conditionnelle, chaque économie converge vers son équilibre.

3. ANALYSE EMPIRIQUE

Le modèle de Solow permet de caractériser, d'une part, l'équilibre de long terme d'une économie et, d'autre part, la dynamique d'ajustement vers cet équilibre. On définit ainsi deux équations, la première correspondant à la caractérisation de l'équilibre, la seconde à la condition de convergence vers cet équilibre.

3.1 L'équation d'équilibre

Afin de déterminer l'équation d'équilibre, on fait l'hypothèse que la fonction de production est de la forme Cobb-Douglas

$$F(K, AN) = K^\alpha (AN)^{(1-\alpha)} \quad (4.14)$$

On en déduit la fonction de production par unité de travail effectif

$$f(\hat{k}) = F\left(\frac{K}{AN}, 1\right) = \hat{k}^\alpha \quad (4.15)$$

Le stock de capital d'équilibre par unité de travail effectif ne varie pas ($\dot{\hat{k}} = 0$). \hat{k}^* est donc obtenu en égalisant l'investissement par unité de travail effectif et la dépréciation du stock de capital par unité de travail effectif

$$s\hat{k}^{*\alpha} = (a + n + \delta)\hat{k}^* \iff \hat{k}^* = \left[\frac{s}{a + n + \delta}\right]^{\frac{1}{1-\alpha}} \quad (4.16)$$

Les niveaux de production par unité de travail effectif \hat{y}^* et par unité de travail $y^* = A\hat{y}^*$, correspondant à cet équilibre sont

$$\hat{y}^* = f(\hat{k}^*) = \left[\frac{s}{a + n + \delta}\right]^{\frac{\alpha}{1-\alpha}} \quad (4.17)$$

et

$$y^* = A \left[\frac{s}{a + n + \delta}\right]^{\frac{\alpha}{1-\alpha}} \quad (4.18)$$

En introduisant l'équation (4.6), l'écriture en logarithme de l'expression précédente donne

$$\text{Log} y^* = \text{Log} A(0) + at + \frac{\alpha}{1-\alpha} \text{Log} s - \frac{\alpha}{1-\alpha} \text{Log}(n + a + \delta) \quad (4.19)$$

Le terme $A(0)$ représente le niveau initial d'efficacité de l'économie, c'est-à-dire l'efficacité avec laquelle les facteurs de production et la technologie de production sont utilisés pour créer de la richesse. Ce niveau n'est pas observé. Un accroissement de 1 % de cet indice induit un accroissement de 1 % de la production d'équilibre. Le ratio $\frac{\alpha}{1-\alpha}$ est l'élasticité de l'output par tête par rapport au taux d'épargne. Cette élasticité est positive et une augmentation du taux d'épargne induit une augmentation du niveau d'équilibre de la production par tête. L'élasticité de l'output par tête d'équilibre par rapport au taux de dépréciation du stock de capital par unité

de travail effectif est égale à $-\frac{\alpha}{1-\alpha}$. Une augmentation de $n + a + \delta$ conduit donc à une baisse de la production par tête d'équilibre.

Si on suppose que $\text{Log}A(0) = b + \varepsilon$ où b est une constante et ε une variable aléatoire, l'équation (4.19) s'écrit (avec $t = 0$ pour simplifier)

$$\text{Log}y^* = b + \frac{\alpha}{1-\alpha} \text{Log}s - \frac{\alpha}{1-\alpha} \text{Log}(n + a + \delta) + \varepsilon \quad (4.20)$$

3.2 L'équation de convergence

La seconde partie de l'analyse empirique a pour but de tester économétriquement l'hypothèse de convergence d'un groupe d'économies vers un équilibre de long terme stable, c'est-à-dire de tester si la vitesse de convergence est positive.

On maintient l'hypothèse d'une fonction de production Cobb-Douglas. Le taux de croissance du stock de capital par unité de travail effectif est alors

$$\gamma_{\hat{k}} = \frac{\dot{\hat{k}}}{\hat{k}} = s\hat{k}^{\alpha-1} - (a + n + \delta) \quad (4.21)$$

et on montre qu'autour du point d'équilibre (cf. annexe 2 à la fin du chapitre)

$$\gamma_{\hat{k}} = \beta \left(\text{Log}\hat{k}^* - \text{Log}\hat{k} \right) \quad (4.22)$$

avec $\beta = -\frac{\partial \gamma_{\hat{k}}}{\partial \text{Log}(\hat{k})} = (1 - \alpha)(a + n + \delta)$

Le taux de croissance et la vitesse de convergence de l'output par unité de travail effectif peuvent être déduits de l'équation (4.22). On montre que le taux de croissance de l'output par unité de travail effectif est égal au taux de croissance du stock de capital par unité de travail effectif multiplié par la part α du capital dans l'output¹⁰

$$\gamma_{\hat{y}} = \frac{\dot{\hat{y}}}{\hat{y}} = \frac{f'(\hat{k})}{f(\hat{k})} \frac{\dot{\hat{k}}}{\hat{k}} = \frac{f'(\hat{k})}{f(\hat{k})} \frac{\dot{\hat{k}}}{\hat{k}} = \frac{\alpha}{\hat{k}} \dot{\hat{k}} = \alpha \gamma_{\hat{k}} \quad (4.23)$$

À l'équilibre, le taux de croissance de la production par unité de travail effectif est égal à celui du stock de capital par unité de travail effectif. Il est donc nul $\gamma_{\hat{y}^*} = \gamma_{\hat{k}^*} = 0$. Le taux de croissance de l'output $\gamma_Y = \gamma_{\hat{y}} + (a + n)$ (cf. annexe 1 à la fin du chapitre), est identique au taux de croissance du stock de capital à l'équilibre $\gamma_{Y^*} = \gamma_{K^*} = a + n$.

¹⁰ Dans le cas d'une fonction Cobb-Douglas

$$\frac{f'(\hat{k})}{f(\hat{k})} = \alpha \frac{\hat{k}^{\alpha-1}}{\hat{k}^\alpha} = \frac{\alpha}{\hat{k}}$$

On montre aussi que la vitesse de convergence de l'output par unité de travail effectif est identique à la vitesse de convergence du stock de capital par unité de travail effectif

$$\gamma_{\hat{y}} = \frac{\dot{\hat{y}}(t)}{\hat{y}(t)} = \alpha\beta \left[\text{Log}\hat{k}^* - \text{Log}\hat{k} \right] = \beta \left[\text{Log}\hat{y}^* - \text{Log}\hat{y} \right] \quad (4.24)$$

Pour des valeurs des paramètres de $\alpha = 1/3$ et $(n + a + \delta) = 0.1$, la vitesse est d'environ 7 %. Si l'économie s'éloigne de son équilibre suite un choc, elle réduira l'écart par rapport à l'équilibre d'environ 7 % par an.

Les différentes analyses empiriques sur la convergence des économies cherchent à déterminer si les économies convergent et à évaluer la vitesse de convergence. Pour cela, elles estiment une équation de convergence dérivée de l'équation (4.24).

L'équation (4.24) est une équation différentielle en $\text{Log}(\hat{y}(t))$ qui admet pour solution (cf. annexe 3 à la fin du chapitre)

$$\text{Log}(\hat{y}(t)) = (1 - e^{-\beta t})\text{Log}(\hat{y}^*) + e^{-\beta t}\text{Log}(\hat{y}(0)) \quad (4.25)$$

D'après cette équation, si la vitesse de convergence est de 7 %, alors la moitié de l'écart initial par rapport à l'équilibre est absorbée en 10 ans. Lorsque $e^{-\beta t} = 1/2$, pour $\beta = 0.07$, on obtient $t = 10$ ans.

L'équation de convergence vers l'output par unité de travail effectif est obtenue en substituant \hat{y}^* par sa valeur donnée par l'équation (4.17)

$$\text{Log}\hat{y}(t) - \text{Log}\hat{y}(0) = -(1 - e^{-\beta t})\text{Log}\hat{y}(0) + \frac{(1 - e^{-\beta t})\alpha}{1 - \alpha} [\text{Log}s - \text{Log}(n + a + \delta)] \quad (4.26)$$

En termes de production par unité de travail, cette équation devient

$$\begin{aligned} \text{Log}y(t) - \text{Log}y(0) = & -(1 - e^{-\beta t})\text{Log}y(0) + \\ & \frac{(1 - e^{-\beta t})\alpha}{1 - \alpha} [\text{Log}s - \text{Log}(n + a + \delta)] + \\ & (1 - e^{-\beta t})\text{Log}A(0) + at \end{aligned} \quad (4.27)$$

Cette équation décrit le processus de convergence d'une économie vers son équilibre macroéconomique. La variable $\text{Log}y(t) - \text{Log}y(0)$ correspond au taux de croissance de l'économie entre la période 0 et t . Cette économie converge vers un équilibre de revenu par tête stable lorsque la vitesse de convergence β est positive. Les variables de conditionnement de l'équilibre sont le taux d'épargne s et la dépréciation du capital par unité de travail effectif $(n + a + \delta)$. Toutes choses égales par ailleurs, un accroissement du taux d'épargne induit une augmentation du taux

de croissance de l'output par tête et, un accroissement de la dépréciation du capital par unité de travail effectif provoque une réduction du taux de croissance. Dans ce modèle théorique la semi-élasticité du taux de croissance de l'output par tête par rapport au taux d'épargne est identique à la semi-élasticité par rapport à la dépréciation du capital par unité de travail effectif au signe près.

Si on reprend l'hypothèse que $\text{Log}A_0 = b + \varepsilon$, l'équation (4.27) devient

$$\begin{aligned} \text{Log}y(t) - \text{Log}y(0) &= (1 - e^{-\beta t})b - (1 - e^{-\beta t})\text{Log}y(0) + \\ &+ (1 - e^{-\beta t})\frac{\alpha}{1 - \alpha} [\text{Log}s - \text{Log}(n + a + \delta)] + at + (1 - e^{-\beta t})\varepsilon_t \end{aligned} \quad (4.28)$$

3.3 Analyse de la convergence

Les données

Les séries ont été extraites de la base de données de la Pen-World-Table¹¹, base de données macroéconomiques internationales sur la période 1950-2000. De cette base, sur la période 1970-2000, on extrait les séries dont les intitulés sont

POP : Population in 000's,

RGDPCH : Real GDP chain per capita — constant prices,

RGDPWOK : Real GDP chain per worker 1985 — constant prices (cette série correspond dans le modèle présenté à l'output par unité de travail noté y),

ki : Investment Share of real GDP per capita — constant prices (cette série permet d'obtenir le taux d'épargne s car dans le modèle l'épargne est égale à l'investissement).

La quantité de travail N à la date t est obtenue à partir des séries *POP*, *RGDPCH* et *RGDPWOK*

$$N_t = \text{POP}_t \frac{\text{RGDPCH}_t}{\text{RGDPWOK}_t}$$

on déduit n , le taux de croissance annuel de la quantité de travail entre la période $t - 1$ et t

$$n_t = \frac{N_t}{N_{t-1}} - 1$$

Enfin, le taux de croissance du progrès technologique et le taux de dépréciation du capital ne sont pas connus précisément. On supposera comme le suggère Mankiw, Romer et Weil (1992)¹² que la somme de ces deux taux est égale à 0.05.

¹¹ Adresse internet : <http://pwt.econ.upenn.edu/>

¹² Mankiw N.G., Romer D., Weil D.N., 1992, « A Contribution to the Empirics of Economic Growth », *Quarterly Journal of Economics*, pp. 407-437. •

Les données nécessaires à l'estimation des équations d'équilibre (équation (4.20)) et de convergence (équation (4.28)) sont présentées sous un format excel dans le fichier chapitre4.xls. On note

$Logy2000_i$ ($= LogRGDPWOK_{(2000)_i}$) le logarithme du PIB par tête du pays i en 2000,

$Logy70_i$ ($= LogRGDPWOK_{(1970)_i}$) le logarithme du PIB par tête du pays i en 1970,

$Logs2000_i$ ($= Logk_{(2000)_i}$) le logarithme du taux d'investissement du pays i en 2000, supposé égal au taux d'épargne dans le modèle,

$Logn2000_i$ ($= Log(n + a + \delta)_{(2000)_i}$) le logarithme de la somme du taux de croissance de population en 2000, du taux de croissance du progrès technologique et du taux de dépréciation du stock de capital (avec $(a + \delta) = 0.05$),

$Logs_i$ ($= Log\bar{k}_i$) le logarithme du taux d'épargne national annuel moyen sur la période 1970-2000,

$Logn_i$ ($= Log(\bar{n} + a + \delta)$) le logarithme de la somme du taux de croissance annuel moyen de la quantité de travail sur la période 1970-2000 et de $(a + \delta)$.

Les pays concernés par l'étude empirique sont les pays de l'OCDE, dans l'ordre du fichier de données : Australie, Autriche, Belgique, Canada, Suisse, Danemark, Espagne, Finlande, France, Grande-Bretagne, Allemagne, Grèce, Hongrie, Irlande, Islande, Italie, Japon, République de Corée, Luxembourg, Mexique, Nouvelle Zélande, Pologne, Portugal, Suède, Turquie, USA. Le fichier de données regroupe un ensemble de pays plus large que celui des pays de l'UE afin d'avoir un échantillon de données plus vaste et de pouvoir tester l'homogénéité du comportement des pays de l'UE au sein des pays de l'OCDE.

Problème 4.1

Estimation de l'équation d'équilibre

1. Estimer avec la méthode des MCO l'équation d'équilibre

$$Logy2000_i = \beta_1 + \beta_2 Logs2000_i + \beta_3 Logn2000_i + \varepsilon \quad i = 1, \dots, 28$$

Commenter d'un point de vue statistique les résultats, donner une mesure de la qualité de l'ajustement et tester la significativité des coefficients et la significativité globale de la régression.

2. Dans le modèle théorique (équation (4.20)) $\beta_2 = -\beta_3$. Tester cette hypothèse et commenter les résultats.
3. Dans le cadre du modèle (4.20), préciser à quelle condition l'estimateur des MCO est convergent.

Problème 4.2

Test de l'hypothèse de convergence

1. Calculer le taux de croissance de l'output par tête entre 1970 et 2000 noté $DLogy$ ($= Logy2000 - Logy70$) pour chaque pays et représenter graphique-

ment celui-ci en fonction du logarithme de son niveau initial ($Logy_{70}$). Estimer l'équation de convergence suivante avec la méthode des MCO

$$DLogy_i = \beta_1 + \beta_2 Logy_{70i} + \varepsilon_i$$

Commenter les résultats d'un point de vue statistique (tests de significativité, qualité d'ajustement, ...).

Calculer la série des valeurs prédites de la variable $DLogy$, on la note $DLogyp$.

2. Le graphique réalisé dans la question précédente fait apparaître que le modèle sous-évalue le taux de croissance du PIB par tête notamment pour 4 pays : Suisse, Mexique, Nouvelle Zélande et Turquie et surévalue le taux de croissance du PIB par tête pour 3 pays : Irlande, République de Corée et Luxembourg. Afin d'améliorer la qualité de l'ajustement, on crée une variable indicatrice notée $DUMS$ qui prend la valeur 1 pour les 4 pays dont le PIB par tête est sous-évalué (0 sinon) et une variable indicatrice $DUMD$, qui prend la valeur 1 pour les 3 pays dont le PIB par tête est surévalué (0 sinon). Estimer l'équation de convergence augmentée de ces variables

$$DLogy_i = \beta_1 + \beta_2 Logy_{70i} + \beta_3 DUMS + \beta_4 DUMD + \varepsilon_i$$

Commenter les résultats et indiquer leur interprétation économique.

3. Estimer l'équation de convergence conditionnelle dérivée du modèle théorique

$$DLogy_i = \beta_1 + \beta_2 Logy_{70i} + \beta_3 Logs_i + \beta_4 Logn_i + \varepsilon_i$$

- a) Tester l'hypothèse $H_0 : \beta_3 + \beta_4 = 0$ contre l'hypothèse $H_1 : \beta_3 + \beta_4 \neq 0$. Sous l'hypothèse H_0 , déterminer la vitesse de convergence et la part du capital dans l'output.
- b) Effectuer le test de modèle emboîté avec pour hypothèse nulle $\beta_3 = \beta_4 = 0$. Il s'agit d'un test de deux contraintes, sous l'hypothèse nulle on retrouve l'équation estimée dans la question 1 du problème 4.2.
4. Estimer l'équation de convergence conditionnelle augmentée des variables indicatrices $DUMS$ et $DUMD$

$$DLogy_i = \beta_1 + \beta_2 Logy_{70i} + \beta_3 Logs_i + \beta_4 Logn_i + \beta_5 DUMS + \beta_6 DUMD + \varepsilon_i$$

- a) Tester l'hypothèse nulle $\beta_3 + \beta_4 = 0$ et calculer la vitesse de convergence et la part du capital dans l'output. Interpréter d'un point de vue économique les résultats.
- b) Calculer la série des valeurs prédites de la variable $DLogy$, on la note $DLogypdum$ puis représenter graphiquement les séries $DLogy$, $DLogyp$, $DLogypdum$. Commenter le graphique.
- c) Calculer la valeur prédite du taux de croissance du PIB par tête d'un pays dont

$$Logy_{70} = 10, Logs = 3.15, logn = -2.85, DMUS = 0 \text{ et } DUMD = 0$$

et donner un intervalle de confiance pour cette prédiction.

d) Répondre à la question précédente lorsque

$$\text{Logy}_{70} = 10, \text{Logs} = 3.15, \text{logn} = -2.85, \text{DMUS} = 1 \text{ et } \text{DUMD} = 0$$

Problème 4.3

Analyse du processus de convergence avec introduction d'effets spécifiques à l'UE

On suppose dans l'estimation du problème 4.2 que tous les pays de l'OCDE ont un comportement homogène. Afin de tester si le comportement des pays de l'UE est identique au comportement des autres pays de l'OCDE, on peut introduire une variable indicatrice pour les pays de l'UE. On se propose alors d'estimer le modèle

$$\begin{aligned} D\text{Logy}_i = & \beta_1 + \beta_2 \text{Logy}_{70i} + \beta_3 \text{Logs}_i + \beta_4 \text{Logn}_i + \beta_5 \text{Dumeu}_i \\ & + \beta_6 (\text{Dumeu}_i \times \text{Logy}_{70i}) + \beta_7 \text{DUMS} + \beta_8 \text{DUMD} + \varepsilon_i \end{aligned}$$

où la variable indicatrice *Dumeu* prend la valeur 1 pour les pays de l'UE, 0 pour les autres pays.

1. Estimer le modèle précédent avec la méthode des MCO et tester à nouveau l'hypothèse $H_0 : \beta_3 + \beta_4 = 0$ contre l'hypothèse $H_1 : \beta_3 + \beta_4 \neq 0$.
2. Interpréter les résultats d'un point de vue économique.

4. ÉLÉMENTS DE CORRIGÉ

Problème 4.1

Estimation de l'équation d'équilibre

1. Les résultats d'estimation du modèle avec des données en coupe transversale sur les 28 pays sont présentés dans le tableau 4.1.

Le tableau 4.1 suggère plusieurs commentaires

- La qualité de l'ajustement est mauvaise car le coefficient de détermination est égal à 0.005367. Seulement 0.5 % de la variation du logarithme du PIB par tête est expliquée par le modèle.
- Le test de Fisher de significativité de la régression indique que l'hypothèse de nullité de tous les coefficients, sauf celui de la constante, n'est pas refusée car la statistique de Fisher calculée ($F = 0.0675$) est plus faible que la valeur tabulée $F(2,25) = 3.39$ pour un seuil de risque égal à 5 %. Autrement dit, le niveau de significativité du test (égal à 93 %) est supérieur à 5 %.
- Les tests de significativité des coefficients indiquent que les coefficients des variables *Logs2000* et *Logn2000* ne sont pas significativement différents de 0 pour un seuil de 5 % (le niveau de significativité des tests est respectivement égal à 84 % et 75 %).

Tableau 4.1 Estimation de l'équation d'équilibre par les MCO

Linear Regression — Estimation by Least Squares					
Dependent Variable LOGY2000					
Usable Observations	28	Degrees of Freedom	25		
Centered R**2	0.005367	R Bar **2	-0.074203		
Uncentered R**2	0.998800	T x R**2	27.966		
Mean of Dependent Variable	10.679621404				
Std Error of Dependent Variable	0.378013552				
Standard Error of Estimate	0.391787523				
Sum of Squared Residuals	3.8374365808				
Regression F(2,25)	0.0675				
Significance Level of F	0.93494022				
Durbin-Watson Statistic	2.412485				

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	9.639385662	3.009782623	3.20269	0.00369159
2.	LOGS2000	0.136460807	0.708884667	0.19250	0.84890521
3.	LOGN2000	-0.211761450	0.660281816	-0.32071	0.75109286

2. Le tableau 4.2 présente les résultats du modèle estimé avec la contrainte $\beta_2 = -\beta_3$.

Tableau 4.2 Estimation de l'équation d'équilibre sous contrainte

Linear Model — Estimation by Restricted Regression					
Dependent Variable LOGY2000					
Usable Observations	28	Degrees of Freedom	26		
Centered R**2	0.005116	R Bar **2	-0.033148		
Uncentered R**2	0.998800	T x R**2	27.966		
Mean of Dependent Variable	10.679621404				
Std Error of Dependent Variable	0.378013552				
Standard Error of Estimate	0.384227750				
Sum of Squared Residuals	3.8384050567				
Regression F(1,26)	0.1337				
Significance Level of F	0.71757535				
Durbin-Watson Statistic	2.424372				

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	9.609238547	2.928144989	3.28168	0.00294016
2.	LOGS2000	0.176899026	0.483777093	0.36566	0.71757535
3.	LOGN2000	-0.176899026	0.483777093	-0.36566	0.71757535

L'application du test de Fisher conduit à ne pas rejeter la contrainte d'égalité en valeur absolue des coefficients des variables *Logs2000* et *Logn2000* car le niveau de significativité du test est égal à 93 %.

$$F(1, 25) = \frac{(3.8384050567 - 3.8374365808) / 1}{3.8374365808 / 25} = 0.00631$$

Le signe des coefficients correspond au signe attendu. Lorsque le taux d'épargne augmente de 1 %, le revenu par tête d'équilibre s'accroît de 17 %. De même, lorsque le taux de croissance de la population augmente du taux de croissance du progrès technique et du taux de dépréciation du stock de capital s'accroît de 1 %, le revenu par tête d'équilibre baisse de 17 %. Cependant, d'un point de vue statistique, les résultats ne sont pas satisfaisants. Le coefficient de détermination est très faible (0.005). Le test de significativité globale de la régression conduit à accepter l'hypothèse nulle selon laquelle tous les coefficients (sauf celui de la constante) sont nuls, le niveau de significativité du test est de 71 %.

Les données ne valident donc pas le modèle théorique présenté.

3. L'estimateur des MCO utilisé précédemment est convergent sous l'hypothèse que les variables taux d'épargne, taux de croissance de la population sont indépendantes de l'aléa. Cette hypothèse d'indépendance, et donc d'exogénéité, peut être remise en question car l'aléa correspond à des facteurs individuels, tel que le niveau de technologie, qui déplacent la fonction de production et peuvent affecter le taux d'épargne ou le taux de croissance de la population.

Plus généralement, la question de la propriété de convergence de l'estimateur des MCO se pose lorsque, des variables explicatives supposées exogènes sont en réalité endogènes et par conséquent, non indépendantes des aléas du modèle (cf. chapitre 6, Les hypothèses du modèle de régression linéaire multiple).

Problème 4.2

Test de l'hypothèse de β -convergence

1. La figure 4.5 représente les taux de croissance de l'output par unité de travail entre 1970 et 2000 des pays en fonction du niveau du logarithme de l'output par unité de travail en 1970.

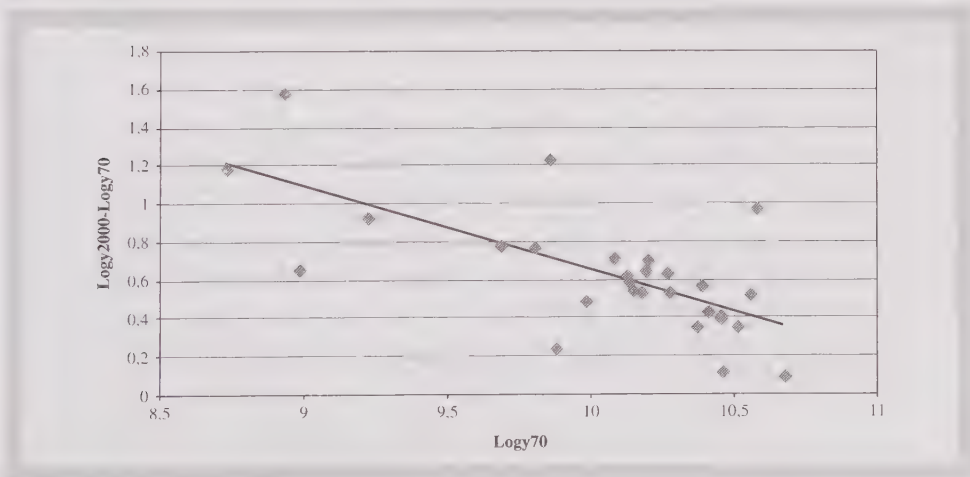


Figure 4.5 Taux de croissance de l'output par unité de travail effectif entre 1970-2000

La figure 4.5 met en évidence un processus de convergence des pays de l'OCDE. Indépendamment des variables de conditionnement de l'équilibre, c'est-à-dire *Logs* et *Logn*, on observe en effet que les pays « les plus pauvres » en 1970 ont eu une croissance entre 1970 et 2000 plus forte que les pays les plus riches. 7 pays parmi les 25 sont éloignés de la droite de régression de l'échantillon : Suisse, Irlande, République de Corée, Luxembourg, Mexique, Nouvelle-Zélande et Mexique. Le taux de croissance du PIB par tête est sous-évalué notamment pour la Suisse, le Mexique, la Nouvelle Zélande et la Turquie et il est surévalué pour l'Irlande, la République de Corée et le Luxembourg.

L'équation de la droite de régression représentée sur le graphique est donnée dans le tableau 4.3.

Tableau 4.3 Test de convergence absolue

Linear Regression — Estimation by Least Squares					
Dependent Variable DLOGY					
Usable Observations	28	Degrees of Freedom	26		
Centered R**2	0.470653	R Bar **2	0.450294		
Uncentered R**2	0.888694	T x R**2	24.883		
Mean of Dependent Variable	0.6236045109				
Std Error of Dependent Variable	0.3276852363				
Standard Error of Estimate	0.2429530067				
Sum of Squared Residuals	1.5346802498				
Regression F(1,26)	23.1171				
Significance Level of F	0.00005578				
Durbin-Watson Statistic	1.990376				

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	4.986482002	0.908576118	5.48824	0.00000931
2.	LOGY70	-0.433857415	0.090236053	-4.80803	0.00005578

La variable *DLogy* représente le taux de croissance de la production par tête entre 1970 et 2000, 47 % de la variation de ce taux est expliquée par le modèle. Les tests de significativité globale de la régression et de significativité du coefficient de *Logy70* sont identiques car le modèle estimé est un modèle de régression simple. Le niveau de significativité du test est égal à 0.00005, il est inférieur à 5 % donc on rejette l'hypothèse nulle selon laquelle le coefficient de la variable *Logy70* est égal à 0.

La pente de la droite de régression est égale à -0.433, la vitesse de convergence peut alors être calculée de la manière suivante

$$[-(1 - e^{-\beta\tau})] = -0.433 \iff \hat{\beta} = -\frac{\text{Log}(1 - 0.433)}{31} = 1.83 \%$$

En cas de choc, l'économie retourne vers son équilibre à une vitesse de 1.83 % par an. Ainsi, la moitié de l'écart à l'équilibre en cas de choc est rattrapé en

37 ans environ

$$e^{-0.0183t} = \frac{1}{2} \iff t = \frac{\text{Log}(1/2)}{-0.0183} = 37.8$$

L'analyse graphique et statistique indique qu'il y a convergence des pays.

2. Lorsqu'on introduit dans l'équation de convergence les variables indicatrices DUMS et DUMD, la qualité de l'ajustement s'améliore. Le coefficient de détermination ajusté passe de 45 % à 94 % (Tableau 4.4). Le test de significativité globale de la régression conduit à rejeter l'hypothèse nulle selon laquelle tous les coefficients (sauf celui de la constante) sont nuls. Chaque coefficient est significativement non nul d'après le test de significativité des paramètres.

Tableau 4.4 Test de convergence lorsque le modèle comporte des variables indicatrices

Linear Regression — Estimation by Least Squares					
Dependent Variable DLOGY					
Usable Observations	28	Degrees of Freedom	24		
Centered R**2	0.949610	R Bar **2	0.943311		
Uncentered R**2	0.989404	T x R**2	27.703		
Mean of Dependent Variable	0.6236045109				
Std Error of Dependent Variable	0.3276852363				
Standard Error of Estimate	0.0780199601				
Sum of Squared Residuals	0.1460907403				
Regression F(3,24)	150.7614				
Significance Level of F	0.00000000				
Durbin-Watson Statistic	2.123746				

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	4.502777607	0.298986852	15.06012	0.00000000
2.	LOGY70	-0.386212449	0.029541295	-13.07365	0.00000000
3.	DUMS	-0.369408344	0.042673889	-8.65654	0.00000001
4.	DUMD	0.535345378	0.049043147	10.91580	0.00000000

L'hypothèse de convergence est retenue car le signe du coefficient de la variable *Logy70* est négatif. La vitesse de convergence est égale à

$$[-(1 - e^{-\beta\tau})] = -0.386 \iff \hat{\beta} = -\frac{\text{Log}(1 - 0.386)}{31} = 1.57 \%$$

et la moitié de l'écart à l'équilibre en cas de choc est rattrapé en 44 ans

$$e^{-0.0157t} = \frac{1}{2} \iff t = \frac{\text{Log}(1/2)}{-0.0157} = 44$$

D'après les résultats, pour les 28 pays sauf la Suisse, l'Irlande, la République de Corée, le Luxembourg, le Mexique, la Nouvelle-Zélande et le Mexique

$$D \widehat{\log y} = 4.502 - 0.386 \text{Log}y_{70}$$

pour la Suisse, le Mexique, la Nouvelle Zélande et la Turquie

$$D \widehat{\log y} = (4.502 - 0.369) - 0.386 \text{Log}y_{70}$$

et pour l'Irlande, la République de Corée et le Luxembourg

$$D \widehat{\log y} = (4.502 + 0.535) - 0.386 \text{Log}y_{70}$$

3. Tests de convergence conditionnelle

- a) Les résultats d'estimations de l'équation de convergence conditionnelle du modèle théorique sont présentés dans le tableau 4.5.

Tableau 4.5 Test de convergence conditionnelle

Linear Regression — Estimation by Least Squares			
Dependent Variable DLOGY			
Usable Observations	28	Degrees of Freedom	24
Centered R**2	0.543958	R Bar **2	0.486953
Uncentered R**2	0.904108	T x R**2	25.315
Mean of Dependent Variable	0.6236045109		
Std Error of Dependent Variable	0.3276852363		
Standard Error of Estimate	0.2347120717		
Sum of Squared Residuals	1.3221541589		
Regression F(3,24)	9.5423		
Significance Level of F	0.00024803		
Durbin-Watson Statistic	1.987995		

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	2.701752283	1.585881779	1.70363	0.10136704
2.	LOGY70	-0.431318391	0.088713620	-4.86192	0.00005901
3.	LOGS	0.470355898	0.274606948	1.71283	0.09963500
4.	LOGN	-0.277409400	0.344156601	-0.80606	0.42812576

F(1,24) = 0.17680 with Significance Level 0.67787652

Le modèle explique 54 % des variations du taux de croissance des pays entre 1970-2000. Le coefficient de la variable *Logy70* est significativement différent de 0 et les coefficients des variables *Logs* et *Logn* sont significativement égaux à 0.

Cependant, le test d'égalité en valeur absolue des coefficients des variables *Logs* et *Logn* est accepté car le niveau de significativité du test est égal à

67 % et donc supérieur à 5 %. De manière équivalente, on peut dire que la statistique de Fisher calculée égale à 0.17 est inférieure à la valeur tabulée du Fisher égale à 4.26 pour un seuil de risque de 5 %. Par conséquent, les résultats du modèle contraint (tableau 4.6) vont être analysés car l'estimateur des MCO sous contrainte est non biaisé lorsque la contrainte est vraie, et possède une variance plus faible que les estimateurs du modèle non contraint (cf. chapitre 3, Le modèle de régression linéaire multiple). Le tableau 4.6 donne les résultats d'estimation lorsque la contrainte sur les coefficients, $\beta_3 + \beta_4 = 0$, est imposée.

Tableau 4.6 Estimation de l'équation théorique de convergence conditionnelle

Linear Model — Estimation by Restricted Regression					
Dependent Variable DLOGY					
Usable Observations	28	Degrees of Freedom	25		
Centered R**2	0.540599	R Bar **2	0.503847		
Uncentered R**2	0.903401	T x R**2	25.295		
Mean of Dependent Variable	0.6236045109				
Std Error of Dependent Variable	0.3276852363				
Standard Error of Estimate	0.2308154396				
Sum of Squared Residuals	1.3318941791				
Regression F(2,25)	14.7093				
Significance Level of F	0.00005990				
Durbin-Watson Statistic	1.933485				

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	2.549477152	1.518346134	1.67911	0.10558706
2.	LOGY70	-0.424699396	0.085856411	-4.94662	0.00004279
3.	LOGS	0.393599693	0.201743963	1.95099	0.06235183
4.	LOGN	-0.393599693	0.201743963	-1.95099	0.06235183

Vitesse annuelle 0.01783

Part du capital 0.48100

Le modèle explique 54 % de la variation du taux de croissance du PIB par unité de travail des pays entre 1970-2000. Le coefficient du logarithme du PIB par unité de travail en 1970 est significativement différent de 0 au seuil de 5 %. Le niveau de significativité du test de Student s'élève à 6.2 % pour les variables *Logs* et *Logn*, l'hypothèse nulle n'est pas acceptée seulement pour un risque α de 10 %. Le signe du coefficient de *Logs* et *Logn* est conforme au résultat attendu. La vitesse de convergence des économies est égale à 1.7 %¹³ par an et la part du capital dans le revenu à 0.48¹⁴. Cette valeur peut apparaître élevée car on estime usuellement que cette part est de 30 % environ.

¹³ $\hat{\beta} = -\frac{\text{Log}(1-0.393)}{31} = 1.7\%$

¹⁴ $(1 - e^{-\beta\tau}) \frac{\alpha}{1-\alpha} = 0.39 \iff \frac{\alpha}{1-\alpha} = \frac{0.39}{0.42} = 0.92 \iff \hat{\alpha} = \frac{0.92}{1+0.92} = 0.48$

- b) Le test de modèle emboîté proposé teste l'hypothèse nulle selon laquelle les coefficients des variables *Logs* et *Logn* sont simultanément nuls. Les résultats d'estimation du modèle sans les deux contraintes sont présentés dans le tableau 4.5 (la somme des carrés des résidus est égale à 1.3221541589) et ceux du modèle avec les deux contraintes dans le tableau 4.3 (la somme des carrés des résidus est égale à 1.5346802498). La statistique de Fisher, distribuée selon un $F(2,24)$, est égale à 1.92

$$F_{stat} = \frac{(1.5346802498 - 1.3221541589)/2}{1.3221541589/24} = 1.92$$

et le niveau de significativité du test est de 16 %. Par conséquent, l'hypothèse nulle n'est pas rejetée. Selon ce résultat, les variables *logs* et *logn* ne seraient pas des variables de conditionnement de l'équilibre.

4. Test de convergence conditionnelle lorsque les variables DUMS et DUMD sont prises en compte.
- a) Seuls les résultats de l'estimation du modèle sous la contrainte $\beta_3 + \beta_4 = 0$ sont analysés (tableau 4.6) car le test de contrainte de Fisher conduit à accepter cette hypothèse avec un niveau de significativité de 93 %. 96 % des variations du taux de croissance du PIB par tête sont expliquées par le modèle. Le test de significativité globale ne retient pas l'hypothèse nulle, de même les tests de significativité des paramètres rejettent l'hypothèse nulle.

Tableau 4.7 Équation de convergence conditionnelle augmentée des variables indicatrices

$F(1,22) = 0.00698$ with Significance Level 0.93418316

Linear Model — Estimation by Restricted Regression					
Dependent Variable DLOGY					
Usable Observations	28	Degrees of Freedom	23		
Centered R**2	0.960218	R Bar **2	0.953299		
Uncentered R**2	0.991635	T x R**2	27.766		
Mean of Dependent Variable	0.6236045109				
Std Error of Dependent Variable	0.3276852363				
Standard Error of Estimate	0.0708139206				
Sum of Squared Residuals	0.1153360610				
Regression F(4,23)	138.7874				
Significance Level of F	0.00000000				
Durbin-Watson Statistic	1.853470				

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	3.306382886	0.554101875	5.96710	0.00000439
2.	LOGY70	-0.377715613	0.027031443	-13.97319	0.00000000
3.	LOGS	0.184577509	0.074531836	2.47649	0.02105619
4.	LOGN	-0.184577509	0.074531836	-2.47649	0.02105619
5.	DUMS	-0.305585074	0.046522913	-6.56849	0.00000106
6.	DUMD	0.555811766	0.045274112	12.27659	0.00000000

Vitesse annuelle 0.01530

Part du capital 0.32826

La vitesse de convergence est estimée à 1.5 % et la part du capital à 30 %. Lorsque le taux d'épargne augmente de 1 % (ou lorsque le taux de croissance de la population augmente du taux de croissance du progrès technique et du taux de dépréciation du stock de capital décroît de 1 %), le taux de croissance du PIB par tête augmente de 0,18 points de pourcentage.

D'après les résultats, pour les 28 pays, sauf la Suisse, l'Irlande, la République de Corée, le Luxembourg, le Mexique, la Nouvelle-Zélande et le Mexique

$$\widehat{D \log y} = 3.306 - 0.377 \text{Log} y_{70} + 0.184 \text{Log} s - 0.184 \text{Log} n$$

pour la Suisse, le Mexique, la Nouvelle Zélande et la Turquie

$$\widehat{D \log y} = (3.306 - 0,305) - 0.377 \text{Log} y_{70} + 0.184 \text{Log} s - 0.184 \text{Log} n$$

et pour l'Irlande, la République de Corée et le Luxembourg

$$\widehat{D \log y} = (3.306 + 0.555) - 0.377 \text{Log} y_{70} + 0.184 \text{Log} s - 0.184 \text{Log} n$$

- b) La figure 4.6 représente la valeur observée du taux de croissance du PIB par tête ($D \log y$), la valeur prédite à partir de l'équation dont les résultats sont présentés dans le tableau 4.3 ($D \log y$) + la valeur prédite à partir de l'équation dont les résultats sont présentés dans le tableau 4.7 ($D \log y_{\text{pdm}}$).

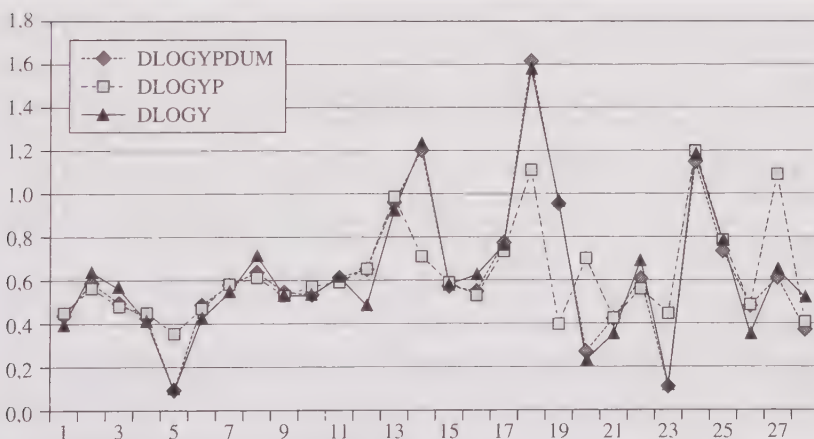


Figure 4.6 Comparaison des valeurs observées et estimées du taux de croissance du PIB/tête

La figure 4.6 montre que le modèle comportant les variables de conditionnement de l'équilibre et les variables indicatrices $DUMS$ et $DUMD$ retrace beaucoup mieux le taux de croissance du PIB par tête des pays.

- c) La valeur prédite du taux de croissance du PIB par tête lorsque

$$\text{Logy}_{70} = 10, \text{Log}s = 3.15, \text{logn} = -2.85, \text{DMUS} = 0 \text{ et } \text{DUMD} = 0$$

est notée $D\text{Logy}1$. Cette valeur est égale à 0.63

$$D\text{Logy}1 = 3.306 - 0.377 \times 10 + 0.184 \times 3.15 - 0.184 \times (-2.85) = 0.63$$

et l'intervalle de confiance est donné par

$$[0.63 \pm 0.072 \times 2.069] = [0.48 ; 0.78]$$

où 0.072 est la valeur de l'écart type estimé de l'erreur de prévision et 2.069 la valeur tabulée de Student pour un seuil de risque égal à 5 % et un degré de liberté correspondant au nombre d'observations moins le nombre de paramètres estimés.

- d) La valeur prédite taux de croissance du PIB par tête lorsque

$$\text{Logy}_{70} = 10, \text{Log}s = 3.15, \text{logn} = -2.85, \text{DMUS} = 1 \text{ et } \text{DUMD} = 0$$

est notée $D\text{Logy}2$. Cette valeur est égale à 0.33

$$D\text{Logy}2 = (3.306 - 0.305) - 0.377 \times 10 + 0.184 \times 3.15 - 0.184 \times (-2.85) = 0.33$$

et l'intervalle de confiance

$$[0.33 \pm 0.082 \times 2.069] = [0.16 ; 0.50]$$

où 0.082 est la valeur de l'écart type estimé de l'erreur de prévision.

Problème 4.3

Analyse du processus de convergence avec introduction d'effets spécifiques aux pays de l'UE

1. Dans l'analyse précédente, on suppose que les pays de l'UE ont le même comportement que les autres pays de l'échantillon. Afin de distinguer les pays de l'UE, on construit une variable indicatrice *dumeu* qui prend la valeur 1 lorsque le pays appartient à l'UE, 0 sinon. Les résultats d'estimation sont présentés dans le tableau 4.8. Les résultats de modèle contraint sont présentés car le test de contraintes de Fisher conduit à retenir l'hypothèse nulle $\beta_3 + \beta_4 = 0$. Le niveau de significativité du test est égal à 57 %.

Tableau 4.8 Équation de convergence avec des effets spécifiques aux pays de l'UE

Linear Model — Estimation by Restricted Regression			
Dependent Variable DLOGY			
Usable Observations	28	Degrees of Freedom	21
Centered R**2	0.962581	R Bar **2	0.951890
Uncentered R**2	0.992132	T x R**2	27.780
Mean of Dependent Variable	0.6236045109		
Std Error of Dependent Variable	0.3276852363		
Standard Error of Estimate	0.0718747433		
Sum of Squared Residuals	0.1084855533		
Regression F(6,21)	90.0349		
Significance Level of F	0.00000000		
Durbin-Watson Statistic	1.921722		

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	3.146234375	0.585559917	5.37304	0.00002498
2.	LOGY70	-0.360425542	0.031338821	-11.50093	0.00000000
3.	LOGS	0.185275152	0.076165542	2.43253	0.02402167
4.	LOGN	-0.185275152	0.076165542	-2.43253	0.02402167
5.	DUMEU	0.660765900	0.874621002	0.75549	0.45834676
6.	DUMEULOGY70	-0.067735050	0.086663371	-0.78710	0.44001877
7.	DUMS	-0.322308330	0.051660153	-6.23901	0.00000345
8.	DUMD	0.563432919	0.046441606	12.13207	0.00000000

2. L'estimation de ce modèle permet de distinguer le comportement des pays de l'UE. La vitesse de convergence et la part du capital estimées peuvent être distinguées pour les pays de l'UE et les autres pays.

Pour les pays de l'UE

$$\hat{\beta} = -\frac{\text{Log}(1 - 0.360 - 0.067)}{31} = 1.8\%$$

$$(1 - e^{-\beta\tau}) \frac{\alpha}{1 - \alpha} = 0.18 \iff \hat{\alpha} = 0.30$$

Pour les autres pays

$$\hat{\beta} = -\frac{\text{Log}(1 - 0.360)}{31} = 1,4\%$$

$$(1 - e^{-\beta\tau}) \frac{\alpha}{1 - \alpha} = 0.18 \iff \hat{\alpha} = 0.34$$

Cependant, les tests de significativité des paramètres indiquent que les variables $Dumeu$ et $Dumeu_i \times Logy_{70i}$ ont des coefficients significativement nuls. Ce résultat implique qu'il n'est pas nécessaire de distinguer le comportement des pays de l'UE.

5. ANNEXES

5.1 *Annexe 1 : Calcul du taux de croissance du stock de capital à l'équilibre*

$K^* = \widehat{k}^* AN$. En divisant l'équation (4.10) par \widehat{k} on obtient

$$\gamma_{\widehat{k}} = \frac{\dot{\widehat{k}}}{\widehat{k}} = \frac{\dot{K}}{AN} \frac{1}{\widehat{k}} - (a + n) \quad (4.29)$$

Or, $\widehat{k} = \frac{K}{AN}$ donc, γ_K le taux de croissance de la variable de stock de capital K s'écrit

$$\gamma_K = \frac{\dot{K}}{K} = \gamma_{\widehat{k}} + (a + n) \quad (4.30)$$

et à l'équilibre $\gamma_{K^*} = (a + n)$ car $\gamma_{\widehat{k}^*} = 0$.

5.2 *Annexe 2 : Détermination du taux de croissance du stock de capital par unité de travail effectif autour de l'équilibre*

L'expression

$$\gamma_{\widehat{k}} = \frac{\dot{\widehat{k}}}{\widehat{k}} = s\widehat{k}^{\alpha-1} - (a + n + \delta)$$

est équivalente à l'expression

$$\gamma_{\widehat{k}} = se^{(\alpha-1)\text{Log}(\widehat{k})} - (a + n + \delta)$$

Pour évaluer la vitesse de convergence β donnée par l'équation (4.13), l'équation précédente est log-linéarisée au point d'équilibre. La linéarisation de la fonction $f(\text{Log}x)$ par rapport à $\text{Log}x$, évaluée au point x^* , c'est-à-dire la log-linéarisation, s'écrit

$$(f(\text{Log}(x)) - f(\text{Log}(x^*))) = \left[\frac{\partial f(\text{Log}(x))}{\partial \text{Log}(x)} \right]_{x=x^*} (\text{Log}(x) - \text{Log}(x^*))$$

D'où :

$$\left(\underbrace{\gamma_{\widehat{k}} - \gamma_{\widehat{k}^*}}_{=0} \right) = \left(s(\alpha - 1)e^{(\alpha-1)\text{Log}\widehat{k}^*} \right) \left(\text{Log}\widehat{k} - \text{Log}\widehat{k}^* \right)$$

$$\gamma_{\widehat{k}} = s(\alpha - 1)\widehat{k}^{*\alpha-1} \left(\text{Log}\widehat{k} - \text{Log}\widehat{k}^* \right)$$

D'après (4.16) $\widehat{k}^* = \left[\frac{s}{a+n+\delta} \right]^{\frac{1}{1-\alpha}}$.

Donc

$$\begin{aligned} \gamma_{\widehat{k}} &= s(\alpha - 1) \left[\frac{s}{a+n+\delta} \right]^{\frac{\alpha-1}{1-\alpha}} \left(\text{Log}\widehat{k} - \text{Log}\widehat{k}^* \right) \\ &= (\alpha - 1)(a+n+\delta) \left(\text{Log}\widehat{k} - \text{Log}\widehat{k}^* \right) \end{aligned}$$

La vitesse de convergence est égale à

$$\beta = -\frac{\partial \gamma_{\widehat{k}}}{\partial \text{Log}(\widehat{k})} = (1 - \alpha)(a+n+\delta)$$

et le taux de croissance du stock de capital par unité de travail effectif autour de l'équilibre est égal à

$$\gamma_{\widehat{k}} = \beta \left(\text{Log}\widehat{k}^* - \text{Log}\widehat{k} \right)$$

5.3 Annexe 3 : Solution de l'équation différentielle de taux de croissance de l'output par unité de travail effectif

L'intégration de l'équation différentielle (4.24) s'effectue de la manière suivante

$$\int \frac{\dot{\widehat{y}}(t)}{\widehat{y}(t)} dt = \int \text{Log}(\dot{\widehat{y}}(t)) dt = \int \beta [\text{Log}(\widehat{y}^*) - \text{Log}(\widehat{y}(t))] dt$$

En regroupant les termes contenant \widehat{y} et en multipliant tous les termes par le facteur d'intégration $e^{\beta t}$, on obtient

$$\int e^{\beta t} \left[\text{Log}(\dot{\widehat{y}}(t)) + \beta \text{Log}(\widehat{y}(t)) \right] dt = \int e^{\beta t} [\beta \text{Log}(\widehat{y}^*)] dt$$

Pour résoudre cette équation, on utilise le fait que

$$\frac{\partial (e^{\beta t} \text{Log}(\hat{y}(t)))}{\partial t} = e^{\beta t} \left[\text{Log}(\dot{\hat{y}}(t)) + \beta \text{Log}(\hat{y}(t)) \right]$$

La solution de l'équation différentielle est donc

$$e^{\beta t} \text{Log}(\hat{y}(t)) = e^{\beta t} \text{Log}(\hat{y}^*) + c \iff \text{Log}(\hat{y}(t)) = \text{Log}(\hat{y}^*) + ce^{-\beta t}$$

avec $c = \text{Log}(\hat{y}(0)) - \text{Log}(\hat{y}^*)$

La solution est finalement donnée par

$$\text{Log}(\hat{y}(t)) = (1 - e^{-\beta t}) \text{Log}(\hat{y}^*) + e^{-\beta t} \text{Log}(\hat{y}(0))$$

5

LES POLITIQUES MONÉTAIRES EUROPÉENNES ET LA RÈGLE DE TAYLOR

1. Introduction
2. Règles monétaires et règle de Taylor
3. Adéquation des politiques monétaires européennes à la règle
4. Éléments de corrigé

OBJECTIFS DU CHAPITRE

D'un point de vue méthodologique

- Estimation d'un modèle de régression linéaire simple
- Pratique des tests de Student et des tests de contraintes de Fisher
- Régressions glissantes
- Test de stabilité du Cusum
- Filtre de Hodrick et Prescott

D'un point de vue économique

- Présentation de la règle de Taylor
- Réflexion sur le SME, l'UEM et la convergence des politiques monétaires

NIVEAU

LICENCE, troisième année et MASTER, 1^{re} année

CHAMP ÉCONOMIQUE TRAITÉ

Macroéconomie monétaire et financière

1. INTRODUCTION

La politique monétaire offre aujourd'hui toutes les apparences d'une belle et plutôt rare unanimité au sein de la communauté économique. En effet, la théorie économique et les autorités monétaires semblent désormais s'accorder pour concevoir et mettre en œuvre une même politique monétaire, définie en tant que règle active centrée sur un objectif d'inflation (« inflation targeting »). Les règles monétaires actives s'opposent à la fois aux politiques discrétionnaires (agissant au coup par coup) et aux règles passives à la Friedman visant une croissance stable de la masse monétaire. Elles apparaissent comme la solution la plus efficace pour une banque centrale cherchant à stabiliser l'inflation à un niveau cible et à limiter les fluctuations macroéconomiques.

La règle de Taylor proposée par J.B. Taylor dans un article de 1993 est la référence centrale de ces règles monétaires actives et l'expression de ces conceptions communes. Elle se présente comme une relation linéaire simple entre le taux d'intérêt à court terme, l'instrument de la politique monétaire, et quelques variables d'inflation et d'activité, dont l'écart entre l'inflation courante et l'objectif d'inflation des autorités monétaires. Formellement, ce n'est donc ni plus ni moins qu'une forme particulière de fonction de réaction de banque centrale, qui est depuis le milieu des années 1980 un instrument bien connu des économètres et modélisateurs pour endogénéiser les taux d'intérêt à court terme, dans un contexte généralisé de régulation monétaire par les taux.

Le succès de la règle de Taylor, outre le fait, comme on le verra, qu'elle synthétise assez bien à un instant donné du temps les positions de la théorie macroéconomique sur la monnaie et la politique monétaire, tient à ce que ses propriétés de stabilisation macroéconomiques ont été identifiées par le biais de simulations sur l'économie américaine et que de surcroît elle correspond bien empiriquement au comportement historique de la réserve fédérale américaine (la FED). Des articles ultérieurs comme celui de Goldman Sachs de 1996, montrant que la règle de Taylor permettait aussi d'expliquer au moins sur certaines périodes le comportement d'autres banques centrales, comme la banque du Japon (la BOJ) et certaines banques centrales européennes, ont achevé de populariser ce modèle et notamment parmi les économistes de banques et d'instituts de conjoncture chargés de décrypter et d'anticiper les comportements de banques centrales.

Ce chapitre propose de réexaminer l'histoire de la politique monétaire européenne, avant le basculement dans l'euro, par référence à la règle de Taylor. La question est donc de savoir dans quelle mesure les différents pays européens ont pu, malgré les contraintes spécifiques du SME et de la construction progressive de l'UEM, se conformer en matière de politique monétaire à la règle normative proposée par Taylor. On y trouvera donc potentiellement aussi des enseignements pour comprendre et interpréter les comportements actuels et futurs de la nouvelle Banque Centrale Européenne (BCE).

Les évaluations économétriques s'inscrivent dans le cadre du modèle linéaire et de l'estimateur des MCO. L'accent est porté ici sur les différentes méthodes d'analyse de la stabilité d'un modèle (estimation par sous-périodes, test du CUSUM à partir de résidus récurifs, régressions glissantes). En outre le filtre de Hodrick

et Prescott est présenté comme méthode de décomposition d'une série économique entre cycle et tendance. On propose ensuite une application aux séries de PIB afin de calculer certaines variables d'intérêt de la règle comme le PIB potentiel et la croissance potentielle.

Le chapitre est organisé de la manière suivante. La section 2 présente la notion de règle monétaire en général et la règle de Taylor en particulier. La section 3 est consacrée aux évaluations empiriques.

2. RÈGLES MONÉTAIRES ET RÈGLE DE TAYLOR

2.1 Règles monétaires actives

La littérature théorique sur les règles monétaires est foisonnante et il n'est pas garanti que les concepts utilisés aujourd'hui soient complètement affinés et définitifs. Le consensus actuel en faveur des règles monétaires s'est construit progressivement au cours des trente dernières années. Nous rappelons ici les éléments et les références incontournables de ce débat théorique.

Deux notions sont au cœur du débat : la neutralité de la monnaie et l'absence de crédibilité ou biais inflationniste des politiques monétaires discrétionnaires.

La neutralité de la monnaie est une idée ancienne des classiques. Elle a été réaffirmée avec force dans les années 1970 et 1980 par l'école des anticipations rationnelles (les nouveaux classiques) et les théoriciens des cycles réels. Les économistes considèrent aujourd'hui que si la monnaie n'est pas complètement neutre à court terme en raison de certaines imperfections (erreurs d'anticipation, rigidité...), elle l'est en revanche à long terme conformément aux preuves empiriques apportées par l'école des cycles réels. Dès lors la politique monétaire doit logiquement se concentrer sur l'ancrage nominal de l'économie, par exemple en définissant un objectif d'inflation maximal (inflation targeting). La logique de soutien à l'activité devient secondaire et doit être entendue comme une possibilité de stabilisation cyclique ou de court terme.

Les travaux de Barro et Gordon (1983)¹ ont montré que dans un régime de politique discrétionnaire (agissant au coup par coup), les autorités monétaires sont toujours soupçonnées par les agents privés de vouloir faire de l'inflation surprise pour stimuler l'activité. Elles sont dès lors peu crédibles, ce qui crée le biais inflationniste de la politique discrétionnaire. On a longtemps cru que ce biais inflationniste ne pouvait être réduit qu'en diminuant le poids accordé à la stabilisation de l'activité, c'est-à-dire en se rapprochant d'une situation de règle passive où les

¹ Barro R.J., Gordon D.B., 1983 « A positive theory of monetary policy in a natural rate model », *Journal of Political Economy*, 91, 586-610.

Barro R.J., Gordon D.B., 1983 « Rule, discretion and reputation in a model of monetary policy », *Journal of Monetary Economics*, 12, 101-121.

autorités monétaires n'ont qu'un objectif d'inflation (Rogoff (1985)²). Il existe pourtant une solution pour sortir de ce conflit entre biais inflationniste et crédibilité. Elle consiste à supposer des règles de comportement actives pour les autorités monétaires où les choix effectués en termes d'inflation et d'activité sont contingents aux «états de la nature» c'est-à-dire aux chocs conjoncturels subis par l'économie. La règle active s'oppose à la règle passive par sa souplesse : l'inflation choisie n'est pas toujours identique à l'inflation cible. La règle active s'oppose à la politique discrétionnaire par l'engagement des autorités sur un mode de comportement immuable, ce qui les préserve de l'incohérence temporelle (Kydland et Prescott (1977)³) et du biais inflationniste. Si la solution des règles actives a été initialement suggérée par Rogoff (1985), Taylor (1993)⁴ est le premier auteur à avoir formulé une règle monétaire active exprimée directement en fonction du taux d'intérêt à court terme, l'instrument privilégié de la politique monétaire.

2.2 Règle de Taylor

Dans son article de 1993, Taylor ne propose pas de dérivation théorique de sa règle monétaire. Il présente simplement le modèle suivant comme règle monétaire supposée adaptée ou optimale pour les États-Unis

$$i_t = 2.2 + \pi_t + 0.5(\pi - 2) + 0.5y_t \quad (5.1)$$

où i_t désigne le taux d'intervention de la banque centrale, π_t le taux d'inflation et y_t l'output gap, c'est-à-dire l'écart relatif entre le PIB et le PIB potentiel de l'économie. La constante 2.2 est supposée représenter la croissance tendancielle de l'économie américaine sur longue période. Elle s'interprète également comme la croissance potentielle de l'économie, c'est-à-dire la croissance soutenable sans inflation excessive. 2 est le niveau d'inflation cible de l'économie officiellement déclaré par la FED.

Le paramètre 0.5 affectant à la fois l'écart à la cible d'inflation et l'output gap est obtenu par le biais de simulations réalisées sur un modèle macroéconomique estimé de l'économie américaine. Taylor cherche précisément un couple de paramètres qui permet au vue de ces simulations dynamiques de stabiliser l'inflation au niveau désiré, tout en assurant une croissance régulière du PIB autour du potentiel. Cependant comme il apparaît toujours possible de diminuer la variance du PIB au prix d'une inflation plus forte et réciproquement, il finit par déterminer ces deux paramètres parmi un sous ensemble de paramètres optimaux comme étant ceux correspondant le mieux au comportement historique de la banque centrale américaine. Il y a donc finalement derrière cette règle de Taylor un critère normatif de stabilisation macroéconomique (activité et inflation) et un critère d'adéquation empirique.

² Rogoff K., 1985, « The optimal degree of commitment to a monetary target », *Quarterly Journal of Economics*, 100, 1169-1990.

³ Kydland F., Prescott E., 1977, « Rules rather than discretion : the inconsistency of optimal plans », *Journal of Political Economy*, 45(3).

⁴ Taylor J.B., 1993, « Discretion versus policy rules in practice », *Carnegie Rochester Conference Series on Public Policy*, 39,195-214.

Le statut théorique de la règle de Taylor a été par la suite affiné par plusieurs contributions de Svensson (1997, 1999)⁵ où il montre que cette règle peut apparaître comme la solution d'un programme de contrôle optimal dans un cadre standard « d'inflation targeting » : la banque centrale cherche à stabiliser l'inflation à un niveau cible et à limiter les fluctuations cycliques de l'activité en stabilisant le produit macroéconomique autour de son niveau potentiel. Le modèle macroéconomique retenu est le modèle canonique généralement utilisé pour évaluer les propriétés des règles monétaires (Taylor (1998))⁶. Il est en fait représentatif de « la nouvelle synthèse néoclassique » qui s'affirme progressivement depuis le milieu des années 1990. Il s'agit précisément d'un modèle d'économie fermée, de courte période ou de cycle en ce sens que le produit macroéconomique est appréhendé par l'output gap, c'est-à-dire de nouveau par l'écart entre le produit et le produit potentiel exogène dont l'évolution relève du processus de croissance de long terme de l'économie.

La fonction d'offre agrégée est donnée par l'équation de Phillips suivante

$$\pi_{t+1} = \pi_t + \alpha_y y_t + \epsilon_{t+1} \quad (5.2)$$

avec $\alpha_y > 0$, π_t l'inflation à la période t (année, trimestre), y_t l'output gap et ϵ_t un choc sur les coûts de production des entreprises supposé *iid* (identiquement distribué) de moyenne nulle et de variance σ_ϵ^2 .

La demande agrégée, exprimée par rapport à l'output gap, est donnée par

$$y_{t+1} = \beta_y y_t - \beta_r (i_t - \pi_{t+1/t} - \bar{r}) + \eta_{t+1} \quad (5.3)$$

avec $\beta_y \geq 0$, $\beta_r > 0$, i_t est le taux d'intérêt nominal à court terme et l'instrument de politique monétaire de la banque centrale. $\pi_{t+1/t}$ est l'inflation anticipée en t pour la période $t+1$. \bar{r} est le taux d'intérêt réel d'équilibre ou encore neutre. η_t est un choc de demande *iid* de moyenne nulle et de variance σ_η^2 .

Les mécanismes de transmission de la politique monétaire sont ici très simples. Ils transitent par la sensibilité de la demande agrégée au taux d'intérêt réel ou plus exactement par l'écart entre le taux réel et le taux réel neutre. L'inflation anticipée pour la période $t+1$ étant dans ce modèle prédéterminée à la date t , toute modification du taux court a bien un effet sur le taux d'intérêt réel de la période.

La fonction objectif de la banque centrale est une fonction de perte intertemporelle

$$L_t = E_t \left(\sum_{\tau=1}^{\infty} \delta^\tau l_{t+\tau,t} \right) \quad (5.4)$$

⁵ Svensson L.E., 1997, « Inflation forecast targeting : implementing and monitoring inflation targets », *European Economic Review*, 4, 1111-1146.

Svensson L.E., 1999, « Inflation targeting as monetary policy rule », *Journal of Monetary Economics*, 43, 607-654.

⁶ Taylor J.B., 1998, « The Robustness and efficiency of monetary policy rules as guidelines for interest rate setting by the european central bank », *Institute for international economics studies*, août.

avec $l_{t+\tau,t}$ la fonction de perte pour la période $t + \tau$ anticipée en t telle que

$$l_t = \frac{1}{2} \left[(\pi_t - \hat{\pi})^2 + \lambda y_t^2 \right] \quad (5.5)$$

δ est un facteur d'escompte psychologique. $\hat{\pi}$ est la cible d'inflation, λ traduit le poids accordé à la stabilisation de l'activité.

Pour déterminer la politique monétaire optimale, la banque centrale minimise L_t par rapport à i_t sous les contraintes données par le modèle dynamique de l'économie (équations (5.2) et (5.3)). Deux pratiques sont alors envisageable : la solution discrétionnaire et la solution de règle. Dans le premier cas, l'optimisation dynamique est renouvelée à chaque période pour déterminer le niveau optimal de l'instrument selon l'état du système économique. La solution de règle consiste à reproduire période après période la condition de premier ordre du programme de contrôle. Il n'y a donc pas de réoptimisation et la banque centrale s'engage à respecter cette condition d'équilibre. Exprimée par rapport à i_t , la condition d'équilibre donne la règle de politique monétaire qui s'écrit

$$i_t = \bar{r} + \pi_t + \lambda_1 (\pi_t - \hat{\pi}) + \lambda_2 y_t \quad (5.6)$$

où λ_1 et λ_2 sont des paramètres positifs dépendants des paramètres du modèle et de λ traduisant le poids accordé à la stabilisation de l'activité. Si $\lambda = 0$, λ_2 prend une valeur minimale non nulle et y_t reste présent dans la règle en tant qu'indicateur de l'inflation future. Svensson qualifie cette situation de «strict inflation targeting». Si $\lambda > 0$, le souci de stabiliser l'activité entraîne des variations moins brutales du taux court. C'est donc le cas général que Svensson qualifie de «flexible inflation targeting».

La présence de l'output gap dans la règle est donc ambiguë puisque compatible avec deux types de fonctions de perte de la banque centrale. C'est sans doute là une raison importante du succès de cette règle de Taylor qui se prête à la fois à des lectures «monétariste» («strict inflation targeting») et «keynésienne» («flexible inflation targeting») de la politique monétaire.

Une difficulté autant pratique que conceptuelle est de définir précisément \bar{r} le taux réel neutre. Il est usuel d'identifier ce taux neutre, par référence à la théorie néoclassique de la croissance, à la croissance potentielle de l'économie. On pose alors $\bar{r} = g$ ou encore $\bar{r}_t = g_t$ si le taux réel neutre et la croissance potentielle sont susceptibles de varier au cours du temps. L'écriture de la règle devient

$$i_t = g_t + \pi_t + \lambda_1 (\pi_t - \hat{\pi}) + \lambda_2 y_t \quad (5.7)$$

Quelle est finalement la lecture de la politique monétaire proposée par la règle ?

Lorsque la cible d'inflation est atteinte et lorsque l'output gap est nul, ce qui est vrai en moyenne ou sur longue période, le taux d'intérêt nominal est tel que le taux d'intérêt réel est égal au taux réel neutre, c'est-à-dire à la croissance potentielle

$$r_t = g_t + \pi_t \text{ tel que } r_t - \pi_t = g_t \quad (5.8)$$

On retrouve ici la règle d'or des modèles de croissance néo-classique, c'est-à-dire l'égalité entre le taux d'intérêt réel et la croissance potentielle de l'économie. À long terme la politique monétaire se fonde donc dans une posture de neutralité vis-à-vis de l'activité. $r_t = g_t + \pi_t$ peut donc être interprété comme un taux nominal de neutralité monétaire.

À court terme en revanche, la politique monétaire est active. Elle sera restrictive ($r_t > g_t + \pi_t$) ou accommodante ($r_t < g_t + \pi_t$) selon le signe et les valeurs prises par l'écart à l'inflation cible et l'output gap. La sensibilité du taux d'intérêt à l'output gap traduit le caractère contracyclique de la politique monétaire.

Enfin on note que si la dérivation théorique des règles monétaires optimales se fait le plus souvent sur la base d'un modèle macroéconomique relativement simple tel que celui par donné (5.2) et (5.3), l'identification des paramètres optimaux des règles et notamment de la règle de Taylor se fait par le biais de simulations réalisées sur des modèles empiriques représentant les différentes économies. L'usage est même de recourir pour un même pays à plusieurs types de modèles. Ceci permet au total d'éprouver la robustesse des différentes spécifications de règle monétaire (Taylor (1998)).

3. ADÉQUATION DES POLITIQUES MONÉTAIRES EUROPÉENNES À LA RÈGLE

L'évaluation empirique de l'adéquation des politiques monétaires européennes à la règle de Taylor est ici proposée dans une optique spécifique. Il s'agit surtout de se doter d'une norme pour interpréter et évaluer les politiques monétaires européennes des trente dernières années. Cette période a en effet été marquée par de nombreuses évolutions tant en ce qui concerne les instruments et les objectifs affichés des politiques monétaires, que les contraintes (Système Monétaire Européen à partir de 1979) et conditions économiques générales encadrant leur mise en œuvre (préparation de l'Union Économique et Monétaire à partir de 1990).

3.1 Données

L'évaluation empirique porte sur les 6 plus gros pays de l'Union Européenne, c'est-à-dire l'Allemagne, la France, l'Italie, l'Espagne, les Pays-Bas et le Royaume-Uni. Tous ces pays ont intégré la zone euro à compter du 1^{er} janvier 1999, à l'exception du Royaume-Uni. Les séries économiques pertinentes sont extraites du serveur Datastream et rassemblées dans le fichier chapitre5.xls. Il comprend

- La série de PIB en volume (*PIB_code* dans la base avec code = DM, FR, ES, IT, NL, UK).
- L'indice des prix à la consommation (Consumer Price Index, *CPI_code* dans la base).
- La série de taux d'intérêt à court terme (*TC_code* dans la base) pour caractériser l'évolution de la politique monétaire. Le taux retenu est le taux au jour le jour sur le marché interbancaire et non pas directement le taux

d'intervention de la Banque Centrale dont le profil d'évolution, par pallier au gré des gestes de politique monétaire, est moins adapté à un traitement économétrique. Ces données sont de fréquences trimestrielles et couvrent pour la plupart des pays et des séries la période 1970 : 1 - 1998 : 3.

3.2 Construction des variables d'intérêt : PIB potentiel, output gap et croissance potentielle

3.2.1 La notion de PIB potentiel

Le PIB potentiel est couramment défini dans la littérature comme la production macroéconomique réalisable sans une accélération de l'inflation (ou plus exactement des prix) au delà de son niveau courant. C'est la production qui résulterait d'une sollicitation normale du potentiel d'offre de l'économie. Elle est donc associée à un certain niveau d'utilisation des facteurs de production. On considère généralement d'une part, la moyenne historique du taux d'utilisation des capacités de production (TUC moyen) pour déterminer le niveau non inflationniste d'utilisation du stock de capital et d'autre part, un niveau d'emploi compatible avec un taux de chômage équivalent au taux de chômage naturel ou NAIRU.

La détermination du PIB potentiel d'une économie pose donc deux difficultés. Il faut tout d'abord se donner une fonction de production macroéconomique, c'est-à-dire identifier les facteurs de production, le travail, capital et d'éventuels facteurs additionnels (énergie...), une spécification de la fonction de production (Cobb-Douglas, CES...) et estimer économétriquement les paramètres de la fonction. Il faut en second lieu, et ce n'est pas la moindre des difficultés, déterminer le NAIRU de cette économie ce qui suppose d'estimer une équation de salaire robuste (courbe de phillips) pour déterminer ensuite le taux de chômage n'accéléralant pas l'inflation. C'est donc là un travail considérable très rarement réalisé par les organismes de conjoncture et de prévision qui préfèrent en général une estimation du PIB potentiel à partir de méthodes purement statistiques de lissage ou encore de filtrage de la série de PIB tel que par exemple le filtre de Hodrick et Prescott.

3.2.2 Évaluation par la méthode de filtrage de Hodrick et Prescott

La méthode de filtrage de Hodrick et Prescott (1980) est aujourd'hui très utilisée en macroéconomie appliquée pour évaluer le PIB potentiel d'une économie, calculer son output gap et sa croissance potentielle. Sa vocation initiale est pourtant différente. C'est au départ avant tout une méthode de décomposition entre cycle et tendance, c'est-à-dire une méthode d'extraction des composantes tendancielle et cyclique d'une série économique. Elle s'est progressivement substituée aux méthodes de régression sur des polynômes du temps présentant l'inconvénient de supposer une tendance déterministe fonction du temps et une composante cyclique aléatoire (le résidu de la régression). Or il est admis depuis la contribution fondamentale de

Nelson et Plosser (1982)⁷ que la composante tendancielle affichée par la plupart des séries économiques est de type aléatoire et non de type déterministe (cf. chapitre 14, Économétrie des séries non stationnaires).

Formellement si y_t^o représente la série de PIB en volume (en logarithme), on pose la décomposition suivante

$$y_t^o = \bar{y}_t + y_t \quad (5.9)$$

où \bar{y}_t désigne la composante tendancielle du PIB et y_t la composante cyclique ou conjoncturelle. Le partage du PIB entre ces deux composantes est obtenu selon le filtre de Hodrick et Prescott par la résolution d'un programme d'optimisation. On minimise par rapport à la chronique des \bar{y}_t une somme pondérée de la variance des écarts conjoncturels et de la variance des changements de rythme de la croissance potentielle. Soit

$$\underset{\{\bar{y}_t\}}{\text{Min}} \left[\sum_{t=1}^T y_t^2 + \lambda \sum_{t=3}^T (\Delta \bar{y}_t - \Delta \bar{y}_{t-1})^2 \right]$$

$$S.C \ y_t^o = \bar{y}_t + y_t$$

Les valeurs extrêmes $\lambda = 0$ et $\lambda \rightarrow \infty$ correspondent respectivement à une composante conjoncturelle nulle ($\bar{y}_t = y_t^o, \forall t$), où le PIB et le PIB tendanciel sont confondus, et à une tendance linéaire, c'est-à-dire une croissance tendancielle à taux constant ($\Delta \bar{y}_t = \Delta \bar{y}, \forall t$). Sur données trimestrielles, la valeur le plus souvent retenue pour λ est 1600, ce qui correspondant au fait de donner la même pondération à un écart conjoncturel de 5 % et à une accélération de la croissance tendancielle de 1/8 de point sur un trimestre.

Le cycle macroéconomique selon ce type d'évaluation est donné par l'évolution de l'écart conjoncturel ou output gap, même si dans les conceptions courantes des conjoncturistes il correspond plutôt au profil d'évolution de la série de taux de croissance du PIB. La composante tendancielle du PIB est interprétée par les macroéconomistes comme la résultante d'influences ou de chocs ayant des effets permanents ou durables sur le PIB. À l'opposé la composante conjoncturelle est interprétée comme la résultante d'influences ou de chocs ayant des effets transitoires. Dès lors l'assimilation de la composante tendancielle du PIB au PIB potentiel s'avère d'autant plus légitime que ces effets permanents sont attribués à des chocs d'offre, c'est-à-dire affectant les comportements d'offre des entreprises (fonction de production, prix relatifs des facteurs...).

3.3 Évaluation empirique

L'évaluation de l'adéquation des politiques monétaires européennes à la règle de Taylor peut se concevoir en fait de deux façons. Il est envisageable tout d'abord

⁷ Nelson J., Plosser C., 1982, « Trend and random walk in macroeconomic time series : some evidence and implications », *Journal of Monetary Economics*, 139-169.

de se référer à la règle stricto sensu en considérant les différents coefficients affectant les variables de la règle comme donnés et notamment ceux décrivant la dimension active de la politique monétaire ($\lambda_1 = \lambda_2 = 0.5$). L'examen de la conformité à la règle se fait alors par de simples simulations algébriques en comparant les taux courts historiques aux taux de Taylor simulés, c'est-à-dire calculés.

Alternativement cet exercice peut s'inscrire dans un cadre économétrique classique en considérant la règle de Taylor comme un modèle à plusieurs variables, c'est-à-dire une spécification, dont il convient d'estimer les coefficients. L'hypothèse d'une stricte conformité à la règle est alors levée et plusieurs possibilités sont envisageables selon que l'on laisse libre, c'est-à-dire déterminés par les estimations économétriques, toute ou partie de ces coefficients.

Problème 5.1

Évaluation par simulation

Les différentes questions sont à traiter successivement pour chaque pays.

1. Construction de l'ensemble des variables d'intérêt du modèle.
 - a) Définir la série de croissance comme le glissement annuel en % (*PIBGA_code*) du PIB. Cette série de croissance n'intervient pas dans la règle mais peut être comparée à la série de croissance potentielle et à celle de l'output gap.
 - b) Définir la série d'inflation comme le glissement annuel en % (*CPIGA_code*) de l'indice des prix à la consommation.
 - c) Définir l'écart à l'inflation cible (*ECINF_code*) en retenant un objectif d'inflation unique pour l'ensemble des pays et constant à 2 % sur toute la période.
 - d) Construire la série de PIB potentiel (*HP_code*) en appliquant la procédure HPFILTER.SRC et définir la croissance potentielle comme le glissement annuel en % (*HPGA_code*) du PIB potentiel et l'output gap (*GAP_code*) comme l'écart en % entre le PIB et le PIB potentiel rapporté au PIB potentiel.

Représenter graphiquement les séries de PIB et de PIB potentiel, de croissance et de croissance potentielle, d'inflation et d'output gap.

2. À partir des équations (5.7) et (5.8), calculer les séries de taux Taylor (*TS_code*) et de taux neutre (*TN_code*) et comparer à l'aide d'un graphique ces deux séries aux taux courts (*TC_code*).
3. Quel type d'indicateur peut-on utiliser pour appréhender la précision de ce modèle explicatif des taux courts ?

Problème 5.2

Évaluation par estimation et tests sur longue période

Le modèle de Taylor peut être lu à partir de l'équation (5.7) comme un modèle linéaire où le vecteur des paramètres β serait prédéterminé avec $\beta' = (1, 1, 0.5, 0.5)$. L'estimation de (5.7) par les MCO doit montrer si les coefficients qui restituent au mieux le comportement des banques centrales correspondent à ces coefficients

prédéterminés. On choisit ici, afin de ne pas dénaturer la règle et de pouvoir respecter le principe de neutralité à long terme de figer à 1 les deux premiers coefficients, de la croissance potentielle et de l'inflation, dont la somme donne le taux nominal de neutralité monétaire.

1. Ce modèle contraint peut être estimé par les MCO en transformant la variable endogène de départ, le taux court (*TC_code*), en une variable d'écart au taux neutre (*EC_code*). On pose donc, en introduisant une constante, le modèle économétrique suivant

$$EC_t = TC_t - TN_t = \alpha + \lambda_1 ECINF_t + \lambda_2 GAP_t + u_t$$

Estimer ce modèle transformé par les MCO et récupérer la série des taux courts estimés (*EC_code*).

2. Que peut représenter d'un point de vue économique le coefficient de la constante ?
3. Le coefficient de détermination de la régression vous semble-t-il adapté pour juger du pouvoir explicatif du modèle de Taylor vis-à-vis des politiques monétaires ?
4. On se propose maintenant de réaliser plusieurs tests sur les paramètres λ_1 et λ_2 décrivant la dimension active de la règle.

Le test de significativité de Student est directement fourni par le logiciel et permet d'obtenir à la fois la réalisation de la T-stat et le risque de rejeter à tort l'hypothèse nulle associée à cette valeur de T-stat (variable Signif, souvent appelée la P-value du test). Interpréter les résultats puis tester successivement les hypothèses nulles $\lambda_1 = 0.5$ puis $\lambda_2 = 0.5$.

Tester, finalement

$$H_o : (\lambda_1, \lambda_2) = (0.5, 0.5)$$

$$H_a : (\lambda_1, \lambda_2) \neq (0.5, 0.5)$$

Construire un tableau de synthèse et conclure (en utilisant l'ensemble des résultats des problèmes 5.1 et 5.2).

Problème 5.3

Analyse de la stabilité

On souhaite maintenant examiner la question de l'adéquation des politiques monétaires européennes à la règle de Taylor de manière plus dynamique par le biais d'estimations sur plusieurs périodes de référence. Cela revient en fait à remettre en cause l'hypothèse d'une stabilité du modèle de comportement des autorités monétaires.

1. Estimations par sous-périodes
 - a) Estimer le modèle de Taylor sur trois périodes de référence de 10 ans : (70.1-79.4, 80.1-89.4, 88.1-98.3) et sur la période débutant avec la phase 1 de l'UEM

(90.1-98.3). Comme précédemment, chaque jeu d'estimation sera complété des tests de Student et de Fisher portant sur l'égalité à 0.5 des coefficients λ_1 et λ_2 .

- b) Construire les tableaux de synthèse de ces estimations et conclure.
2. Réaliser sur l'ensemble de la période le test de stabilité du CUSUM fondé sur des estimations et résidus récurrents du modèle. Interpréter les résultats. Procéder ensuite à des estimations glissantes du modèle sur des périodes de 10 ans. La première période d'estimation sera 70.1-79.4 et les dates de début et de fin de période augmenteront d'un trimestre à chaque itération. On récupérera à chaque estimation les coefficients estimés et leur écart type. Représenter graphiquement l'évolution des coefficients estimés encadrés par un intervalle de confiance à \pm deux fois l'écart type. Conclure sur l'évolution des politiques monétaires et leur éventuelle convergence.

4. ÉLÉMENTS DE CORRIGÉ

Problème 5.1

Évaluation par simulation

1. La croissance d'une série de données trimestrielles x_t en glissement annuel en % est calculée de la manière suivante

$$\left(\frac{x_t - x_{t-4}}{x_{t-4}} \right) \cdot 100$$

C'est cette formule qui est utilisée ici.

Pour calculer le PIB potentiel et l'output gap il faut mettre en œuvre le filtre de Hodrick et Prescott sur le logiciel Winrats s'obtient par la procédure HP-FILTER (fichier source HPFILTER.SRC). Par exemple, son application sur la série de PIB allemand donne les résultats représentés dans la figure 5.1. On remarquera que si le PIB tendanciel-potentiel (noté HP_dm) apparaît bien comme un lissage de la série de PIB, il en va de même de la série de croissance potentielle, définie comme le glissement annuel de la série de PIB potentiel, vis-à-vis de la croissance. Ainsi, toute élévation durable du niveau de la croissance ($PIBGA_dm$) se traduit par une élévation de la croissance potentielle ($HPGA_dm$) et réciproquement. La figure 5.1 c reproduit l'évolution de l'output gap (GAP_dm) conjointement à celui de l'inflation ($CPIGA_dm$). On identifie ainsi pour le cas de l'Allemagne le profil d'évolution de l'ensemble des variables qui déterminent selon la règle de Taylor l'évolution des taux d'intérêt.

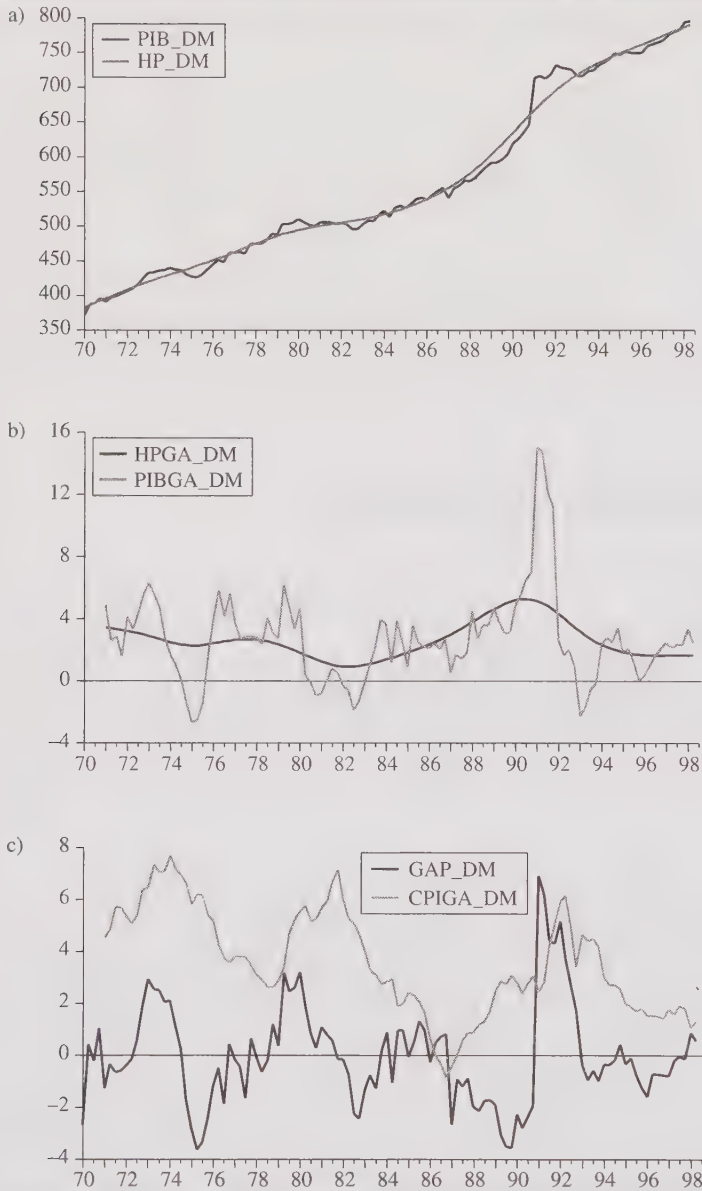


Figure 5.1 Allemagne — variables d'intérêt. a) PIB et PIB potentiel, b) croissance et croissance potentielle, c) inflation et output gap

2. La comparaison graphique des taux simulés (taux Taylor et taux neutre) et des taux historiques est représentée figure 5.2.

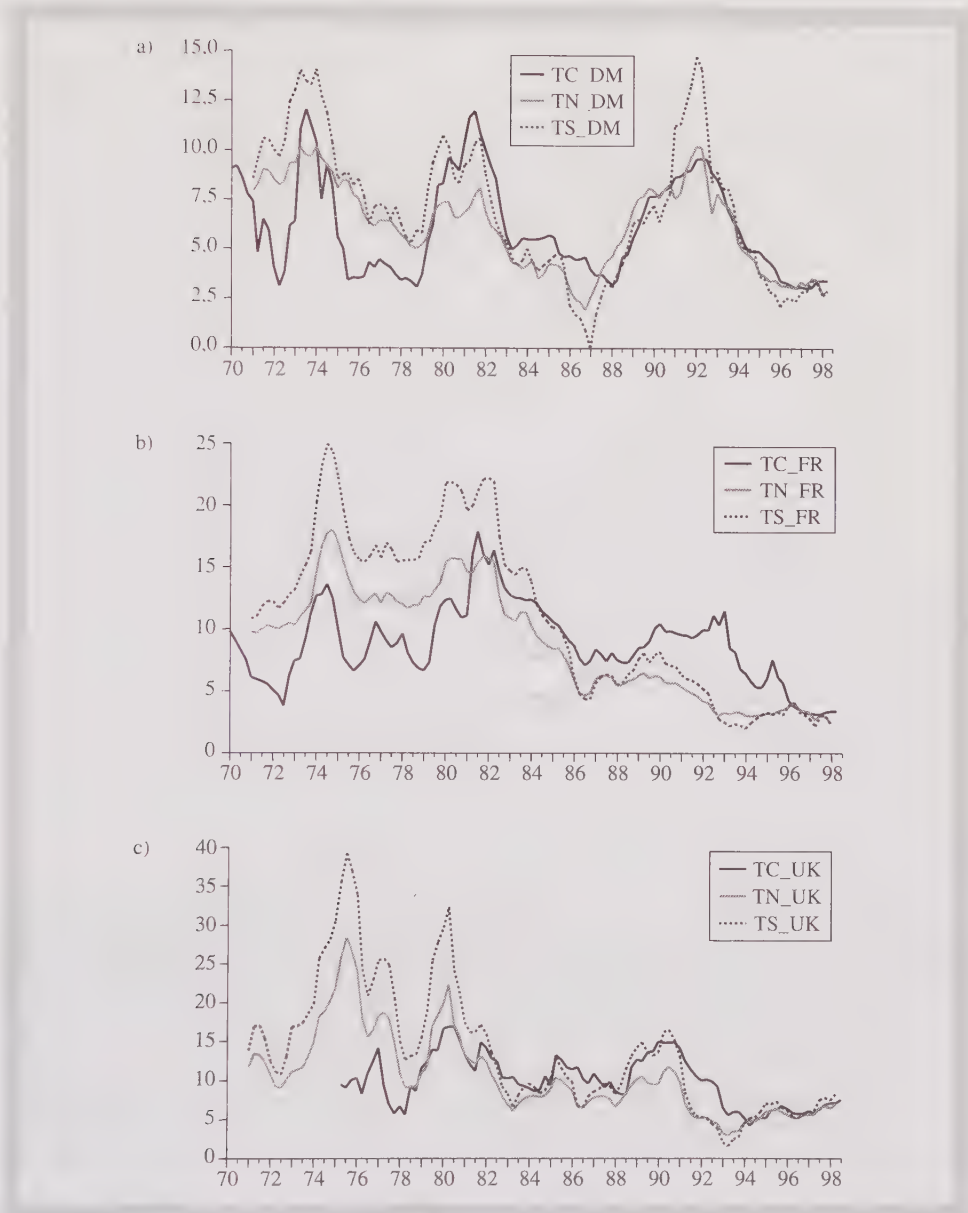


Figure 5.2 Taux courts, taux neutres et taux de Taylor. a) Allemagne, b) France, c) Royaume-Uni

La figure 5.2 fournit trois principales indications.

- i) Dans l'ensemble l'adéquation des politiques monétaires à la règle de Taylor semble plutôt médiocre : les écarts entre les taux observés et simulés sont élevés et le sens d'évolution des taux est moyennement reproduit par le modèle.

- ii) Il semble en fait que la règle de Taylor soit davantage adaptée au cas Allemand.
- iii) De plus l'adéquation des politiques monétaires à la règle semble très variable au cours du temps. Pour l'ensemble des pays à l'exception des Pays-Bas, la précision du modèle aurait tendance à s'accroître sur la période. On assiste globalement à une convergence des taux courts vers les taux Taylor avec des degrés et des rythmes variables selon les pays et une accélération en fin de période.

Cette convergence s'effectue en outre selon deux régimes. Jusqu'en 1980, c'est-à-dire grossièrement jusqu'à la création du SME (mars 1979), les taux courts sont nettement inférieurs aux taux de Taylor et aux taux neutres : les politiques monétaires sont alors clairement accommodantes. Il apparaît en particulier que les deux chocs pétroliers de 1974 et 1979 auraient nécessité des réactions monétaires plus vigoureuses, c'est-à-dire des hausses de taux plus importantes. À partir du début des années 1980, la situation a tendance à s'inverser avec des politiques monétaires qui deviennent dans l'ensemble restrictives. Les taux courts sont supérieurs aux taux de Taylor et aux taux neutres et les écarts entre ces derniers se réduisent au gré de la baisse des taux d'inflation. On a là finalement, et notamment pour le cas de la France, une illustration simple des stratégies de désinflation compétitive menée par plusieurs pays européens au cours de cette période.

Deux situations particulières méritent d'être précisées. Dans le cas de l'Allemagne tout d'abord, l'adéquation vis-à-vis de la règle s'affirme nettement à partir de 1988 avec un degré de précision du modèle qui semble à peu près constant sur la période 1988-1998, si l'on excepte les deux années ayant suivi la réunification : la hausse des taux Taylor est alors en grande partie due à une hausse artificielle de l'output gap résultant d'une rupture dans la définition du PIB qui devient pan-allemand à partir de 1990.

La cas du Royaume-Uni est également intéressant dans la mesure où l'on observe une convergence quasi parfaite des taux courts vers les taux de Taylor (TS et TN) à partir de 1994, c'est-à-dire environ un an après la sortie de l'économie britannique du SME en septembre 1992.

- 3. Pour dépasser ces appréciations qualitatives sur l'adéquation des politiques monétaires à la règle, il faut construire un indicateur définissant le niveau de précision du modèle. Il est usuel, dès lors que l'on raisonne sur un modèle qui n'est pas estimé où lorsque l'on évalue un modèle estimé sur une autre période que sa période d'estimation, de se référer à l'erreur quadratique moyenne ou $RMSE$ (Root Mean Square Error). Soit $e_t = i_t - \hat{i}_t$, l'écart entre le taux court observé et le taux simulé, le $RMSE$ est donné par

$$RMSE = \sqrt{\frac{\sum_{t=1}^T e_t^2}{T}}$$

Par rapport à la définition des variables dans le fichier de données et le programme, on note $EC = TC - TN$ et $ER = TC - TS$.

Cet indicateur est très proche de l'écart type des résidus d'un modèle estimé où par construction les résidus sont de moyenne nulle. La somme des erreurs au carré est

ici divisée par le nombre de périodes et non pas par le nombre de degré de liberté comme dans un modèle estimé.

Pour disposer d'une information sur l'évolution dans le temps de la précision du modèle, les statistiques ont été calculées sur des périodes glissantes d'une durée arbitrairement fixée à 3 ans (12 trimestres). L'exercice est réalisé à la fois pour le taux neutre et la version complète de la règle de Taylor (figure 5.3).

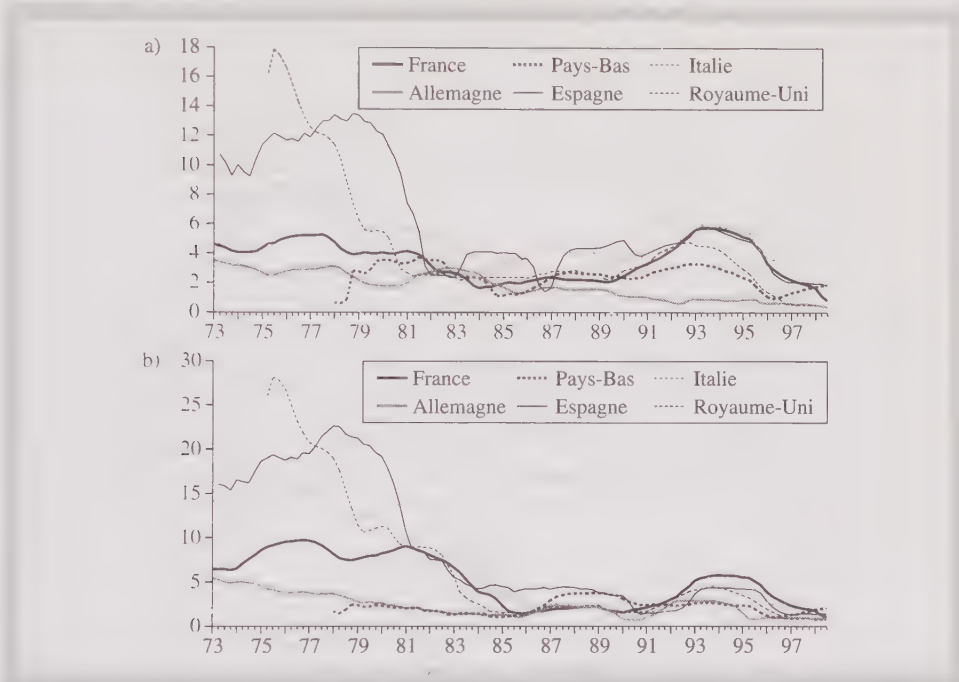


Figure 5.3 Indicateurs de précision — règle de neutralité. a) Précision de la règle de Taylor simulée — $\lambda_1 = \lambda_2 = 0.0$ (RMSE sur 12 trimestres), b) Précision de la règle de Taylor simulée — $\lambda_1 = \lambda_2 = 0.5$ (RMSE sur 12 trimestres)

Ces calculs confirment les indications fournies par les premiers graphiques. La précision est plus élevée dans le cas de l'Allemagne, elle s'accroît au cours du temps et particulièrement en fin de période. Il apparaît en outre que pour la plupart des pays la règle de neutralité monétaire est aussi précise que la version complète de la règle de Taylor.

Problème 5.2

Évaluation par estimation et tests sur longue période

1. Le modèle contraint estimé donne \widehat{EC}_t (noté *ECS_code* dans le programme)

$$\widehat{EC}_t = \hat{\alpha} + \hat{\lambda}_1 ECINF_t + \hat{\lambda}_2 GAP_t$$

On souhaite obtenir la série des taux courts estimés selon le modèle de Taylor.

Or $\widehat{EC}_t = TC_t - TN_t$, soit encore puisque TN_t est connue :

$\widehat{EC}_t = TC_t - TN_t$, ce qui donne : $\widehat{TC}_t = \widehat{EC}_t + TN_t$ (on le note *ECS_code*)

dans le programme). Le taux court estimé est donc logiquement obtenu en ajoutant le taux neutre à la variable d'écart estimée.

2. Il est toujours préférable d'introduire une constante dans la définition du modèle économétrique puisque l'absence de cette variable est de nature à biaiser les estimateurs des coefficients et celui de la variance des résidus. Le coefficient de la constante n'a pas a priori ici d'interprétation économique particulière dans la mesure où cette variable est exclue du modèle théorique. Il pourra toutefois être interprété comme le niveau moyen d'une ou plusieurs variables manquantes ou comme une erreur de mesure systématique sur certaines variables. En particulier en posant une cible d'inflation constante à 2 % sur l'ensemble de la période, l'écart à l'inflation cible est sans doute surévalué en moyenne. Un coefficient négatif (respectivement positif) pour la constante pourra donc s'interpréter raisonnablement comme un objectif d'inflation supérieur (inférieur) à 2 % ou encore comme un biais accommodant (restrictif) de la politique monétaire.
3. Le coefficient de détermination du modèle (R^2) mesure la part de la variance des écarts au taux neutre expliquée par le modèle, c'est-à-dire en fait l'apport spécifique des variables $ECINF$ et GAP dans l'explication des taux courts. Si l'essentiel de la politique monétaire est expliqué par l'évolution du taux neutre, l'apport des ces variables sera faible, sans que le modèle soit pour autant mauvais. Cette statistique ne peut donc pas ici, compte tenu de la transformation effectuée sur la variable endogène, être utilisée pour juger de la qualité des ajustements. La seule possibilité est d'utiliser la statistique d'écart type estimé des résidus. En effet

$$\hat{\sigma}_{\hat{u}} = \sqrt{\frac{\sum_{t=1}^T \hat{u}_t^2}{T-3}}$$

avec

$$\hat{u}_t = EC_t - \widehat{EC}_t$$

c'est-à-dire

$$\hat{u}_t = TC_t - TN_t - (\widehat{TC}_t - \widehat{TN}_t) = TC_t - \widehat{TC}_t$$

Le résidu de la régression est bien l'écart entre le taux court et le taux estimé. $\hat{\sigma}_{\hat{u}}$ reste donc pertinent pour juger de la précision du modèle.

La figure 5.4 représente les taux courts, taux de Taylor et taux estimés pour l'Allemagne, la France et le Royaume-Uni. Paradoxalement, à la lecture des figures 5.2 et 5.4, ces versions estimées, c'est-à-dire optimisées, de la règle de Taylor n'apparaissent guère supérieures à la simple simulation de la règle. La détermination de coefficients optimaux pour l'ensemble de la période semble contraignante et ne permet pas de retrouver la convergence des taux courts vers les taux Taylor observée dans l'exercice précédent. L'hypothèse d'une instabilité des paramètres se trouve donc renforcée. L'autocorrélation des résidus constatée pour la plupart des pays (test de Durbin-Watson) traduit bien un problème général de spécification du modèle tel qu'une ou plusieurs variables manquantes ou plus radicalement un phénomène d'instabilité des

paramètres. On pourrait vérifier du reste que la procédure courante de correction de l'autocorrélation par l'estimation d'un modèle quasi-différencié n'est pas ici pour la plupart des pays efficace, ce qui montre bien que le caractère radical du problème. Les estimateurs des coefficients restent dans ce contexte sans biais mais ne sont plus de variance minimale. Les tests qui comme le test de student utilisent l'écart type estimé des estimateurs n'auront donc pas la précision souhaitée.

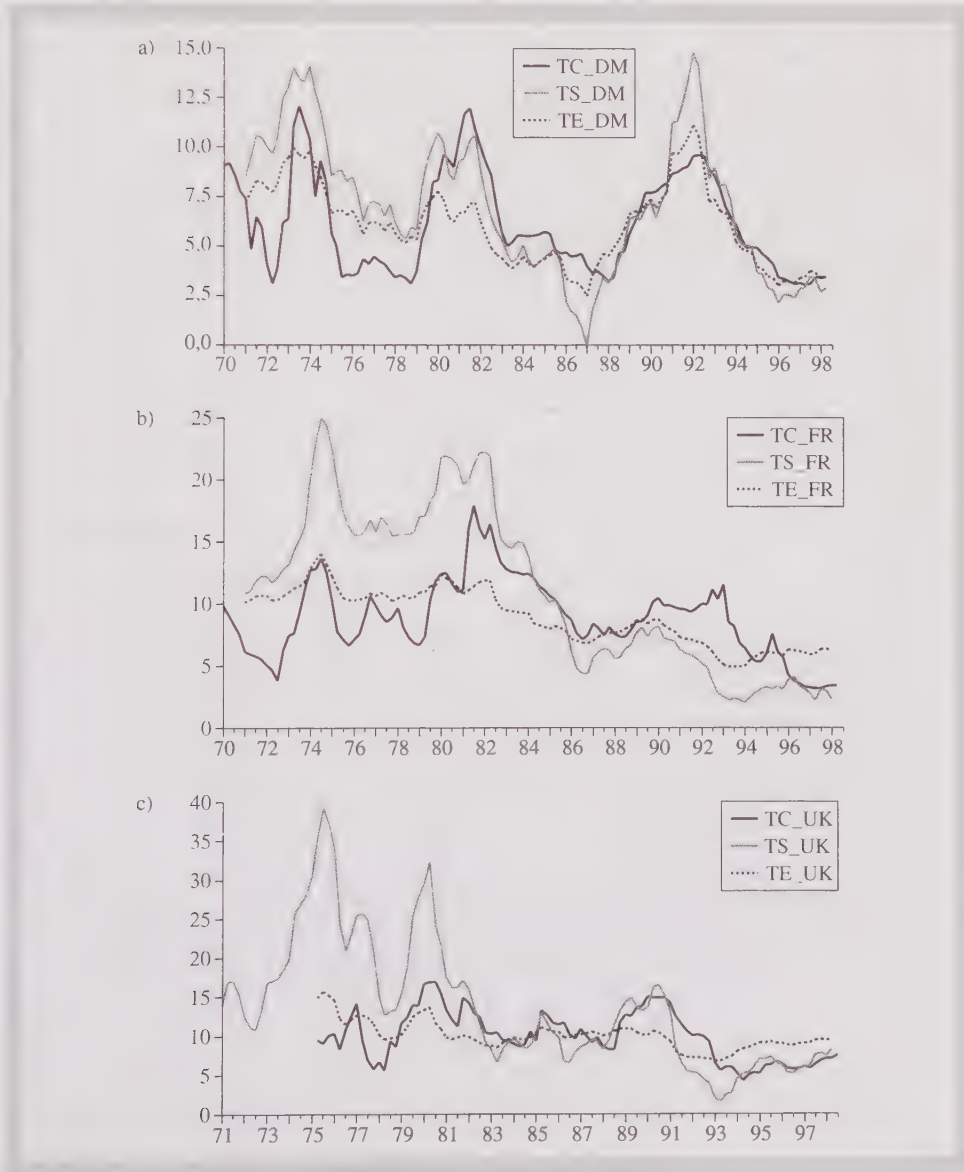


Figure 5.4 Taux courts, taux de Taylor et taux estimés. a) Allemagne, b) France, c) Royaume-Uni

La valeur des écarts types estimés confirme que le modèle est plus adapté aux cas de l'Allemagne des Pays-Bas et de la France. Le test de significativité globale de Fisher permet de conclure pour les 6 pays à l'apport des variables *ECINF* et *GAP* dans l'explication des taux courts (variable Signif < 0.05).

La constante est significative (variable Signif < 0.05) pour tous les pays à l'exception de l'Allemagne. Les coefficients sont élevés (de 1.42 à 6.04) et traduisent sans doute le biais restrictif moyen des politiques monétaires de ces pays au cours de la période.

À l'exception des Pays-Bas les paramètres λ_1 sont négatifs et significativement différents de 0 au seuil de 5 % (variable Signif < 0.05), ce qui est en contradiction avec les fondements théoriques du modèle. Ces coefficients sont toutefois tous supérieurs à -1, ce qui permet de maintenir une indexation positive des taux courts sur l'inflation. En revanche, le paramètre λ_2 ressort bien dans l'ensemble avec le signe positif attendu. Toutefois à l'exception des cas allemand et néerlandais, les estimateurs sont peu précis (écarts types élevés) de sorte qu'on ne peut pas rejeter pour un seuil de 5 % l'hypothèse de nullité. Pour la France, l'Espagne, l'Italie et le Royaume-Uni, ces coefficients ne peuvent pas être considérés comme significativement différents de 0.5 en raison de l'imprécision des estimateurs. Dans le cas des Pays-Bas le coefficient est réellement proche de 0.5.

Compte tenu de ces estimations, l'hypothèse nulle du test de Fisher d'égalité du couple de coefficients à 0.5 est rejetée pour les 6 pays (variable Signif < 0.05). La dimension active de la politique monétaire n'est donc pas selon ces évaluations conforme à la règle de Taylor.

Problème 5.3

Analyse de la stabilité

1. L'estimation du modèle par sous-période (tableaux 5.2, 5.3, 5.4, 5.5) confirme bien dans l'ensemble les hypothèses d'instabilité des coefficients formulées précédemment. L'instabilité concerne aussi bien le coefficient de la constante plutôt négatif durant les années 1970 (biais accommodant des politiques monétaires) et positif sur les périodes suivantes, que les paramètres de réaction à l'écart d'inflation et à l'output gap. On note de plus avec intérêt la situation allemande sur les deux dernières périodes d'estimation. La baisse de l'écart type des résidus traduit la plus grande précision du modèle. La constante et λ_1 ne sont pas significativement différents de 0. Le coefficient λ_2 estimé à 0.12 traduit une réaction très faible à l'output gap. Ces résultats révèlent en fait une forme de convergence de la politique monétaire allemande vers une règle de neutralité. C'est donc une confirmation des résultats obtenus lors de l'évaluation par simulation.
2. Test du CUSUM et régressions glissantes.

Le test formel d'instabilité du modèle, c'est-à-dire de l'ensemble de ces coefficients, est donné par le test du CUSUM de Brown, Durbin et Evans. Rappelons que le test s'appuie sur une série d'estimations et de résidus récurrents où la période d'estimation n'est pas fixe mais croissante de 1, ..., r avec $r = K + 1, \dots, T$. La dernière estimation porte donc sur l'ensemble de la période de référence. La statistique du test repose sur une somme de résidus

récur­sifs normalisés W_r qui doivent sous l'hypothèse nulle de stabilité des coefficients du modèle se situer à l'intérieur du corridor $[-C_\alpha, C_\alpha]$.

Tableau 5.2 Estimation sur la première sous-période

70.1-79.4		ALL	FR	PB	ES	IT	RU
$\widehat{\sigma}_{\hat{u}}$		1.54	1.21	2.50	7.97	1.93	2.24
α	$\hat{\alpha}$	-2.53	-3.53	0.29	3.31	-2.18	1.57
	T-stat	-4.29	-5.87	0.51	0.51	-2.94	0.88
λ_1	$\hat{\lambda}_1$	0.16	-0.10	0.48	-0.74	-0.33	-0.65
	T-stat	0.93	-1.30	0.17	-1.90	-5.59	-5.41
	T-stat ($\lambda_1 = 0.5$)	-1.92	-7.69	0.00	-3.18	-13.94	-9.52
λ_2	$\hat{\lambda}_2$	0.56	0.92	1.23	-0.67	0.14	1.07
	T-stat	3.83	5.13	2.04	-0.45	0.87	3.30
	T-stat ($\lambda_2 = 0.5$)	0.42	2.36	1.22	-0.79	-2.22	1.76
Fisher	Signif	0.17	0.00	0.51	0.01	0.00	0.00

Tableau 5.3 Estimation sur la deuxième sous-période

80.1-89.4		ALL	FR	PB	ES	IT	RU
$\widehat{\sigma}_{\hat{u}}$		1.18	1.57	1.13	2.72	1.51	1.34
α	$\hat{\alpha}$	1.17	3.00	2.13	4.51	5.76	3.46
	T-stat	5.40	7.19	11.52	4.36	12.34	10.27
λ_1	$\hat{\lambda}_1$	0.30	-0.32	-0.01	-0.28	-0.40	-0.33
	T-stat	3.36	-5.36	-0.06	-2.48	-9.18	-7.02
	T-stat ($\lambda_1 = 0.5$)	-2.19	-13.64	-7.49	-6.86	-20.62	-17.37
λ_2	$\hat{\lambda}_2$	0.52	0.17	0.42	1.52	-1.25	-0.08
	T-stat	4.18	0.61	3.03	3.00	-5.84	-0.62
	T-stat ($\lambda_2 = 0.5$)	0.19	-1.17	-0.56	2.01	-8.17	-4.24
Fisher	Signif	0.09	0.00	0.00	0.00	0.00	0.00

Tableau 5.4 Estimation sur la troisième sous-période

88.1-98.3		ALL	FR	PB	ES	IT	RU
$\widehat{\sigma}_{\bar{u}}$		0.69	2.04	2.13	1.75	1.52	1.30
α	$\hat{\alpha}$	0.23	2.86	0.69	1.25	3.18	0.50
	T-stat	1.74	8.03	1.84	0.62	5.84	1.62
λ_1	$\hat{\lambda}_1$	-0.16	1.38	0.56	0.77	0.37	0.71
	T-stat	-1.78	3.05	1.11	3.91	2.04	7.35
	T-stat ($\lambda_1 = 0.5$)	-7.05	1.95	0.12	1.39	-0.70	2.24
λ_2	$\hat{\lambda}_2$	0.12	-0.64	0.63	-0.23	-0.27	-1.06
	T-stat	2.58	-1.84	1.48	-0.95	-1.03	-6.26
	T-stat ($\lambda_2 = 0.5$)	-7.69	-3.28	0.31	-3.00	-2.89	-9.20
Fisher	Signif	0.00	0.00	0.92	0.01	0.00	0.00

Tableau 5.5 Estimation sur la quatrième sous-période (début de l'UEM)

90.1-98.3		ALL	FR	PB	ES	IT	RU
$\widehat{\sigma}_{\bar{u}}$		0.58	2.08	1.88	1.80	1.61	1.22
α	$\hat{\alpha}$	0.38	3.00	-0.41	1.13	3.12	0.35
	T-stat	3.28	8.07	-0.93	1.72	5.34	1.20
λ_1	$\hat{\lambda}_1$	-0.15	1.70	2.20	0.88	0.42	0.68
	T-stat	-1.87	3.43	3.54	4.03	2.10	7.23
	T-stat ($\lambda_1 = 0.5$)	-8.02	2.42	2.74	1.74	-0.38	1.92
λ_2	$\hat{\lambda}_2$	0.05	-0.56	0.29	-0.19	-0.25	-1.30
	T-stat	1.24	-1.56	0.75	-0.75	-0.86	-7.13
	T-stat ($\lambda_2 = 0.5$)	-9.62	-2.95	-0.51	-2.73	-2.75	-9.86
Fisher	Signif	0.00	0.01	0.03	0.02	0.02	0.00

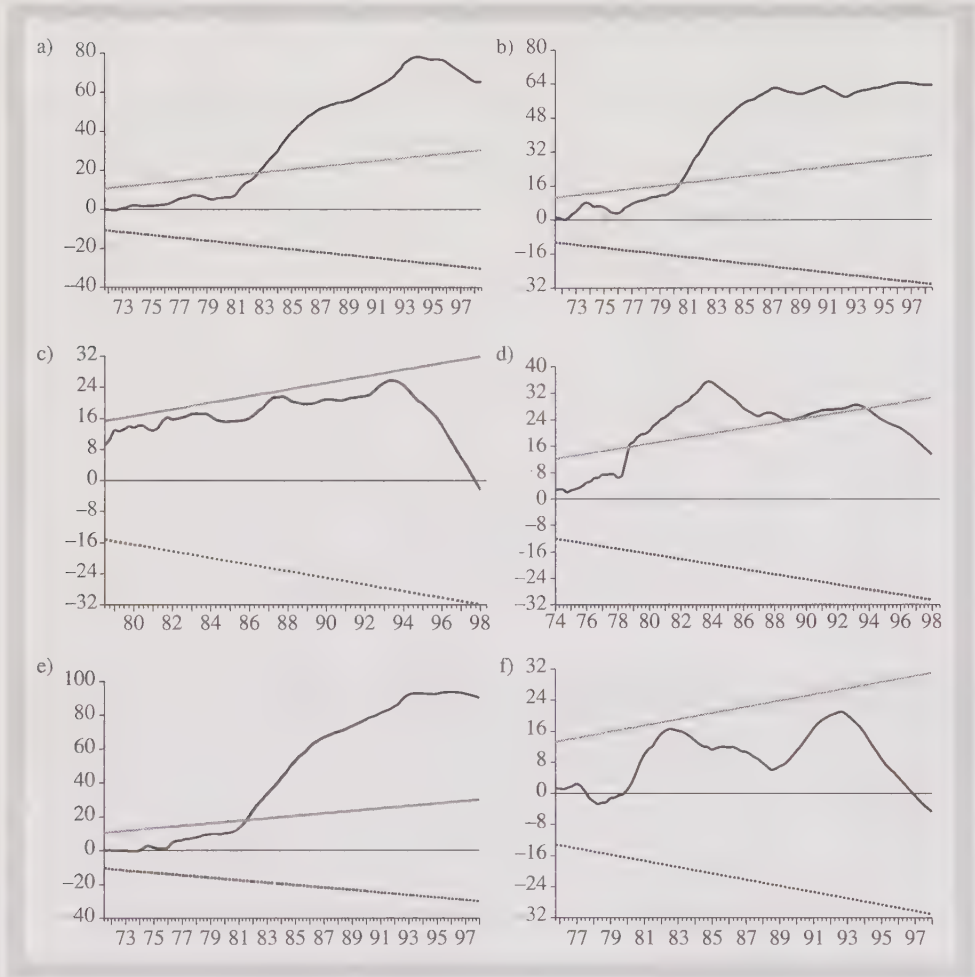


Figure 5.5 Test de stabilité des paramètres — Test du CUSUM. a) France, b) Allemagne, c) Pays-Bas, d) Espagne, e) Italie, f) Royaume-Uni

La statistique du test sort du corridor pour 4 des 6 pays étudiés. Les modèles sont donc globalement instables.

Ce résultat est confirmé par les estimations glissantes de la règle. La figure 5.6 représente pour l'Allemagne et le Royaume-Uni les résultats d'estimations des paramètres λ_1 et λ_2 .

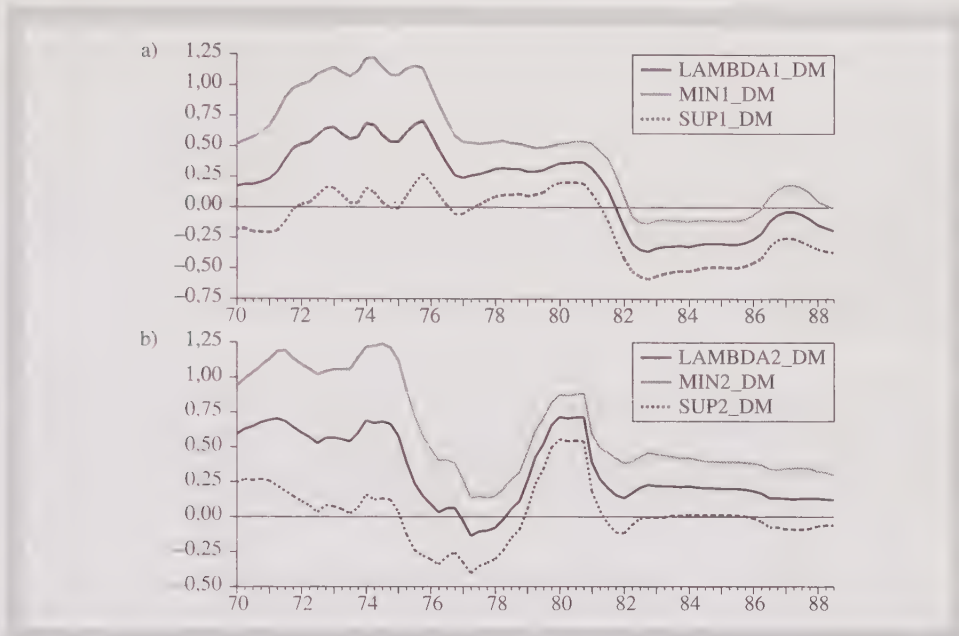


Figure 5.6 Régressions glissantes — Cas de l'Allemagne... a) Dynamique du coefficient lambda 1, b) Dynamique du coefficient lambda 2

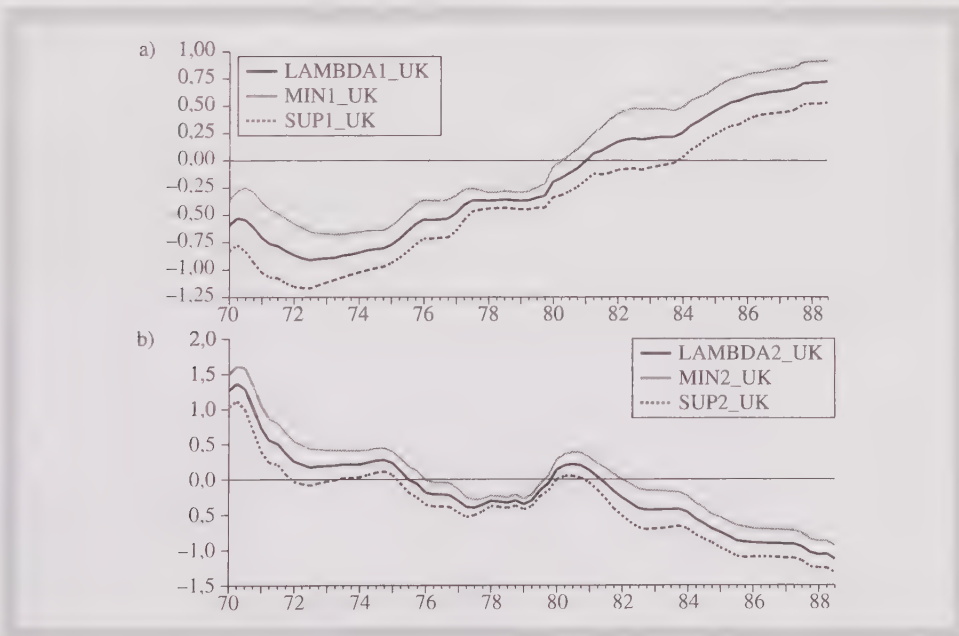


Figure 5.7 Régressions glissantes — Cas du Royaume-Uni. a) Dynamique du coefficient lambda 1, b) Dynamique du coefficient lambda 2

La dynamique des coefficients de réaction λ_1 et λ_2 présente une forte variabilité. Il n'y a pas de réelles convergences vers les valeurs de Taylor (0.5) ou vers des niveaux communs. Tout au plus peut-on constater pour les dernières estimations une tendance à la hausse vers des valeurs positives du paramètre de réaction à l'excès d'inflation par rapport à la cible de 2 %, et ce pour tous les pays à l'exception de l'Allemagne qui semble adopter depuis 1988 une politique monétaire proche de la neutralité. Il semble finalement que la convergence des politiques monétaires vers une plus grande adéquation avec la règle de Taylor soit uniquement le fait d'une plus grande conformité avec la dimension neutre de cette règle.

Quels sont finalement les enseignements principaux qui se dégagent des évaluations statistiques et économétriques proposées dans ce chapitre ?

À l'exception de l'Allemagne et du Royaume-Uni à partir de 1993, les principaux pays européens ont pratiqué des politiques monétaires assez éloignées de ce qu'aurait donné une règle active comme la règle de Taylor. Ce résultat n'est somme toute pas très surprenant. La règle de Taylor est récente et traduit une conception moderne de la politique monétaire. Elle ne pouvait être précisément et uniformément appliquée avant même d'être popularisée comme outil théorique. D'autre part les conditions d'exercice de la politique monétaire n'étaient pas non plus favorables à une forte adéquation avec ce type de règle monétaire (évolution des instruments, contraintes du SME, préparation de l'UEM). Au total, les résultats obtenus s'interprètent assez bien par la thèse de l'asymétrie du SME : le pays leader du système de change détermine son taux d'intérêt de manière autonome en fonction de ses propres conditions économiques, les autres pays ont tendance à suivre pour respecter les contraintes de parité du système.

Ces évaluations montrent aussi que concernant l'Allemagne, une règle de neutralité semble suffisante pour décrire le comportement de la banque centrale. La dimension active de la politique monétaire pour réduire le taux d'inflation et stabiliser l'activité n'apparaît absolument pas. En d'autres termes, cela signifie que la Bundesbank avait une attitude passive en se contentant d'ajuster ses taux directeurs à l'évolution d'une inflation déterminée indépendamment des évolutions monétaires (inflation importée, politique salariale...). Les autres pays en revanche ont clairement adopté à partir du milieu des années 1980 des politiques monétaires restrictives : les taux directeurs étaient fixés à des niveaux supérieurs aux taux de neutralité et aux taux de Taylor pour une inflation cible de 2 %. C'est sans doute à ce prix qu'à pu s'opérer la convergence des taux d'inflation au cours des années 1990.

6

LES HYPOTHÈSES DU MODÈLE DE RÉGRESSION LINÉAIRE MULTIPLE

1. La normalité des aléas
2. L'homoscédasticité et la non-autocorrélation des aléas
3. Les variables explicatives du modèle

OBJECTIFS DU CHAPITRE

Examiner les hypothèses du modèle de régression linéaire et les tests associés

- Hypothèse de normalité des aléas
- Test de Jarque-Bera
- Estimateur du maximum de vraisemblance
- Tests de Wald, du multiplicateur de Lagrange et du ratio de vraisemblance
- Estimateur des Moindres Carrés Généralisés
- Hypothèse d'homoscédasticité des aléas
- Tests de White, Breush-Pagan, Goldfeld-Quandt
- Hypothèse de non-autocorrélation des aléas
- Test de Durbin-Watson, statistique h de Durbin
- Méthodes de correction de l'autocorrélation d'ordre 1
- Correction de White de la matrice de variance-covariance
- Hypothèses sur les variables explicatives du modèle
- Estimateur des Variables Instrumentales
- Test d'Hausman
- Estimateur des éléments généralisés
- Test de validité des instruments

NIVEAU

LICENCE, troisième année et MASTER, première année

Les paramètres du modèle de régression linéaire multiple peuvent être estimés par la méthode des MCO sous les hypothèses 1 à 6 (cf. chapitre 3, Le modèle de régression linéaire multiple). En particulier, l'estimateur des MCO est distribué suivant une loi normale lorsque les aléas sont également distribués suivant une loi normale. Par ailleurs, cet estimateur est efficace lorsque les aléas sont homoscédastiques et non autocorrélés. Enfin, il est non biaisé et convergent lorsque l'hypothèse d'indépendance des variables explicatives et des aléas est vérifiée.

Dans ce chapitre ces trois hypothèses fondamentales du modèle de régression linéaire multiple sont successivement examinées.

1. LA NORMALITÉ DES ALÉAS

On reprend ici le modèle de régression linéaire multiple. L'hypothèse de normalité est essentielle pour mener à bien les tests statistiques usuels pour une taille d'échantillon finie. La normalité des aléas implique la normalité des estimateurs MCO¹ et la possibilité de réaliser les tests usuels.

Dans un premier temps on présente le test de normalité des aléas de Jarque-Béra et, dans un second temps, on montre que, sous l'hypothèse de normalité, l'estimateur des MCO des paramètres du modèle est identique à l'estimateur du maximum de vraisemblance. Enfin, trois tests d'hypothèses asymptotiquement équivalents, dérivés de l'estimateur du maximum de vraisemblance, sont présentés.

1.1 Test de l'hypothèse de normalité des aléas

On vérifie l'hypothèse de normalité des aléas en s'appuyant sur les caractéristiques de la distribution normale : le coefficient d'asymétrie θ_1 (Skewness) et le coefficient d'aplatissement θ_2 (Kurtosis). Pour une loi normale $\theta_1 = 0$ et $\theta_2 = 3$.

Soient $\hat{\theta}_1$ et $\hat{\theta}_2$ tels que

$$\hat{\theta}_1 = \frac{\frac{\sum_{i=1}^N e_i^3}{N}}{\left(\frac{\sum_{i=1}^N e_i^2}{N}\right)^{3/2}}, \quad \hat{\theta}_2 = \frac{\frac{\sum_{i=1}^N e_i^4}{N}}{\left(\frac{\sum_{i=1}^N e_i^2}{N}\right)^2}$$

¹ On montre que asymptotiquement, lorsque la taille de l'échantillon tend vers l'infini, l'estimateur des MCO est distribué selon une loi normale même si les aléas ne sont pas distribués selon une loi normale.

La statistique de Jarque-Béra, notée JB , s'écrit

$$JB = N \left(\frac{\widehat{\theta}_1^2}{6} + \frac{(\widehat{\theta}_2 - 3)^2}{24} \right) \quad (6.1)$$

La statistique JB suit une loi du χ^2 à 2 degrés de liberté. Si $JB > \chi_\alpha^2(2)$ alors l'hypothèse H_0 est refusée.

1.2 *Le maximum de vraisemblance*

Lorsque l'hypothèse de normalité des aléas est vérifiée, on montre que l'estimateur des MCO est identique à l'estimateur du maximum de vraisemblance.

1.2.1 *L'estimateur du maximum de vraisemblance*

L'estimateur du maximum de vraisemblance des paramètres d'une population correspond aux valeurs des paramètres qui génèrent le plus souvent l'échantillon observé. Si les aléas sont indépendamment distribués selon une loi normale avec une espérance nulle et une variance σ^2 alors les variables y_i ($i = 1, \dots, N$) sont également indépendamment distribuées selon une loi normale avec

$$E(y_i) = \beta_1 + \beta_2 x_{2i} + \dots + \beta_K x_{Ki} = \beta' x_i$$

et

$$var(y_i) = \sigma^2$$

La densité de probabilité s'écrit

$$f(y_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{y_i - \beta' x_i}{\sigma} \right)^2} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\varepsilon_i^2}{2\sigma^2}} \quad (6.2)$$

Les variables étant distribuées de manière indépendante, la fonction de vraisemblance est égale au produit des densités de probabilité $L = \prod_{i=1}^N f(y_i)$. Il est

équivalent de maximiser L ou $\text{Log}L = \sum_{i=1}^N \text{Log}f(y_i)$ et l'estimateur du maximum de vraisemblance est obtenu par la maximisation de la fonction $\text{Log}L$

$$\begin{aligned} \text{Max}_{\beta, \sigma^2} \text{Log}L &= -\frac{N}{2} [\text{Log}2\pi + \text{Log}\sigma^2] - \frac{1}{2\sigma^2} \varepsilon' \varepsilon \\ &= -\frac{N}{2} [\text{Log}2\pi + \text{Log}\sigma^2] - \frac{1}{2\sigma^2} (Y - X\beta)' (Y - X\beta) \end{aligned} \quad (6.3)$$

Les conditions du premier ordre permettent de déterminer les paramètres β et σ^2

$$\boxed{\begin{aligned} \widehat{\beta}_{MV} &= (X'X)^{-1}X'Y \\ \widehat{\sigma}_{MV}^2 &= \frac{e'e}{N} \end{aligned}} \tag{6.4}$$

et la matrice de variance-covariance asymptotique (*Asym.Var*) de l'estimateur du maximum de vraisemblance est donnée par l'inverse de la matrice d'information de Fisher

$$- \begin{bmatrix} E \left(\frac{\partial^2 \text{Log}L}{\partial \beta \partial \beta'} \right) & E \left(\frac{\partial^2 \text{Log}L}{\partial \beta \partial \sigma^2} \right) \\ E \left(\frac{\partial^2 \text{Log}L}{\partial \sigma^2 \partial \beta'} \right) & E \left(\frac{\partial^2 \text{Log}L}{\partial (\sigma^2)^2} \right) \end{bmatrix}^{-1} = \begin{bmatrix} \sigma^2 (X'X)^{-1} & 0 \\ 0 & \frac{2\sigma^4}{N} \end{bmatrix} \tag{6.5}$$

L'estimateur du maximum de vraisemblance $\widehat{\beta}_{MV}$ est égal à l'estimateur des MCO $\widehat{\beta}_{MCO}$.

Il est connu pour ses propriétés asymptotiques, c'est-à-dire quand la taille de l'échantillon tend vers $+\infty$. C'est un estimateur distribué asymptotiquement selon une loi normale et efficace.

Donc sous l'hypothèse de normalité, l'estimateur des MCO de β est également un estimateur asymptotiquement efficace. De même asymptotiquement, l'estimateur des MCO de la variance des aléas est identique à celui du maximum de vraisemblance (quand $N \rightarrow \infty$, $\widehat{\sigma}_{MV}^2 = \widehat{\sigma}_{MCO}^2$).

1.2.2 Tests d'hypothèses basés sur l'estimateur du maximum de vraisemblance

On présente usuellement trois tests d'hypothèses asymptotiquement équivalents basés sur l'estimateur du maximum de vraisemblance : le test de Wald, le test du ratio de vraisemblance et le test du multiplicateur de Lagrange. Supposons que le modèle a été estimé par la méthode du maximum de vraisemblance et que l'on souhaite tester J contraintes linéaires

$$H_0 : R\beta = q \quad \text{contre} \quad H_1 : R\beta \neq q$$

où R est une matrice d'ordre (J, K) et q un vecteur composé de J lignes.

1.2.3 Test de Wald

Le test de Wald a pour but de déterminer si l'écart entre $R\widehat{\beta}$ et q est significativement égal à 0.

L'estimateur du maximum de vraisemblance est distribué asymptotiquement selon une loi normale

$$\widehat{\beta}_{MV} \rightsquigarrow N \left(\beta, \sigma^2 (X'X)^{-1} \right)$$

et

$$\sqrt{N} \left(\widehat{\beta}_{MV} - \beta \right) \sim N \left(0, \sigma^2 \left(\frac{X'X}{N} \right)^{-1} \right)$$

Par conséquent

$$\sqrt{N} R \left(\widehat{\beta}_{MV} - \beta \right) = \sqrt{N} \left(R \widehat{\beta}_{MV} - q \right) \sim N \left(0, \sigma^2 R \left(\frac{X'X}{N} \right)^{-1} R' \right)$$

d'où la statistique de Wald

$$N \left(R \widehat{\beta}_{MV} - q \right)' \left(\widehat{\sigma}^2 R \left(\frac{X'X}{N} \right)^{-1} R' \right)^{-1} \left(R \widehat{\beta}_{MV} - q \right) \sim \chi_{\alpha}^2(J)$$

car $P \lim \widehat{\sigma}^2 \left(\frac{X'X}{N} \right)^{-1} = \sigma^2 \left(\frac{X'X}{N} \right)^{-1}$.

La distribution limite de la statistique de Wald est donc un $\chi_{\alpha}^2(J)$. Cette statistique est égale au nombre de contraintes multipliées par la statistique de Fisher F_{cal} du test de plusieurs contraintes (cf. chapitre 3 Le modèle de régression linéaire multiple)

$$\boxed{\text{Wald} = \left(R \widehat{\beta}_{MV} - q \right)' \left(\widehat{\sigma}^2 R \left(\frac{X'X}{N} \right)^{-1} R' \right)^{-1} \left(R \widehat{\beta}_{MV} - q \right) = J \cdot F_{cal}} \quad (6.6)$$

Sous l'hypothèse de normalité des aléas, la distribution exacte de la statistique F_{cal} est $F(J, N - K)$. Ici on montre que la distribution limite de la statistique $J \cdot F_{cal}$ est un $\chi^2(J)$. Par conséquent le test de plusieurs contraintes de Fisher est valide, indépendamment de la taille de l'échantillon, sous l'hypothèse de normalité des aléas.

Enfin si la statistique de Wald est inférieure à la valeur du $\chi_{\alpha}^2(J)$ alors les contraintes ne sont pas rejetées.

1.2.4 Test du ratio de vraisemblance

Le test du ratio de vraisemblance se base sur la comparaison des valeurs prises par le logarithme de la vraisemblance selon que le modèle est contraint ou non contraint. Le logarithme de la vraisemblance du modèle de régression linéaire multiple est donné par l'équation (6.3).

Soient $\widehat{\beta}_{MV}$ l'estimateur du modèle non contraint et $e'e$ la somme des carrés des résidus correspondante²

$$\text{Log}L(\widehat{\beta}_{MV}) = -\frac{N}{2} [\text{Log}2\pi + \text{Log}\widehat{\sigma}^2] - \frac{1}{2\widehat{\sigma}^2} e'e \quad (6.7)$$

$$= -\frac{N}{2} \left[\text{Log}2\pi + \text{Log}\frac{e'e}{N} \right] - \frac{N}{2} \quad (6.8)$$

De même, soient $\widehat{\beta}_{MV}^*$ l'estimateur du modèle contraint et e^*e^* la somme des carrés des résidus correspondante³

$$\begin{aligned} \text{Log}L(\widehat{\beta}_{MV}^*) &= -\frac{N}{2} [\text{Log}2\pi + \text{Log}\widehat{\sigma}^{*2}] - \frac{1}{2\widehat{\sigma}^{*2}} e^*e^* \\ &= -\frac{N}{2} \left[\text{Log}2\pi + \text{Log}\frac{e^*e^*}{N} \right] - \frac{N}{2} \end{aligned} \quad (6.9)$$

Sous l'hypothèse H_0 la statistique du ratio de vraisemblance (LR) a une distribution limite identique à la statistique de Wald, elle s'écrit

$$LR = -2 \left(\text{Log}L(\widehat{\beta}_{MV}^*) - \text{Log}L(\widehat{\beta}_{MV}) \right) \rightsquigarrow \chi^2_\alpha(J) \quad (6.10)$$

la statistique LR peut également s'écrire sous la forme

$$\boxed{LR = N (\text{Log}(e^*e^*) - \text{Log}(e'e))} \quad (6.11)$$

Si $LR < \chi^2_\alpha(J)$ alors l'hypothèse H_0 n'est pas rejetée.

1.2.5 Test du multiplicateur de Lagrange

Le test du multiplicateur de Lagrange, appelé aussi le test du Score, est basé sur l'utilisation de la condition du premier ordre du problème de maximisation du logarithme de la fonction de vraisemblance. $\widehat{\beta}_{MV}$, l'estimateur non contraint du modèle, satisfait la condition $\frac{\partial \text{Log}L}{\partial \beta} = 0$.

² On note que $\widehat{\sigma}^2 = \frac{e'e}{N}$ et $(Y - X\widehat{\beta})' (Y - X\widehat{\beta}) = e'e$ d'où $\frac{1}{2\widehat{\sigma}^2} e'e = \frac{N}{2}$

³ On note que :

$\widehat{\sigma}^{*2} = \frac{e^*e^*}{N}$ et $(Y - X\widehat{\beta}^*)' (Y - X\widehat{\beta}^*) = e^*e^*$ d'où $\frac{1}{2\widehat{\sigma}^{*2}} e^*e^* = \frac{N}{2}$

Le test propose d'évaluer cette condition en $\hat{\beta}_{MV}^*$ et de calculer la statistique du multiplicateur de Lagrange donnée par

$$LM = \left[\frac{\partial \text{Log} L}{\partial \beta} \right]'_{\beta = \hat{\beta}_{MV}^*} \left[-E \left(\frac{\partial^2 \text{Log} L}{\partial \beta \partial \beta'} \right) \right]^{-1}_{\beta = \hat{\beta}_{MV}^*} \left[\frac{\partial \text{Log} L}{\partial \beta} \right]_{\beta = \hat{\beta}_{MV}^*}$$

Or

$$\left[\frac{\partial \text{Log} L}{\partial \beta} \right]_{\beta = \hat{\beta}_{MV}^*} = \left[\frac{1}{\sigma^2} X' [Y - X\beta] \right]_{\beta = \hat{\beta}_{MV}^*} = \frac{N}{e'^* e^*} X' e^*$$

et

$$\left[-E \left(\frac{\partial^2 \text{Log} L}{\partial \beta \partial \beta'} \right) \right]^{-1}_{\beta = \hat{\beta}_{MV}^*} = \left[\sigma^2 (X' X)^{-1} \right]_{\beta = \hat{\beta}_{MV}^*} = \frac{e'^* e^*}{N} (X' X)^{-1}$$

Par conséquent la statistique du multiplicateur de Lagrange s'écrit

$$LM = \left[\frac{N}{e'^* e^*} X' e^* \right]' \left[\frac{e'^* e^*}{N} (X' X)^{-1} \right] \left[\frac{N}{e'^* e^*} X' e^* \right]$$

$$\boxed{LM = \frac{N}{e'^* e^*} (X' e^*)' (X' X)^{-1} (X' e^*) = N R_{e^*}^2} \quad (6.12)$$

où $R_{e^*}^2$ ⁴ correspond au coefficient de détermination d'une estimation par les MCO d'un modèle dans lequel la variable endogène est e^* , et les variables exogènes X . Cette statistique suit la même distribution que la statistique de Wald et du ratio de vraisemblance. Si $LM < \chi_{\alpha}^2(J)$ alors on ne refuse pas les contraintes.

⁴ Dans un modèle sur données centrées $Y = X\beta + \varepsilon$.
Le coefficient de détermination s'écrit

$$\begin{aligned} R^2 &= \frac{\hat{Y}' \hat{Y}}{Y' Y} = \frac{(X \hat{\beta})' (X \hat{\beta})}{Y' Y} = \frac{\hat{\beta}' (X' X) \hat{\beta}}{Y' Y} = \frac{\hat{\beta}' X' Y}{Y' Y} \\ &= \frac{[(X' X)^{-1} X' Y]' X' Y}{Y' Y} \end{aligned}$$

On obtient ainsi pour des données centrées le coefficient de détermination

$$R^2 = \frac{(X' Y)' (X' X)^{-1} (X' Y)}{Y' Y}$$

Si la variable endogène centrée est notée e^* , on retrouve la statistique LM au facteur N près.

2. L'HOMOSCÉDASTICITÉ ET LA NON-AUTOCORRÉLATION DES ALÉAS

Dans cette section est discutée l'hypothèse d'homoscédasticité et de non-autocorrélation des aléas, c'est-à-dire l'hypothèse 2 selon laquelle : $E(\varepsilon\varepsilon') = \sigma^2 I$. Sous cette hypothèse l'estimateur des MCO est efficace, autrement dit il a la variance la plus faible parmi les estimateurs linéaires et non biaisés de β . Lorsque cette hypothèse n'est pas vérifiée il existe un autre estimateur linéaire et non biaisé à variance plus faible, il s'agit de l'estimateur des Moindres Carrés Généralisés (MCG). Dans un premier temps cet estimateur est présenté. Dans un second temps on considère deux cas d'application de cet estimateur : l'hétéroscédasticité des aléas ($\text{var}(\varepsilon_i) = \sigma_i^2$) et l'autocorrélation d'ordre 1 ($\text{cov}(\varepsilon_i, \varepsilon_{i-1}) \neq 0$).

2.1 L'estimateur des MCG

Le modèle linéaire généralisé s'écrit

$$Y = X\beta + \varepsilon \quad \text{avec} \quad E(\varepsilon\varepsilon') = \sigma^2 \Omega \quad (6.13)$$

où $\Omega \neq I$. Ω est une matrice symétrique et définie positive d'ordre (N, N) .

2.1.1 L'estimateur des MCO et le modèle linéaire généralisé

L'estimateur des MCO, $\hat{\beta}_{MCO}$, reste un estimateur linéaire non biaisé et convergent, sous les hypothèses 1, 3, 4 et 5. En revanche, $\hat{\beta}_{MCO}$ n'est pas efficace lorsque les aléas ne respectent pas l'hypothèse 2.

Sous l'hypothèse $E(\varepsilon\varepsilon') = \sigma^2 \Omega$, la matrice de variance-covariance de l'estimateur des MCO s'écrit

$$\text{Var}(\hat{\beta}_{MCO}) = E \left[\left(\hat{\beta}_{MCO} - \beta \right) \left(\hat{\beta}_{MCO} - \beta \right)' \right]$$

avec $\hat{\beta}_{MCO} - \beta = (X'X)^{-1}X'\varepsilon$ d'où

$$\begin{aligned} \text{Var}(\hat{\beta}_{MCO}) &= E \left[\left((X'X)^{-1}X'\varepsilon \right) \left((X'X)^{-1}X'\varepsilon \right)' \right] \\ &= \sigma^2 (X'X)^{-1}X'\Omega X(X'X)^{-1} \end{aligned} \quad (6.14)$$

2.1.2 L'estimateur des MCG

L'estimateur des MCG est obtenu en transformant le modèle (6.13) de manière à obtenir un modèle de régression linéaire multiple dont les aléas vérifient les hypothèses 1, 2 et 3.

Ω étant une matrice symétrique définie positive, sa matrice inverse Ω^{-1} est également symétrique et définie positive et il existe $\Omega^{-1/2}$ symétrique définie positive telle que

$$\Omega^{-1} = \Omega^{-1/2} \Omega^{-1/2}$$

En prémultipliant le modèle initial par $\Omega^{-1/2}$ (transformation d'Aitken), on obtient un modèle dont les aléas vérifient les hypothèses 1,2 et 3.

Soient $Y_* = \Omega^{-1/2}Y$, $X_* = \Omega^{-1/2}X$ et $\varepsilon_* \stackrel{\text{d}}{=} \Omega^{-1/2}\varepsilon$, le modèle transformé s'écrit

$$Y_* = X_*\beta + \varepsilon_* \quad (6.15)$$

avec

$$\begin{aligned} E(\varepsilon_*) &= \Omega^{-1/2}E(\varepsilon) = 0 \\ \text{Var}(\varepsilon_*) &= E(\Omega^{-1/2}\varepsilon\varepsilon'\Omega^{-1/2}) \\ &= \Omega^{-1/2} \underbrace{E(\varepsilon\varepsilon')}_{\sigma^2\Omega} \Omega^{-1/2} \\ &= \sigma^2 I \end{aligned}$$

L'estimateur des MCG est égal à l'estimateur des MCO des paramètres du modèle transformé. Il minimise $\varepsilon_*'\varepsilon_* = \varepsilon'\Omega^{-1}\varepsilon$

$$\boxed{\hat{\beta}_{MCG} = (X_*'X_*)^{-1}X_*'Y_* = \left((X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y \right)} \quad (6.16)$$

Il est BLUE

$$E\left(\hat{\beta}_{MCG}\right) = \beta \quad (6.17)$$

et sa matrice de variance-covariance est donnée par

$$\boxed{\text{Var}\left(\hat{\beta}_{MCG}\right) = \sigma^2(X_*'X_*)^{-1} = \sigma^2(X'\Omega^{-1}X)^{-1}} \quad (6.18)$$

Si de plus les aléas vérifient l'hypothèse 6, l'estimateur des MCG est distribué selon une loi normale. Enfin si $\lim_{N \rightarrow +\infty} S_{X_*} = \lim_{N \rightarrow +\infty} \frac{X_*'X_*}{N} = V_{X_*}$, avec V_{X_*} une matrice finie et définie positive, alors l'estimateur des MCG est convergent.

On note que l'application de la méthode des MCG exige de connaître à une constante près la matrice Ω . Lorsque celle-ci est inconnue, on doit chercher un estimateur $\hat{\Omega}$ de cette matrice de variance-covariance afin d'appliquer la transformation d'Aitken. L'application de l'estimateur des MCO sur ce modèle transformé fournit alors l'estimateur des Moindres Carrés Quasi-Généralisés (MCQG)

$$\boxed{\begin{aligned} \hat{\beta}_{MCQG} &= \left((X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}Y \right) \\ \text{Var}\left(\hat{\beta}_{MCQG}\right) &= \sigma^2(X'\hat{\Omega}^{-1}X)^{-1} \end{aligned}} \quad (6.19)$$

La matrice $\widehat{\Omega}$ est généralement estimée à partir des résidus des MCO (cf. cas de l'hétéroscédasticité et de l'autocorrélation), par conséquent le terme $(X' \widehat{\Omega}^{-1} X)^{-1} X' \widehat{\Omega}^{-1}$ est aléatoire et non indépendant des aléas du modèle. L'estimateur des MCQG est généralement biaisé, mais $\widehat{\beta}_{MCQG}$ est asymptotiquement équivalent à $\widehat{\beta}_{MCG}$ si les deux conditions suivantes sont vérifiées

$$P \lim \frac{(X' \widehat{\Omega}^{-1} X)^{-1}}{N} = \lim \frac{(X' \Omega^{-1} X)^{-1}}{N} \quad (6.20)$$

et

$$P \lim \frac{X' \widehat{\Omega}^{-1} \varepsilon}{N} = P \lim \frac{X' \Omega^{-1} \varepsilon}{N} \quad (6.21)$$

Asymptotiquement l'estimateur des MCQG est donc convergent sous ces deux conditions et il est efficace si $P \lim \widehat{\Omega} = \Omega$. Enfin, asymptotiquement l'estimateur centré dilaté $\sqrt{N} (\widehat{\beta}_{MCQG} - \beta)$ converge en loi vers la même loi asymptotique que $\sqrt{N} (\widehat{\beta}_{MCG} - \beta)$

$$\sqrt{N} (\widehat{\beta}_{MCQG} - \beta) \sim N(0, \sigma^2 V_{X_*}^{-1}) \quad (6.22)$$

2.2 Cas de l'hétéroscédasticité des aléas

Les aléas sont hétéroscédastiques lorsque $var(\varepsilon_i) = \sigma_i^2 \neq \sigma^2$, $i = 1, \dots, N$. Généralement ce problème apparaît sur des données individuelles (données en coupe transversale) mais il peut cependant se présenter sur des données temporelles (données en coupe longitudinale). Il existe plusieurs tests⁵ statistiques de détection de l'hétéroscédasticité. Trois tests sont présentés ici : le test de White, le test de Breusch-Pagan et le test de Goldfeld-Quandt.

2.2.1 Application des tests statistiques d'homoscédasticité des aléas

Les tests statistiques d'homoscédasticité portent sur l'hypothèse

$$H_0 : \sigma_1^2 = \dots = \sigma_m^2 = \sigma^2 \quad (m \leq N)$$

c'est-à-dire sur l'hypothèse d'une variance des aléas identique pour chaque individu. La première étape de tous les tests consiste à estimer le modèle par la méthode des

⁵ Dans une première étape, il est intéressant de représenter graphiquement le carré des résidus MCO (e_i^2) en fonction des valeurs estimées de la variable endogène (\widehat{y}_i) ou d'une variable explicative (x_{ki}), afin d'identifier une relation possible, qui peut-être à l'origine de la non constance de la variance des aléas.

MCO afin de calculer les résidus du modèle (e_i). Le carré de ces résidus est, dans une seconde étape, utilisé comme estimateur de σ_i^2 .

1. **Test de White** : Le test de White est basé sur la comparaison des variances estimées des aléas lorsque le modèle est estimé par les MCO sous l'hypothèse d'homoscédasticité et lorsque le modèle est estimé par les MCO sous l'hypothèse d'hétéroscédasticité. Lorsque l'hypothèse H_0 est vérifiée, l'écart entre les deux variances estimées devrait être significativement nul asymptotiquement.

En pratique, ce test consiste à estimer par les MCO l'équation

$$e_i^2 = \delta_0 + \delta_1 Z_{i1} + \dots + \delta_p Z_{ip} + u_i \quad (6.23)$$

où les variables $Z_{ik}, k = 1, \dots, p$ sont les variables explicatives du modèle, leurs carrés et leurs produits et u représente l'aléa.

Par exemple, si on teste l'homoscédasticité dans un modèle avec une constante et deux variables explicatives x_{i2} et x_{i3} , $p = 5$

$$Z_{i1} = x_{i2}, Z_{i2} = x_{i3}, Z_{i3} = x_{i2}^2, Z_{i4} = x_{i3}^2, Z_{i5} = x_{i2}x_{i3}$$

De l'estimation par les MCO de l'équation (6.23) on déduit le coefficient de détermination noté $R_{e^2}^2$. Asymptotiquement, sous l'hypothèse H_0 la statistique

$$W_h = NR_{e^2}^2 \quad (6.24)$$

est distribuée selon un $\chi^2(p)$. Si $NR_{e^2}^2 < \chi_\alpha^2(p)$ alors l'hypothèse d'homoscédasticité est retenue.

2. **Test de Breush-Pagan** : Le test de Breush-Pagan est basé sur le principe que sous l'hypothèse H_0 l'estimateur des MCO ne doit pas différer significativement de l'estimateur du maximum de vraisemblance, qui admet la possibilité d'une hétéroscédasticité des aléas. Par rapport au test de White, l'hypothèse H_1 du test de Breush et Pagan est définie de manière précise⁶

$$H_1 : \sigma_i^2 = \sigma^2 f(\delta_0 + \delta_1 Z_{i1} + \dots + \delta_p Z_{ip})$$

où f est une fonction continue dérivable et $Z_{ik}, k = 1, \dots, p$ est un vecteur de variables indépendantes qui correspond généralement aux variables explicatives du modèle, ou à une fonction connue de ces variables. Si les aléas sont homoscédastiques alors $\delta_1 = \dots = \delta_p = 0$.

Le test de Breush-Pagan consiste à appliquer le test du multiplicateur de Lagrange. On montre que asymptotiquement, sous l'hypothèse H_0 , la statistique LM

$$LM = \frac{1}{2} g' Z (Z'Z)^{-1} Z'g \quad (6.25)$$

⁶ Dans le test de White l'hypothèse H_1 indique la présence d'une hétéroscédasticité des aléas, mais la forme que prend cette hétéroscédasticité n'est pas spécifiée. Le test de White est un test général de détection de l'hétéroscédasticité.

converge en loi vers un $\chi^2(p)$ où g est le vecteur des observations $g_i = \frac{e_i^2}{\frac{e' e}{N}} - 1$

et Z la matrice d'ordre $(N, p+1)$ suivante

$$Z_{(N, p+1)} = \begin{pmatrix} 1 & Z_{11} & \cdot & \cdot & Z_{p1} \\ 1 & Z_{12} & \cdot & \cdot & Z_{p2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & Z_{1N} & \cdot & \cdot & Z_{pN} \end{pmatrix}$$

Si $LM < \chi^2_\alpha(p)$ alors l'hypothèse d'homoscédasticité n'est pas refusée.

Cette statistique peut également être calculée à partir de la somme des carrés expliquée par le modèle

$$g_i = \delta_0 + \delta_1 Z_{i1} + \dots + \delta_p Z_{ip} + u_i = \delta' Z_i + u_i \quad (6.26)$$

$$LM = \frac{1}{2} \widehat{g}' \widehat{g} \quad (6.27)$$

- 3. Test de Goldfeld-Quantt :** Le test de Goldfeld-Quantt est recommandé lorsqu'on suppose que la variance des aléas est fonction d'une variable explicative x_k du modèle. Dans ce cas les observations sont ordonnées suivant les valeurs croissantes de x_k . Si σ^2 est une fonction croissante de x_k alors les observations seront également ordonnées suivant les valeurs croissantes de la variance des aléas.

Le test de Goldfeld-Quantt consiste à scinder l'échantillon en deux parties égales en omettant C observations au milieu de l'échantillon. Si les aléas sont homoscédastiques alors la variance sur les deux échantillons sera identique. Sinon, si la variance est une fonction croissante de x_k , alors la variance sur le premier échantillon (σ_1^2) sera plus faible que sur le second (σ_2^2). Le test d'homoscédasticité correspond à un test d'égalité de variances. L'estimation par les MCO du modèle sur les deux échantillons, composés chacun de $(N - C)/2$ observations, permet de calculer

$$\widehat{\sigma}_1^2 = \frac{e_1' e_1}{\frac{N - C}{2} - K} \quad \text{et} \quad \widehat{\sigma}_2^2 = \frac{e_2' e_2}{\frac{N - C}{2} - K} \quad (6.28)$$

Sous l'hypothèse H_0

$$\frac{\widehat{\sigma}_j^2}{\sigma^2} \sim > \frac{\chi^2 \left(\frac{N - C}{2} - K \right)}{\frac{N - C}{2} - K} \quad j = 1, 2 \quad (6.29)$$

et la statistique du test de Goldfeld-Quandt est donnée par⁷

$$F_{cal} = \frac{\hat{\sigma}_2^2}{\hat{\sigma}_1^2} = \frac{e_2' e_2}{e_1' e_1} \sim F \left(\frac{N-C}{2} - K, \frac{N-C}{2} - K \right) \quad (6.30)$$

Si la statistique F_{cal} est inférieure au $F_\alpha \left(\frac{N-C}{2} - K, \frac{N-C}{2} - K \right)$, alors l'hypothèse d'hétéroscédasticité est rejetée.

Pour effectuer ce test il reste à définir le nombre C . Ce nombre d'observations omises de l'échantillon n'est pas donné clairement. Enlever ces observations conduit à accroître la puissance du test, car les écarts entre les deux variances sont plus importants s'il y a hétéroscédasticité. Dans la pratique il est conseillé d'enlever 1/6^{ème} des observations.

2.2.2 Estimation des paramètres en cas d'hétéroscédasticité des aléas

Le problème ici consiste à déterminer un estimateur BLUE des paramètres du modèle $Y = X\beta + \varepsilon$ dans lequel $var(\varepsilon_i) = \sigma_i^2$. La matrice de variance-covariance des aléas de ce modèle s'écrit sous la forme

$$Var(\varepsilon) = \begin{bmatrix} \sigma_1^2 & 0 & \cdot & \cdot & 0 \\ 0 & \sigma_2^2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \sigma_N^2 \end{bmatrix} = \Omega \neq \sigma^2 I \quad (6.31)$$

Ce modèle constitue un cas particulier du modèle linéaire généralisé (6.13) dans lequel σ^2 est égal à 1 et Ω est une matrice diagonale (les covariances entre les aléas sont nulles). L'estimateur des MCG est un estimateur BLUE des paramètres du modèle. La matrice de transformation $\Omega^{-1/2}$ s'écrit

$$\Omega^{-1/2} = \begin{bmatrix} 1/\sigma_1 & 0 & \cdot & \cdot & 0 \\ 0 & 1/\sigma_2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & 1/\sigma_N \end{bmatrix}$$

et pour chaque observation i , le modèle transformé s'écrit

$$\frac{y_i}{\sigma_i} = \beta_1 \frac{1}{\sigma_i} + \beta_2 \frac{x_{2i}}{\sigma_i} + \dots + \beta_K \frac{x_{Ki}}{\sigma_i} + \frac{\varepsilon_i}{\sigma_i} \quad (6.32)$$

L'estimateur des MCG est obtenu en appliquant les MCO sur (6.32), modèle dans lequel les observations x_{ki} et les aléas ε_i sont pondérés par l'écart type des aléas.

⁷ Si σ^2 est une fonction décroissante de la variable x_k et que les observations ont été rangées suivant les valeurs croissantes de x_k , on a $\sigma_1^2 > \sigma_2^2$ et on doit calculer la statistique $F_{cal} = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$.

L'estimateur des MCG dans ce cas est également appelé l'estimateur des Moindres Carrés Pondérés.

Plus particulièrement, si les tests d'hétéroscédasticité indiquent par exemple que $\sigma_i^2 = \sigma^2 x_{2i}$ on doit estimer par la méthode des MCO le modèle

$$\frac{y_i}{\sqrt{x_{2i}}} = \beta_1 \frac{1}{\sqrt{x_{2i}}} + \beta_2 \frac{x_{2i}}{\sqrt{x_{2i}}} + \dots + \beta_K \frac{x_{Ki}}{\sqrt{x_{2i}}} + \frac{\varepsilon_i}{\sqrt{x_{2i}}} \quad (6.33)$$

dans ce modèle

$$\text{var} \left(\frac{\varepsilon_i}{\sqrt{x_{2i}}} \right) = \frac{1}{x_{2i}} \text{var}(\varepsilon_i) = \sigma^2$$

Enfin, lorsque σ_i^2 est inconnu, les MCQG s'appliquent. Il faut alors, dans un premier temps, déterminer un estimateur convergent de la variance des aléas. Par exemple si la variance des aléas est une forme linéaire des variables explicatives du modèle telle que

$$\sigma_i^2 = \delta_1 + \delta_2 x_{2i} + \dots + \delta_K x_{Ki} = x_i' \delta$$

on estime avec la méthode des MCO l'équation

$$e_i^2 = \delta_1 + \delta_2 x_{2i} + \dots + \delta_K x_{Ki} + u_i = x_i' \delta + u_i \quad (6.34)$$

où u_i est un aléa et e_i^2 le carré des résidus du modèle (6.13) estimé avec la méthode des MCO. Si le modèle (6.34) satisfait les hypothèses 1, 2, 3, 4 et 5 alors l'estimateur des MCO $\hat{\delta}$ de δ est convergent et $\hat{e}_i^2 = x_i' \hat{\delta}$ est un estimateur convergent de σ_i^2 .

Finalement, on calcule l'estimateur des MCQG en estimant par les MCO l'équation

$$\frac{y_i}{\hat{e}_i} = \beta_1 \frac{1}{\hat{e}_i} + \beta_2 \frac{x_{2i}}{\hat{e}_i} + \dots + \beta_K \frac{x_{Ki}}{\hat{e}_i} + \frac{\varepsilon_i}{\hat{e}_i} \quad (6.35)$$

Cet estimateur est asymptotiquement équivalent à l'estimateur des MCG.

2.3 Cas de l'autocorrélation d'ordre 1 des aléas

Il y a autocorrélation des aléas lorsque la covariance des aléas est non nulle. Dans ce cas, la matrice de variance-covariance des aléas n'est plus diagonale et le modèle peut s'écrire sous la forme d'un modèle de régression linéaire généralisé (6.13). Le problème de l'autocorrélation se rencontre sur des données temporelles.

Ici on traite le cas de l'autocorrélation d'ordre 1 des aléas

$$\varepsilon_i = \rho \varepsilon_{i-1} + u_i$$

où u_i est un aléa dont l'espérance est nulle ($E(u_i) = 0$), la variance constante ($\sigma_{u_i}^2 = \sigma_u^2 \forall i = 1, \dots, N$) et la covariance nulle ($cov(u_i, u_j) = 0, i \neq j$). La matrice de variance-covariance des aléas structurels s'écrit alors : $E(\varepsilon\varepsilon') = \sigma_u^2\Omega$.

Dans un premier temps un test de détection de l'autocorrélation d'ordre 1 des aléas est présenté⁸. Dans un second temps la matrice de variance-covariance des aléas ε est déterminée. Enfin, l'estimation des paramètres d'un modèle linéaire avec autocorrélation d'ordre 1 des aléas est examinée.

2.3.1 Test de l'absence d'autocorrélation d'ordre 1 des aléas : le test de Durbin-Watson

Le test de Durbin-Watson consiste à vérifier si dans la relation $\varepsilon_i = \rho\varepsilon_{i-1} + u_i$ le coefficient d'autocorrélation ρ est significativement égal à 0. On teste l'hypothèse $H_0 : \rho = 0$.

La première étape consiste à estimer le modèle $Y = X\beta + \varepsilon$ avec la méthode des MCO et à évaluer les résidus e_i . Ensuite à l'aide de ces résidus, on calcule la statistique de Durbin-Watson

$$d = \frac{\sum_{i=2}^N (e_i - e_{i-1})^2}{\sum_{i=1}^N e_i^2} \quad (6.36)$$

La distribution de cette statistique dépend de manière complexe des observations sur les variables explicatives du modèle. Elle est calculée à partir des résidus déterminés pour X donnée. De ce fait, il n'existe pas une valeur critique unique comme dans les tests de Fisher, de Student ou du Chi-deux. Cependant, sous H_0 , on montre que la distribution exacte de la statistique d se situe entre les distributions de deux autres statistiques notées d_l , pour la borne la plus faible, et d_u pour la borne la plus élevée. Pour un risque de première espèce, on peut déterminer les deux valeurs seuils $d_{l\alpha}$ et $d_{u\alpha}$. Les lois de ces deux statistiques ne dépendent que du nombre de variables explicatives et du nombre d'observations, leurs valeurs sont données dans la table de Durbin-Watson.

⁸ Comme pour l'hétéroscédasticité, une manière simple de tester dans un premier temps l'autocorrélation des aléas consiste à représenter les résidus au cours du temps ou e_i en fonction de e_{i-1} , afin de détecter une relation éventuelle entre les résidus.

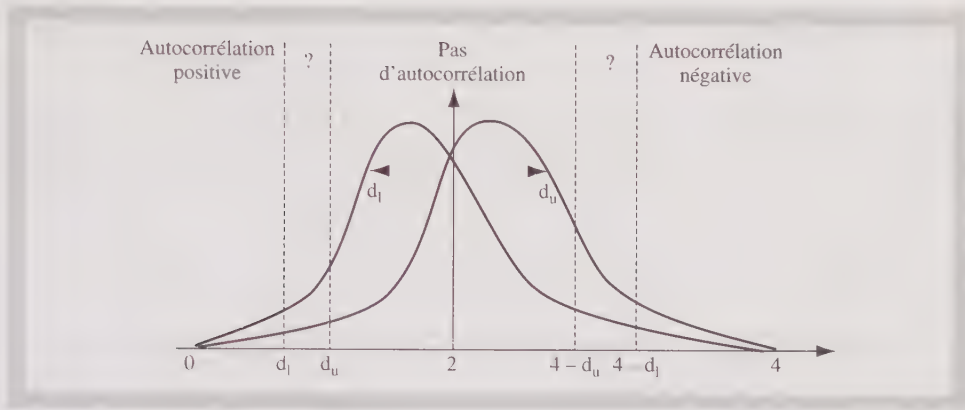


Figure 6.1 Test de Durbin-Watson

La règle de décision selon la valeur de la statistique d est la suivante

$0 < d < d_{l\alpha}$	rejet de H_0	autocorrélation positive
$d_{l\alpha} \leq d \leq d_{u\alpha}$	pas de décision	—
$d_{u\alpha} < d < 4 - d_{u\alpha}$	non rejet de H_0	pas d'autocorrélation
$4 - d_{u\alpha} \leq d \leq 4 - d_{l\alpha}$	pas de décision	—
$4 - d_{l\alpha} < d < 4$	rejet de H_0	autocorrélation négative

La statistique de Durbin-Watson est comprise entre 0 et 4. En effet, la relation entre d et $\hat{\rho}^9$, estimateur convergent du coefficient d'autocorrélation des aléas, est donnée par

$$d \simeq 2(1 - \hat{\rho})$$

or $-1 \leq \hat{\rho} \leq 1$ d'où si $\hat{\rho} = -1$ alors $d \simeq 4$, si $\hat{\rho} = 1$ alors $d \simeq 0$, enfin si $\hat{\rho} = 0$ alors $d \simeq 2$.

Pour appliquer ce test, les conditions suivantes doivent être respectées : le modèle de régression doit inclure une constante et les variables explicatives doivent être certaines. En particulier, ce test ne peut pas s'appliquer lorsque le modèle est autorégressif, c'est-à-dire lorsque la variable endogène décalée apparaît comme variable explicative. Par exemple dans le modèle dynamique

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 y_{i-1} + \varepsilon_i \quad i = 1, \dots, N$$

⁹ $\hat{\rho} = \left(\frac{\sum_{i=2}^N e_i e_{i-1}}{\sum_{i=2}^N e_{i-1}^2} \right)$ ce coefficient n'est pas utilisé pour le test car sa distribution n'est pas déterminée contrairement à celle de d .

il faut utiliser la statistique h de Durbin pour tester l'autocorrélation d'ordre 1 des aléas

$$h = \left(1 - \frac{d}{2}\right) \sqrt{\frac{N}{1 - N \left(\widehat{\text{var}}\left(\widehat{\beta}_3\right)\right)}} \sim N(0,1) \quad (6.37)$$

l'hypothèse nulle $\rho = 0$ est acceptée si $|h| \leq 1.645$ pour un risque de première espèce de 5 %. Le test est unilatéral, l'hypothèse alternative est $\rho > 0$ ou $\rho < 0$.

Si $N \left(\widehat{\text{var}}\left(\widehat{\beta}_3\right)\right)$ prend une valeur supérieure à 1, on ne peut pas calculer la statistique h de Durbin. On peut procéder alors de la manière suivante

1. estimer le modèle dynamique avec la méthode des MCO
2. estimer le modèle suivant

$$e_i = \alpha_1 + \alpha_2 x_{2i} + \alpha_3 y_{i-1} + \alpha_4 e_{i-1} + \varepsilon_i \quad i = 1, \dots, N$$

où e_i est le résidu MCO du modèle dynamique

3. si le coefficient de la variable e_{i-1} est significativement différent de 0 on peut rejeter l'hypothèse nulle de non-autocorrélation.

Enfin, il existe une version du test de Durbin-Watson, lorsque les données sont trimestrielles, donnée par Wallis

$$\varepsilon_i = \rho \varepsilon_{i-4} + u_i$$

Il s'agit de tester, sous H_0 , si les aléas sont non autocorrélés d'ordre 4 ($\rho = 0$). On doit calculer une statistique analogue à celle de Durbin-Watson¹⁰

$$d_4 = \frac{\sum_{i=5}^N (e_i - e_{i-4})^2}{\sum_{i=1}^N e_i^2} \quad (6.38)$$

¹⁰ La table statistique peut être trouvée dans W.H. Greene, 1993, *Econometric Analysis*, 2^e édition, Macmillan.

2.3.2 Détermination de la matrice de variance-covariance des aléas

Lorsqu'un problème d'autocorrélation d'ordre 1 est détecté l'estimateur des MCO du modèle n'est pas efficace¹¹. Il faut appliquer les MCG et par conséquent définir la matrice Ω . Dans le cas d'une autocorrélation d'ordre 1, lorsque $|\rho| < 1$ et donc que $\lim_{s \rightarrow \infty} \rho^s = 0$, les aléas peuvent s'écrire sous la forme¹²

$$\varepsilon_i = \sum_{s=0}^{\infty} \rho^s u_{i-s}$$

Sous l'hypothèse que les aléas u_i sont des bruits blancs, on a donc

1. $E(\varepsilon_i) = \sum_{s=0}^{\infty} \rho^s E(u_{i-s}) = 0$,
2. $var(\varepsilon_i) = \sigma_u^2 / (1 - \rho^2)$, c'est-à-dire que les aléas sont homoscédastiques.
En effet,

$$\begin{aligned} var(\varepsilon_i) &= E(\varepsilon_i^2) = E\left(\left(\sum_{s=0}^{\infty} \rho^s u_{i-s}\right)^2\right) \\ &= \sum_{s=0}^{\infty} \rho^{2s} E(u_{i-s}^2) = \sigma_u^2 \sum_{s=0}^{\infty} \rho^{2s} = \frac{\sigma_u^2}{1 - \rho^2} \end{aligned}$$

3. $cov(\varepsilon_i, \varepsilon_{i-s}) = E(\varepsilon_i \varepsilon_{i-s}) = \rho^s (\sigma_u^2 / (1 - \rho^2))$.

En effet

$$\begin{aligned} E(\varepsilon_i \varepsilon_{i-s}) &= E((\rho^s \varepsilon_{i-s} + \rho^{s-1} u_{i-(s-1)} + \dots + \rho u_{i-1} + u_i) \cdot (\varepsilon_{i-s})) \\ &= \rho^s var(\varepsilon_{i-s}) \end{aligned}$$

car ε_i ne dépend que des valeurs passées et présentes du bruit blanc u_i ¹³.

Ainsi, la matrice de variance-covariance des aléas ε_i s'écrit

$$Var_{(N,N)}(\varepsilon) = E(\varepsilon \varepsilon') = \sigma_u^2 \Omega \quad (6.39)$$

$$= \sigma_u^2 \begin{bmatrix} 1/(1-\rho^2) & \rho/(1-\rho^2) & \cdot & \cdot & \rho^{N-1}/(1-\rho^2) \\ \rho/(1-\rho^2) & 1/(1-\rho^2) & \cdot & \cdot & \rho^{N-2}/(1-\rho^2) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho^{N-1}/(1-\rho^2) & \rho^{N-2}/(1-\rho^2) & \cdot & \cdot & 1/(1-\rho^2) \end{bmatrix}$$

¹¹ On ne considère pas ici un modèle dynamique, cf. section suivante.

¹² $\varepsilon_i = \rho \varepsilon_{i-1} + u_i = \rho^2 \varepsilon_{i-2} + \rho u_{i-1} + u_i = \dots$

$= \rho^s \varepsilon_{i-s} + \rho^{s-1} u_{i-(s-1)} + \dots + \rho u_{i-1} + u_i$

¹³ $E((\rho^{s-1} u_{i-(s-1)} + \dots + \rho u_{i-1} + u_i) \cdot (\varepsilon_{i-s})) = 0$

La matrice de transformation $\Omega^{-1/2}$ est alors égale à

$$\Omega^{-1/2} = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & \cdot & \cdot & \cdot & 0 \\ -\rho & 1 & \cdot & \cdot & \cdot & 0 \\ 0 & -\rho & 1 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & -\rho & 1 \end{bmatrix} \quad (6.40)$$

2.3.3 Estimation des paramètres en cas d'autocorrélation des aléas

Afin d'obtenir un estimateur efficace on applique les MCG, c'est-à-dire les MCO sur le modèle transformé. Dans le cas d'un modèle avec autocorrélation d'ordre 1 des aléas, le modèle transformé s'écrit

$$Y_* = X_*\beta + \varepsilon_* \quad (6.41)$$

avec

$$Y_* = \begin{bmatrix} (\sqrt{1-\rho^2}) y_1 \\ y_2 - \rho y_1 \\ \cdot \\ y_N - \rho y_{N-1} \end{bmatrix}, \varepsilon_* = \begin{bmatrix} (\sqrt{1-\rho^2}) \varepsilon_1 \\ \varepsilon_2 - \rho \varepsilon_1 \\ \cdot \\ \varepsilon_N - \rho \varepsilon_{N-1} \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ u_N \end{bmatrix} \text{ et}$$

$$X_* = \begin{bmatrix} \sqrt{1-\rho^2} & (\sqrt{1-\rho^2}) x_{21} & \cdot & \cdot & (\sqrt{1-\rho^2}) x_{K1} \\ (1-\rho) & x_{22} - \rho x_{21} & \cdot & \cdot & x_{K2} - \rho x_{K1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ (1-\rho) & x_{2N} - \rho x_{2N-1} & \cdot & \cdot & x_{KN} - \rho x_{KN-1} \end{bmatrix}$$

Par conséquent, pour la première observation, le modèle s'écrit

$$\begin{aligned} (\sqrt{1-\rho^2}) y_1 &= \beta_1 (\sqrt{1-\rho^2}) + \beta_2 (\sqrt{1-\rho^2} x_{21}) \\ &\quad + \dots + \beta_K (\sqrt{1-\rho^2} x_{K1}) + (\sqrt{1-\rho^2}) \varepsilon_1 \end{aligned} \quad (6.42)$$

Et pour les autres observations

$$\begin{aligned} y_i - \rho y_{i-1} &= \beta_1 (1-\rho) + \beta_2 (x_{2i} - \rho x_{2i-1}) + \dots + \beta_K (x_{Ki} - \rho x_{Ki-1}) + u_i \\ i &= 2, \dots, N \end{aligned} \quad (6.43)$$

La méthode des MCG consiste à appliquer les MCO sur ce modèle transformé.

Lorsque le coefficient d'autocorrélation est inconnu on applique l'estimateur des MCQG. Il existe plusieurs procédures. Elles sont différentes car

- elles n'utilisent pas le même nombre d'observations. Par exemple les procédures de Cochrane-Orcutt et de Hildreth-Lu considèrent les observations $i = 2, \dots, N$, tandis que la procédure de Prais-Winsten utilise l'ensemble des observations,
- elles n'ont pas recours aux mêmes estimateurs convergents de ρ pour transformer le modèle.

Une méthode très utilisée est la méthode de Cochrane-Orcutt. Elle procède en plusieurs étapes

1. Estimer avec la méthode des MCO le modèle initial et évaluer les résidus e_i sur les observations $i = 2, \dots, N$.
2. Estimer ρ par $\hat{\rho}_{co}$, où $\hat{\rho}_{co}$ est déterminé de manière à minimiser la fonction

$$\text{Min}_{\rho} \sum_{i=2}^N (e_i - \rho e_{i-1})^2$$

on obtient ainsi

$$\hat{\rho}_{co} = \frac{\sum_{i=2}^N (e_i \cdot e_{i-1})}{\sum_{i=2}^N e_{i-1}^2}$$

3. Appliquer les MCO sur le modèle transformé, appelé le modèle quasi-différencié

$$y_i - \hat{\rho}_{co} y_{i-1} = \beta_1 (1 - \hat{\rho}_{co}) + \beta_2 (x_{2i} - \hat{\rho}_{co} x_{2i-1}) + \dots \\ + \beta_K (x_{Ki} - \hat{\rho}_{co} x_{Ki-1}) + u_i \quad i = 2, \dots, N$$

On obtient ainsi un estimateur asymptotiquement identique à l'estimateur des MCG.

4. On utilise généralement la méthode de Cochrane-Orcutt itérative afin d'améliorer les performances de l'estimateur. Dans l'étape 3 on calcule alors les résidus du modèle.

$$e_i = y_i - \hat{y}_i$$

avec

$$\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_{2i} + \dots + \hat{\beta}_K x_{Ki}$$

On reprend ensuite la procédure à l'étape 2. La procédure s'arrête lorsque $\hat{\rho}_{co}$ ne varie plus de manière significative entre deux itérations.

La méthode de Prais-Winsten procède comme la méthode de Cochrane Orcutt mais elle prend en compte la première observation. Le coefficient d'autocorrélation ρ est estimé de manière à minimiser la fonction

$$\underset{\rho}{\text{Min}} (1 - \rho^2) e_1^2 + \sum_{i=2}^N (e_i - \rho e_{i-1})^2$$

on obtient ainsi

$$\widehat{\rho} = \frac{\sum_{i=2}^N (e_i \cdot e_{i-1})}{\sum_{i=3}^N e_{i-1}^2}$$

La méthode de Hildreth-Lu estime le modèle aux quasi-différences ($i = 2, \dots, N$), par la méthode des MCO, pour différentes valeurs de ρ comprises entre -1 et 1. On choisit $\widehat{\rho}$ qui minimise la somme des carrés des résidus des modèles ainsi estimés.

2.3.4 Préviation en présence d'autocorrélation des aléas

Pour réaliser une prévision lorsque les aléas sont autocorrélés d'ordre 1 on doit utiliser les estimations du modèle transformé

$$\widehat{y}_{*i} = \widehat{\beta}' x_{*i}$$

où $\widehat{\beta}'$ est l'estimateur des MCO du modèle transformé (6.41).

Lorsque ρ est connu, pour la période $N + 1$ la prévision optimale de y étant donné x_{*N+1} s'écrit

$$\widehat{y}_{*N+1} = \widehat{\beta}' x_{*N+1} \Leftrightarrow \widehat{y}_{N+1} - \rho y_N = \widehat{\beta}' x_{N+1} - \widehat{\beta}' \rho x_N$$

soit

$$\widehat{y}_{N+1} = \widehat{\beta}' x_{N+1} + \underbrace{\rho (y_N - \widehat{\beta}' x_N)}_{e_N}$$

Plus généralement, pour la période $N + j$ la prévision optimale est donnée par

$$\widehat{y}_{N+j} = \widehat{\beta}' x_{N+j} + \underbrace{\rho^j (y_N - \widehat{\beta}' x_N)}_{e_N} \quad (6.44)$$

Ce résultat est justifié car $E(\varepsilon_{N+1}/\varepsilon_N) = \rho \varepsilon_N$.

2.4 La correction de White

Si la structure de la matrice Ω est parfaitement inconnue, l'estimateur des MCO reste le seul disponible. Il est non biaisé, convergent et distribué selon une loi normale mais il n'est pas efficace. Sa matrice de variance-covariance est donnée par l'équation (6.14). La stratégie consiste alors à déterminer un estimateur convergent de la matrice de variance-covariance asymptotique de l'estimateur des MCO.

Dans le cas d'aléas hétéroscédastiques la matrice de variance-covariance de l'estimateur des MCO est donnée par

$$(X'X)^{-1} (X'\Omega X) (X'X)^{-1} \quad (6.45)$$

avec

$$\Omega = \begin{bmatrix} \sigma_1^2 & 0 & \cdot & \cdot & 0 \\ 0 & \sigma_2^2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \sigma_N^2 \end{bmatrix}$$

Soit

$$\Sigma = \frac{X'\Omega X}{N} = \frac{1}{N} \sum_{i=1}^N \sigma_i^2 x_i x_i' \quad (6.46)$$

White a montré, sous des conditions très générales, que la matrice

$$\widehat{\Sigma} = \frac{1}{N} \sum_{i=1}^N e_i^2 x_i x_i' \quad (6.47)$$

où e_i^2 est le carré des résidus MCO, est un estimateur convergent de Σ . Ainsi, lorsque la forme de l'hétéroscédasticité des aléas est inconnue, White propose d'estimer le modèle par les MCO et de calculer la matrice de variance-covariance de $\widehat{\beta}_{MCO}$ selon la formule

$$N (X'X)^{-1} \widehat{\Sigma} (X'X)^{-1} \quad (6.48)$$

L'extension de la procédure de White au cas de l'autocorrélation est plus difficile car un problème de convergence de l'estimateur de la matrice $\frac{X'\Omega X}{N}$ se pose.

3. LES VARIABLES EXPLICATIVES DU MODÈLE

Dans cette section on s'intéresse à l'hypothèse 3 du modèle de régression linéaire multiple. Sous cette hypothèse les variables explicatives du modèle sont

supposées non aléatoires. Dans un premier temps, sont examinées les propriétés de l'estimateur des MCO lorsque les variables explicatives sont aléatoires. Dans un second temps, l'estimateur des Variables Instrumentales est présenté. Ensuite, deux cas d'application de l'estimateur des Variables Instrumentales sont discutés. Pour terminer, l'estimateur des Moments Généralisés est présenté.

3.1 Variables explicatives aléatoires et estimateur des MCO

Les variables explicatives peuvent être des variables aléatoires. Dans ce cas les propriétés de l'estimateur des MCO diffèrent selon la relation entre les variables explicatives et les aléas : ils peuvent être indépendants, non indépendants mais non corrélés, ou encore corrélés. Lorsque les variables explicatives sont aléatoires l'hypothèse 5 devient l'hypothèse 5' sur le modèle de régression linéaire multiple.

Hypothèse 5'

$p \lim S_X = p \lim \frac{X'X}{N} = V_X$ avec V_X une matrice finie, définie positive.

L'estimateur des MCO des paramètres du modèle $Y = X\beta + \varepsilon$ est une fonction linéaire des aléas et peut s'écrire sous la forme

$$\widehat{\beta}_{MCO} = \beta + (X'X)^{-1} X'\varepsilon \quad (6.49)$$

Conditionnellement aux réalisations x_i , on a

$$E(\widehat{\beta}_{MCO}/X) = \beta + (X'X)^{-1} X'E(\varepsilon/X) \quad (6.50)$$

et en utilisant la loi des espérances itérées

$$E(\widehat{\beta}_{MCO}) = \beta + E_X \left((X'X)^{-1} X'E(\varepsilon/X) \right) \quad (6.51)$$

On distingue alors 3 cas

1. Si les variables explicatives sont indépendantes des aléas c'est-à-dire $E(\varepsilon/X) = 0$, alors l'estimateur des MCO est non biaisé $E(\widehat{\beta}_{MCO}) = \beta$. Sa matrice de variance-covariance conditionnelle s'écrit

$$Var(\widehat{\beta}_{MCO}/X) = \sigma^2 (X'X)^{-1}$$

et sa matrice de variance-covariance non conditionnelle

$$\begin{aligned} Var(\widehat{\beta}_{MCO}) &= E_X \left(Var(\widehat{\beta}_{MCO}/X) \right) + Var_X \left(\underbrace{E(\widehat{\beta}_{MCO}/X)}_{=\beta} \right) \\ &= E \left(Var(\widehat{\beta}_{MCO}/X) \right) = E \left(\sigma^2 (X'X)^{-1} \right) \\ &= \sigma^2 E \left((X'X)^{-1} \right) \end{aligned} \quad (6.52)$$

Donc si les variables explicatives et les aléas sont indépendants, sous les hypothèses 1, 2, 4, 5' et 6 l'estimateur des MCO est BLUE, distribué selon une loi normale et asymptotiquement efficace.

2. Si les variables explicatives sont non indépendantes ($E(\varepsilon/X) \neq 0$) mais non corrélées aux aléas

$$\text{cov}(x_{ki}, \varepsilon_i) = E(x_{ki} \cdot \varepsilon_i) = 0, \quad k = 1, \dots, K \text{ et } i = 1, \dots, N \quad (6.53)$$

lorsque $N \rightarrow \infty$

$$P \lim \frac{X' \varepsilon}{N} = 0 \quad (6.54)$$

alors l'estimateur des MCO est convergent

$$P \lim \widehat{\beta}_{MCO} = \beta + P \lim \left(\frac{X'X}{N} \right)^{-1} P \lim \left(\frac{X'\varepsilon}{N} \right) = \beta \quad (6.55)$$

Il est asymptotiquement distribué selon une loi normale et asymptotiquement efficace. L'estimateur des MCO perd ses bonnes propriétés à distance finie.

3. Enfin si les variables explicatives et les aléas sont corrélés alors l'estimateur des MCO n'est pas convergent.

Dans ce cas $P \lim \frac{X'\varepsilon}{N} \neq 0$ et $P \lim \widehat{\beta}_{MCO} \neq \beta$. Le problème consiste alors à définir un estimateur convergent de β .

3.2 L'estimateur des Variables Instrumentales

L'estimateur des Variables Instrumentales est appliqué lorsque la covariance entre les aléas et les variables explicatives est non nulle. L'idée est de rechercher des variables fortement corrélées aux variables explicatives mais non corrélées aux aléas. Soit Z la matrice des observations des variables instrumentales d'ordre (N, K) , la matrice Z est déterminée de manière à vérifier les deux conditions suivantes lorsque $N \rightarrow \infty$

$$P \lim \left(\frac{Z'X}{N} \right) = V_{Z'X} \quad \text{et} \quad P \lim \left(\frac{Z'\varepsilon}{N} \right) = 0 \quad (6.56)$$

L'estimateur des Variables Instrumentales $\widehat{\beta}_{VI}$ s'écrit

$$\widehat{\beta}_{VI} = (Z'X)^{-1} Z'Y = \beta + (Z'X)^{-1} Z'\varepsilon \quad (6.57)$$

Ses propriétés asymptotiques sont les suivantes

$$p \lim \widehat{\beta}_{VI} = \beta \quad \text{car} \quad p \lim \left(\frac{Z'\varepsilon}{N} \right) = 0 \quad (6.58)$$

et sa matrice de variance-covariance asymptotique peut être estimée avec

$$\boxed{\text{Asym.var} \left(\widehat{\beta}_{VI} \right) = \widehat{\sigma}^2 (Z'X)^{-1} (Z'Z) (X'Z)^{-1}} \quad (6.59)$$

$$\text{où } \widehat{\sigma}^2 = \frac{(Y - X\widehat{\beta}_{VI})' (Y - X\widehat{\beta}_{VI})}{N}$$

L'estimateur des Variables Instrumentales est convergent et distribué selon une loi normale sous les hypothèses 1, 2, 4, 5' et 6.

Afin de gagner en précision, il est conseillé de déterminer une matrice des instruments Z qui contient L colonnes avec $L > K$. L'estimateur des Variables Instrumentales s'écrit dans ce cas

$$\boxed{\widehat{\beta}_{VI} = (\widehat{X}'X)^{-1} \widehat{X}'Y = (X'P_ZX)^{-1} (X'P_Zy)} \quad (6.60)$$

avec $\widehat{X} = Z(Z'Z)^{-1}Z'X$ et $P_Z = Z(Z'Z)^{-1}Z'$

On remarque que $(Z'Z)^{-1}Z'X$ est l'estimateur des MCO des paramètres du modèle

$$X = Z\delta + \varepsilon$$

$$\widehat{\delta}_{MCO} = (Z'Z)^{-1}Z'X$$

et \widehat{X} est donné par

$$\widehat{X} = Z\widehat{\delta}_{MCO} = \underbrace{Z}_{(N,L)} \left[\underbrace{(Z'Z)^{-1}}_{(L,L)} \underbrace{Z'X}_{(L,K)} \right]$$

\widehat{X} est une matrice d'ordre (N, K) comme la matrice X . On peut vérifier que

$$\boxed{\widehat{\beta}_{VI} = (\widehat{X}'X)^{-1} \widehat{X}'Y = (\widehat{X}'\widehat{X})^{-1} \widehat{X}'Y} \quad (6.61)$$

car $\widehat{X}'X = \widehat{X}'\widehat{X}$.

Sa matrice de variance-covariance asymptotique s'écrit

$$\boxed{\text{Asym.var} \left(\widehat{\beta}_{VI} \right) = \widehat{\sigma}^2 (\widehat{X}'\widehat{X})^{-1} = \widehat{\sigma}^2 (X'P_ZX)^{-1}} \quad (6.62)$$

Il n'est pas toujours facile de déterminer si les variables explicatives du modèle sont corrélées avec les aléas et de faire un choix entre l'estimateur des Variables Instrumentales et l'estimateur des MCO. Le test d'Hausman apporte des éléments de réponse à cette interrogation.

3.3 Test d'Hausman

Avec le test d'Hausman on teste l'hypothèse H_0 d'une covariance nulle entre les variables explicatives et les aléas. Sous l'hypothèse H_0 , les estimateurs des MCO et des Variables Instrumentales sont convergents, l'estimateur des MCO étant plus précis. Sous l'hypothèse H_1 , la covariance est non nulle et l'estimateur des MCO est biaisé et non convergent, tandis que l'estimateur des Variables Instrumentales est convergent. Le test consiste à vérifier si l'écart entre les deux estimateurs est significatif ou non. S'il n'est pas significatif, on accepte H_0 . On calcule la statistique de Wald

$$Wald = \left(\hat{\beta}_{VI} - \hat{\beta}_{MCO} \right)' \left(\hat{\sigma}^2 \left(\left(\hat{X}'\hat{X} \right)^{-1} - (X'X)^{-1} \right) \right)^{-1} \left(\hat{\beta}_{VI} - \hat{\beta}_{MCO} \right) \quad (6.63)$$

avec $\hat{\sigma}^2$ un estimateur convergent de σ^2 . Sous l'hypothèse H_0 , on peut utiliser $\hat{\sigma}^2 = \frac{e'e}{N-K}$.

Le degré de liberté de cette statistique est égal à K lorsque les matrices X et Z n'ont pas de variables communes, c'est-à-dire lorsque sous H_0 on teste l'hypothèse $cov(x_{ki}, \varepsilon_i) = 0, \forall k = 1, \dots, K$. Si $Wald < \chi^2_{\alpha}(K)$ alors l'hypothèse H_0 n'est pas refusée et on peut appliquer l'estimateur des MCO.

Cependant usuellement X et Z comportent des variables communes¹⁴. Dans ce cas, lorsque le nombre de variables présentes dans les deux matrices est égal à K_0 , le degré de liberté du test est égal à $K - K_0$. Autrement dit, l'hypothèse d'une covariance nulle entre les variables explicatives et les aléas est testée pour un nombre de variables $K - K_0$, les K_0 variables sont supposées avoir une covariance nulle avec les aléas.

Pour terminer, on note qu'il est équivalent de réaliser le test d'exogénéité d'Hausman d'une variable x_k ou de mettre en œuvre une régression du modèle initial augmenté des résidus MCO obtenus lors de l'estimation de la variable x_k en fonction des variables de la matrice Z . La non significativité du paramètre associé à ces résidus conduit à ne pas refuser l'hypothèse nulle d'exogénéité.

¹⁴ Par exemple, la constante est généralement présente dans les deux matrices.

Jude-Henri Jeaurio

3.4 Applications de l'estimateur des Variables Instrumentales

Deux cas sont examinés : le cas d'une erreur de mesure sur une variable explicative et le cas d'un modèle dynamique dans lequel les aléas sont autocorrélés.

3.4.1 Cas d'une erreur de mesure sur une variable explicative

On considère un modèle de régression linéaire simple centré, dans lequel la variable explicative x_i^* n'est pas directement observée

$$y_i = \beta x_i^* + \varepsilon_i \quad i = 1, \dots, N \quad (6.64)$$

on observe la variable x

$$x = x^* + v \quad (6.65)$$

où v est un aléa distribué selon une loi normale avec une espérance nulle, une variance notée σ_v^2 et une covariance nulle. De plus, on suppose que ε et v sont des variables aléatoires indépendantes. Étant donné les observations dont on dispose, on estime le modèle

$$y = \beta x - \beta v + \varepsilon \quad (6.66)$$

Soit $w = -\beta v + \varepsilon$, la covariance entre x et w étant non nulle l'estimateur des MCO est non convergent. En effet

$$\text{cov}(x, w) = \text{cov}(x^* + v, -\beta v + \varepsilon) = -\beta \sigma_v^2 \quad (6.67)$$

et

$$\widehat{\beta}_{MCO} = \frac{\sum_{i=1}^N x_i y_i}{\sum_{i=1}^N x_i^2} = \beta + \frac{\sum_{i=1}^N x_i w_i}{\sum_{i=1}^N x_i^2} \quad (6.68)$$

la probabilité limite de l'estimateur des MCO est donnée par

$$P \lim \widehat{\beta}_{MCO} = \beta + \frac{P \lim \frac{\sum_{i=1}^N x_i w_i}{N}}{P \lim \frac{\sum_{i=1}^N x_i^2}{N}} = \beta - \beta \frac{\sigma_v^2}{\sigma_x^2} \quad (6.69)$$

avec $P \lim \frac{\sum_{i=1}^N x_i^2}{N} = \sigma_x^2$. Par conséquent, lorsque la variable explicative est mesurée avec une erreur l'estimateur des MCO est non convergent, $P \lim \widehat{\beta}_{MCO} \neq \beta$. Dans ce cas il faut rechercher des variables fortement corrélées à cette variable explicative mais non corrélées à l'aléa afin d'appliquer l'estimateur des Variables Instrumentales.

3.4.2 Cas d'un modèle dynamique avec autocorrélation des aléas

Les modèles dynamiques dont les aléas sont autocorrélés doivent être estimés avec l'estimateur des Variables Instrumentales.

Soit le modèle de régression

$$y_i = \beta_1 + \beta_2 x_i + \beta_3 y_{i-1} + \varepsilon_i \quad i = 1, \dots, N \tag{6.70}$$

Si les aléas du modèle sont autocorrélés d'ordre 1

$$\text{cov}(\varepsilon_i, \varepsilon_{i-1}) = \rho \tag{6.71}$$

alors la covariance entre y_{i-1} et ε_i est différente de 0

$$\text{cov}(y_{i-1}, \varepsilon_i) = \text{cov}(\varepsilon_i, \varepsilon_{i-1}) = \rho \tag{6.72}$$

Il faut dans ce cas appliquer l'estimateur des Variables Instrumentales. Par exemple, la variable x_{i-1} est une variable fortement corrélée à la variable y_{i-1} mais non corrélée aux aléas du modèle.

$$p \lim \frac{\sum_{i=1}^N x_{i-1} y_{i-1}}{N} = \sigma_{xy} \quad \text{et} \quad p \lim \frac{\sum_{i=1}^N x_{i-1} \varepsilon_i}{N} = 0$$

Dans ce cas

$$X = \begin{pmatrix} 1 & x_1 & \cdot \\ 1 & x_2 & y_1 \\ \cdot & x_3 & y_2 \\ \cdot & \cdot & \cdot \\ 1 & x_N & y_{N-1} \end{pmatrix} \quad \text{et} \quad Z = \begin{pmatrix} 1 & x_1 & \cdot \\ 1 & x_2 & x_1 \\ \cdot & x_3 & x_2 \\ \cdot & \cdot & \cdot \\ 1 & x_N & x_{N-1} \end{pmatrix}$$

et l'estimateur des Variables Instrumentales est égal à $\widehat{\beta}_{VI} = (Z'X)^{-1} Z'Y$.

De manière générale, dans un modèle dynamique, il faut dans un premier temps tester l'hypothèse d'autocorrélation des aléas du modèle avec la statistique h de Durbin (6.37). Et si l'hypothèse d'autocorrélation n'est pas rejetée, il faut estimer le modèle avec l'estimateur des Variables Instrumentales. Une procédure alternative consiste à appliquer le test d'Hausman afin de tester l'hypothèse d'une covariance nulle entre la variable endogène décalée et l'aléa du modèle dynamique.

3.5 L'estimateur de la Méthode des Moments Généralisés

La Méthode des Moments Généralisés s'applique dans le cadre du modèle linéaire généralisé lorsque les variables explicatives sont corrélées aux aléas.

Le modèle s'écrit

$$\begin{cases} Y = X\beta + \varepsilon \\ E(\varepsilon\varepsilon') = \sigma^2\Omega \\ P \lim \frac{X'\varepsilon}{N} \neq 0 \end{cases} \quad (6.73)$$

L'estimateur des Variables Instrumentales du modèle (6.73) est convergent mais asymptotiquement il n'est pas le plus précis. L'estimateur des moments généralisés est convergent et donne une meilleure précision, il est donné par

$$\widehat{\beta}_{GMM} = (X'P_M X)^{-1} X'P_M Y \quad (6.74)$$

Avec

$$P_M = Z(Z'\Omega Z)^{-1} Z'$$

et Z la matrice des instruments telle que

$$P \lim \frac{Z'X}{N} = V_{Z'X} \quad \text{et} \quad P \lim \frac{Z'\varepsilon}{N} = 0$$

Sa matrice de variance-covariance asymptotique s'écrit

$$\text{Var}_{as} \left(\widehat{\beta}_{GMM} \right) = \sigma^2 (X'P_M X)^{-1} \quad (6.75)$$

Les éléments de la matrice Ω étant inconnu ont procédé en 2 étapes :

1. On calcule $\widehat{\Omega}$ un estimateur convergent de Ω (Ω peut être obtenu à partir des résidus du modèle (6.73) estimés avec l'estimateur des variables instrumentales)
2. Ayant $\widehat{\Omega}$ on calcule $\widehat{\beta}_{GMM} = (X'\widehat{P}_M X)^{-1} X'\widehat{P}_M Y$ avec $\widehat{P}_M = Z (Z'\widehat{\Omega} Z)^{-1} Z'$

Pour tester l'exogénéité des variables explicatives on peut utiliser le test d'Hausman et calculer

$$Wald = \left(\hat{\beta}_{GMM} - \hat{\beta}_{MCG} \right)' \left(\widehat{Var}_{as} \left(\hat{\beta}_{GMM} \right) - \widehat{Var}_{as} \left(\hat{\beta}_{MCG} \right) \right)^{-1} \left(\hat{\beta}_{GMM} - \hat{\beta}_{MCG} \right) \quad (6.76)$$

Avec

$$\left(\widehat{Var}_{as} \left(\hat{\beta}_{GMM} \right) - \widehat{Var}_{as} \left(\hat{\beta}_{MCG} \right) \right) = \hat{\sigma}^2 \left(X'Z(Z'\hat{\Omega}Z)^{-1}Z'X \right)^{-1} - \left(X'\hat{\Omega}X \right)^{-1}$$

Par ailleurs, lorsque la matrice des instruments est de dimension (N, p) avec $p > K$ on peut tester la validité des instruments avec le test de Sargan

Sous l'hypothèse nulle les variables instrumentales et les aléas sont non corrélés asymptotiquement.

On calcule la statistique S

$$S = \frac{1}{\hat{\sigma}^2} e' \hat{P}_M e \quad (6.77)$$

Sous H_0 , S est distribué asymptotiquement selon un χ^2 à $p - K$ degrés de libertés.

Si S est inférieur à la valeur tabulée on accepte l'hypothèse de validité des instruments.

7

LE MODÈLE D'ÉVALUATION DES ACTIFS FINANCIERS

1. Introduction
2. Le MEDAF : hypothèses et relations d'équilibre
3. Estimations et validation sur les sociétés du CAC 40
4. Éléments de corrigé

OBJECTIFS DU CHAPITRE

D'un point de vue méthodologique

- Estimation d'un modèle de régression simple
- Recherche d'une forme testable d'un modèle théorique
- Test d'autocorrélation d'ordre 1 (Durbin-Watson)
- Méthodes de correction de l'autocorrélation

D'un point de vue économique

- Présentation du MEDAF
- Application et validation sur des sociétés composant l'indice CAC 40

NIVEAU

LICENCE, troisième année

CHAMP ÉCONOMIQUE TRAITÉ

Microéconomie financière

1. INTRODUCTION

Développé dans les années 1960 par trois économistes américains (Sharpe, Lintner et Mossin)¹, le modèle d'évaluation des actifs financiers (MEDAF) apparaît en première approche comme le contrepoint financier de la théorie de l'équilibre général (Walras, Arrow-Debreu). Le modèle d'équilibre général propose une explication de la formation des prix sur les différents marchés microéconomiques des biens et services. Le MEDAF (Capital Asset Pricing Model ou CAPM, en anglais) s'attache à décrire les conditions d'équilibre et la formation des prix sur les marchés d'actifs financiers. S'il constitue donc par son objet et son ambition, une construction tout à fait centrale de l'analyse économique, le MEDAF n'en demeure pas moins un instrument extrêmement utile sur un plan pratique, pour tout investisseur ayant à effectuer des choix d'investissement en univers incertain et en particulier pour les gestionnaires de portefeuilles sur les marchés d'actions (OPCVM, mutual funds, fonds de pension...).

Le résultat principal du MEDAF est l'identification d'une relation linéaire qui doit prévaloir, à l'équilibre des marchés financiers, entre l'excès de rendement d'un titre risqué par rapport au taux d'intérêt sans risque (la prime de risque) et l'excès de rendement du portefeuille de marché par rapport à ce même taux sans risque. Le coefficient affectant l'excès de rendement du portefeuille de marché par rapport au taux sans risque est le fameux bêta, bien connu des praticiens. Il mesure la sensibilité du rendement de chaque titre par rapport au rendement du portefeuille de marché.

Le coefficient bêta s'interprète également comme le risque systématique du titre considéré, c'est-à-dire la variance du rendement du titre imputable à la variance du rendement du portefeuille de marché. Il existe ainsi, à l'équilibre des marchés, une relation croissante entre l'espérance de rendement des titres et leur risque systématique. Cette relation permet donc de nuancer les enseignements de la théorie de base du portefeuille (Markowitz) en montrant que s'il y a bien une relation croissante entre rendement et risque, le marché rémunère uniquement le risque systématique des titres et pas leur risque intrinsèque ou diversifiable.

L'application économétrique développée dans ce chapitre cherche à estimer et tester économétriquement le MEDAF en s'appuyant pour cela sur les sociétés de l'indice français CAC 40. Les séries de cours et de rendements boursiers nécessaires à l'estimation du modèle sont observées selon une fréquence hebdomadaire.

Sur un plan économétrique, l'application se situe dans le cadre théorique du modèle de régression simple. Elle permet d'illustrer un problème courant en économétrie appliquée qu'est la difficulté d'estimer en l'état le modèle proposé par la théorie économique. Il convient donc de présenter et de discuter les transformations effectuées sur le modèle économique pour procéder à son estimation économétrique. L'estimation du MEDAF par les MCO donne lieu à une double validation. Dans un premier temps, la validation économétrique de la relation estimée nous impose d'effectuer le test d'autocorrélation de Durbin-Watson et de corriger le cas échéant

¹ Voir notamment W. Sharpe, 1964, « Capital asset prices : a theory of market equilibrium under conditions of risk », *Journal of Finance*, septembre.

les autocorrélations observées par l'estimation, selon les procédures courantes, d'un modèle quasi-différencié. Les résultats obtenus fournissent alors une illustration du théorème de Gauss-Markov précisant les propriétés des estimateurs des MCO et des MCG dans le cas d'une matrice de variance-covariance des aléas non scalaire. Au terme de la validation statistique de la relation estimée, on procède à une validation économique du MEDAF en confrontant les résultats d'estimation sur les titres de l'indice CAC 40 aux principaux résultats théoriques du modèle.

Le chapitre est organisé de la manière suivante. La section 2 présente les principales hypothèses et relations d'équilibre du MEDAF. La section 3 est consacrée aux estimations économétriques sur les sociétés cotées composant l'indice CAC40.

2. LE MEDAF : HYPOTHÈSES ET RELATIONS D'ÉQUILIBRE

Le MEDAF cherche à comprendre la formation des prix et des rendements des titres négociables sur les marchés financiers. Il décrit pour cela les comportements d'achat et de vente d'investisseurs supposés rationnels en s'appuyant directement sur la théorie du portefeuille de Markowitz (1952-1959)².

2.1 Hypothèses

Trois hypothèses principales suffisent pour décliner les propriétés d'équilibre du MEDAF.

1. Tous les investisseurs présents sur le marché ont un comportement conforme à celui décrit par la théorie du portefeuille. Ils forment un système de demande pour n actifs risqués ($i = 1, \dots, n$) en maximisant l'espérance de l'utilité de leur richesse ou de manière équivalente un critère espérance-variance du rendement total de leur portefeuille. Soit formellement le programme suivant

$$\underset{\alpha_i}{\text{Max}} U = E(\tilde{R}) - \frac{\theta}{2} V(\tilde{R})$$

où

\tilde{R} désigne le rendement aléatoire du portefeuille tel que $\tilde{R} = \sum_{i=1}^n \alpha_i \tilde{R}_i$

α_i est la part de la richesse investie sur le titre i de rendement \tilde{R}_i

$E(\tilde{R})$ et $V(\tilde{R})$ représentent l'espérance et la variance du rendement du portefeuille. $V(\tilde{R})$ s'interprète aussi comme le risque du portefeuille.

Le paramètre θ correspond à l'aversion au risque des investisseurs.

² H. Markowitz, 1952, « Portfolio selection », *Journal of Finance*, mars.

H. Markowitz, 1959, *Portfolio selection, efficient diversification of investments*, John Wiley and Sons, New-York.

2. Ces investisseurs ont le même horizon d'investissement (1 semaine, 1 mois...) qui correspond à la période sur laquelle sont calculés les rendements des titres. Les anticipations des investisseurs sur cette période sont homogènes, ce qui signifie qu'ils intègrent dans leur programme d'investissement le même vecteur d'espérance de rendement $\mu' = (\mu_1, \dots, \mu_n)$ et la même matrice de variance-covariance Ω (d'éléments σ_{ij}). Concrètement l'expression des rendements varie avec le type de placement considéré (actions, obligations, actifs immobiliers...). Pour les placements en actions, le rendement correspond au taux de croissance des prix sur la période d'épargne corrigé du versement éventuel de dividendes. Le caractère aléatoire des rendements et la nature risquée du placement sont dus au fait que les prix de fin de période sont inconnus au début de la période d'épargne.
3. Il existe dans le système financier un actif sans risque de rendement r .

2.2 Relations d'équilibre

Deux types de relation d'équilibre peuvent être dérivés des hypothèses précédentes.

2.2.1 La droite des marchés de capitaux

Les investisseurs ayant des comportements parfaitement homogènes, ils définissent ex-ante la même frontière efficiente qui représente un sous-ensemble de portefeuilles offrant un meilleur compromis entre le rendement et le risque. On montre simplement que cette frontière efficiente correspond dans le plan (μ, σ) à la portion supérieure d'une hyperbole. En présence d'un actif sans risque de rendement r , les investisseurs peuvent combiner les portefeuilles efficients et l'actif sans risque de sorte que la frontière efficiente devient une demi-droite passant par r et offrant un point de tangence avec la portion d'hyperbole. Ce point de tangence correspond à un portefeuille risqué particulier demandé par tous les investisseurs. À l'équilibre, ce portefeuille correspond nécessairement au portefeuille de marché noté M dont l'espérance de rendement est μ_M et le risque σ_M .

On montre dès lors sans difficulté que l'équation de la demi-droite des portefeuilles efficients est donnée par

$$\mu - r + \left(\frac{\mu_M - r}{\sigma_M} \right) \sigma \quad (7.1)$$

où μ désigne le rendement anticipé des portefeuilles efficients et σ le risque de ces portefeuilles efficients.

Cette équation montre qu'il existe à l'équilibre une relation linéaire et croissante entre le rendement anticipé et le risque des portefeuilles efficients. Sa représentation dans le plan (μ, σ) est appelée la droite des marchés de capitaux. La pente de cette droite $\left(\frac{\mu_M - r}{\sigma_M} \right)$ est le prix de marché du risque pour ces portefeuilles. C'est le rapport de la prime de risque du marché (ou excès de rendement anticipés par rapport au taux sans risque $(\mu_M - r)$) sur le niveau du risque de marché σ_M . La

prime de risque de marché doit être positive pour que les investisseurs averses au risque détiennent des actifs risqués.

2.2.2 La relation d'équilibre du MEDAF

La relation d'équilibre donnée par la droite de marché (7.1) est pertinente pour caractériser les portefeuilles efficaces, mais ne concerne pas les actifs individuels de l'univers d'investissement. Le MEDAF en tant que modèle d'équilibre du marché des actifs financiers doit permettre d'exprimer la valeur des rendements d'équilibre des actifs risqués et indirectement celle des prix d'équilibre. On peut montrer que cette relation s'écrit

$$\mu_i = r + \left(\frac{\mu_M - r}{\sigma_M^2} \right) \sigma_{iM} \quad i = 1, \dots, n \quad (7.2)$$

avec μ_i le rendement anticipé de l'actif et σ_{iM} la covariance entre ce rendement et celui du portefeuille de marché.

En posant

$$\beta_i = \frac{\sigma_{iM}}{\sigma_M^2}$$

la relation devient

$$\mu_i = r + (\mu_M - r) \beta_i \quad i = 1, \dots, n \quad (7.3)$$

ou de manière équivalente

$$E(\tilde{R}_i) - r = [E(\tilde{R}_M) - r] \beta_i \quad i = 1, \dots, n \quad (7.4)$$

Cette relation d'équilibre est donc une relation linéaire entre la prime de risque de chaque titre et la prime de risque du marché. Elle ne doit pas être confondue avec ce que l'on appelle le modèle de marché de Sharpe dont elle est pourtant proche. En effet, Sharpe postule directement le modèle statistique suivant pour expliquer la rentabilité des titres

$$\tilde{R}_i = \alpha_i + \beta_i \tilde{R}_M + \tilde{\varepsilon}_i$$

où α_i (sans rapport avec le terme α_i de la section 2.1) et β_i sont des constantes propres à chaque titre et $\tilde{\varepsilon}_i$ un aléa non corrélé avec \tilde{R}_M d'espérance nulle.

En espérance, la relation donne

$$\mu_i = \alpha_i + \beta_i \mu_M$$

et en variance

$$\sigma_i^2 = \sigma_{\varepsilon_i}^2 + \beta_i^2 \sigma_M^2$$

Le risque d'un titre se décompose ainsi en un risque propre au titre dit intrinsèque ($\sigma_{\varepsilon_i}^2$) et un risque systématique lié au marché ($\beta_i^2 \sigma_M^2$). Le risque intrinsèque est aussi qualifié de risque diversifiable dans la mesure où il peut être dilué au sein d'un portefeuille.

Si comme Sharpe, on interprète le coefficient β de la relation (7.4) comme le risque systématique des titres, le MEDAF montre que l'excès de rendement accordé à chaque titre dépend de l'excès de rendement accordé au marché et de son risque systématique. En d'autres termes, à l'équilibre seul le risque systématique se trouve rémunéré. Ce résultat vient donc nuancer la relation croissante entre espérance de rendement et risque donnée par la frontière efficiente.

D'un point de vue pratique, la mesure du coefficient β et sa position par rapport à 1 deviennent cruciales pour appréhender le caractère défensif ou agressif des différents titres et déterminer les excès de rendement attendus en fonction d'un scénario sur les performances du marché.

3. ESTIMATIONS ET VALIDATION SUR LES SOCIÉTÉS DU CAC 40

3.1 *Passage du modèle théorique au modèle économétrique*

La relation d'équilibre du MEDAF (7.4) doit subir deux transformations avant de faire l'objet d'une estimation et de tests économétriques.

Les espérances mathématiques de rendement pour les titres risqués et le portefeuille de marché, qui représentent les rendements attendus en début de période sur ces placements, ne sont pas observables. La seule solution envisageable pour l'économètre consiste à retenir des séries de rendement constatés en fin de période. On passe ainsi d'une formulation ex-ante de l'équilibre à une formulation ex-post.

Il faut de plus introduire une constante dans le modèle économétrique pour ne pas prendre le risque de biaiser l'estimateur du coefficient β et celui de la variance des résidus. En outre la décomposition de la variance et l'interprétation du coefficient de détermination (R^2) n'a de sens qu'en présence d'un terme constant dans le modèle économétrique. On pose donc le modèle linéaire simple suivant pour estimer la relation d'équilibre du MEDAF

$$(R_i - r)_t = \alpha_i + \beta_i(R_m - r)_t + u_{i,t} \quad (7.5)$$

où R_i désigne le rendement constaté de l'actif i , et $u_{i,t}$ la perturbation aléatoire du modèle.

S'il est utile d'introduire une constante dans le modèle, on attend en revanche que son coefficient α soit non significativement différent de 0. Dans le cas contraire

(rejet de l'hypothèse nulle pour le test de significativité de Student), cela révélerait l'existence d'un excès de rendement anormal sur le titre i considéré et serait en contradiction avec les résultats du MEDAF.

Notons enfin qu'en estimant ce modèle avec la méthode des MCO, on doit obtenir par définition des estimateurs

$$\hat{\beta}_i = \frac{Cov((R_i - r), (R_m - r))}{Var(R_m - r)} \quad (7.6)$$

ce qui ne correspond à la définition théorique du $\beta = \frac{Cov(R_i, R_m)}{Var(R_m)}$ que sous l'hypothèse où la covariance (empirique) entre le taux sans risque et les rendements risqués et la variance (empirique) du rendement sans risque sont nulles, c'est-à-dire si le taux sans risque est constant sur la période d'estimation. Ce cas de figure n'est pas envisageable concrètement, même en travaillant sur une période courte avec des données quotidiennes. Le taux sans risque doit représenter le rendement d'un placement accessible à la plupart des investisseurs. Il faut donc considérer des placements du marché monétaire (TCN (Bons du trésor, Certificats de Dépôts, Billets de trésorerie), interbancaire) d'une maturité équivalente à la fréquence des séries de rendements retenues. Dans tous les cas, ces taux évoluent de manière quasi quotidienne selon les conditions de marché.

Dans le modèle théorique, le taux sans risque est par définition de variance nulle et de covariances nulles avec les actifs risqués. C'est une donnée non aléatoire qui se présente aux investisseurs au moment où ils doivent définir leurs stratégies d'investissement et leurs portefeuilles optimaux. Finalement, l'estimation économétrique du MEDAF nous fait glisser d'une représentation instantanée de l'équilibre du marché à une représentation en moyenne sur la période d'estimation.

3.2 Données

L'application économétrique proposée porte sur un sous ensemble de sociétés du CAC 40 dans sa composition du mois de juillet 1999. Les séries brutes, nécessaires à la construction des séries de rendement, sont extraites de la base Datastream et sont rassemblées dans le fichier chapitre7.xls³. Elles sont de fréquence hebdomadaire et couvrent la période janvier 1990 - juillet 1999 pour les sociétés présentes dans l'indice depuis au moins le 29.12.1989. Plusieurs sociétés n'entrent dans la base qu'à compter de leur intégration dans le CAC.

On sélectionne un panier de titres appartenant à des secteurs industriels différents : Banque — BNP, Automobile — Peugeot, Distribution — Carrefour, Telecom — France telecom, Informatique — Cap Gemini, Pétrole — Elf Média — Canal+.

³ Le fichier de données correspondant à l'ensemble des sociétés du CAC 40 dans sa composition du mois de juillet 1999 est également disponible sur le site web.

Les données brutes comprennent les séries suivantes

cac = valeur de l'indice CAC 40 en points

cacpi = série de price index de la série de CAC 40

cacri = série de return index de la série de CAC 40

pi1m = taux de placement sans risque (taux PIBOR à un mois)

et pour chaque société du panier repérée par le préfixe code représentant les trois premières lettres du nom des firmes (soit code = BNP, PEU, CAR, FRA, CAP, ELF, CAN)

code = valeur du titre en euro

codepi = série de price index de la société

coderi = série de return index de la société

On note que

- (i) pour les séries correspondant à la valeur des titres en euro la cotation des titres en euro n'a commencé que le 4.01.1999, de ce fait les cours antérieurs en francs ont été convertis en euro sur la base de la parité officielle;
- (ii) les séries de Price Index sont des séries de prix basées à 100 à une date de référence qui peut être variable selon les titres;
- (iii) les séries de Return Index correspondent aux séries de Price Index augmentées des versements de dividendes. Il n'y a pas de prise en compte des dividendes pour l'indice CAC 40. Les séries de Price Index et de Return Index sont confondues.

Ces séries hebdomadaires correspondent à des valeurs fin de période, c'est-à-dire datées du vendredi. Selon cette convention le taux de rendement de l'indice à la période t sera défini comme le taux de croissance de la valeur de l'indice entre t et $t - 1$. Les taux de rendement des titres seront définis comme le taux de croissance du Return Index de chacun des titres. Pour la plupart des séries, la période d'estimation sera 5.01.1991 - 2.07.1999 et portera donc sur 496 observations.

3.3 Mise en œuvre de l'évaluation

Problème 7.1

Estimation par les MCO

- 1-a) Procéder à une visualisation des séries brutes de la base de données. Le taux d'intérêt sans risque sera converti en un taux de dimension hebdomadaire pour être homogène avec les séries de rendement des titres.
- 1-b) Calculer pour les titres du panier les variables d'intérêt du modèle (taux de rendement et primes de risque ou excès de rendement). Construire pour deux sociétés les graphiques des excès de rendement et des excès de rendement du marché.

2. Estimer l'équation (7.5) par les MCO pour les sept sociétés du panier. Commenter les résultats. Récupérer pour chaque modèle la série des résidus estimés. Construire les graphiques de ces séries de résidus et comparer les graphiques entre eux.
3. Quel est l'intérêt du test de significativité globale du Fisher dans le cadre de ce modèle linéaire simple ?

Problème 7.2

Test et correction de l'autocorrélation

1. Tester l'autocorrélation d'ordre 1 des résidus.
2. Corriger lorsque c'est nécessaire l'autocorrélation à l'ordre 1 des résidus en estimant un modèle quasi-différencié selon la procédure itérative de Cochrane-Orcutt. Comparer les résultats avec les estimations précédentes. Pour un des titres concernés par l'autocorrélation, comparer les méthodes de Cochrane-Orcutt et Hildreth-Lu.

Problème 7.3

Tests sur les coefficients α et β

1. Procéder pour les 7 titres aux tests de significativité sur les coefficients α et β . Interpréter les résultats.
2. Procéder au seuil de 5 % au test de Student bilatéral d'égalité à 1 du coefficient β . Conclure sur les caractéristiques des titres.

4. ÉLÉMENTS DE CORRIGÉ

Problème 7.1

Estimation par les MCO

- 1-a) La figure 7.1 représente les évolutions de l'indice CAC 40 et du taux sans risque. La première série affiche une croissance tendancielle sur la période (série non stationnaire, cf. chapitre 14, Économétrie des séries non stationnaires) qui traduit la rentabilité positive des placements en actions sur une période longue. Au contraire le taux d'intérêt sans risque décroît de manière quasi continue. Cette évolution traduit fondamentalement l'orientation de la politique monétaire et la désinflation des années 1990 (cf. chapitre 5, Les politiques monétaires européennes et la règle de Taylor). Il montre bien que contrairement à ce qui est supposé dans le modèle théorique, la série de taux sans risque sera de variance non nulle sur la période d'estimation.

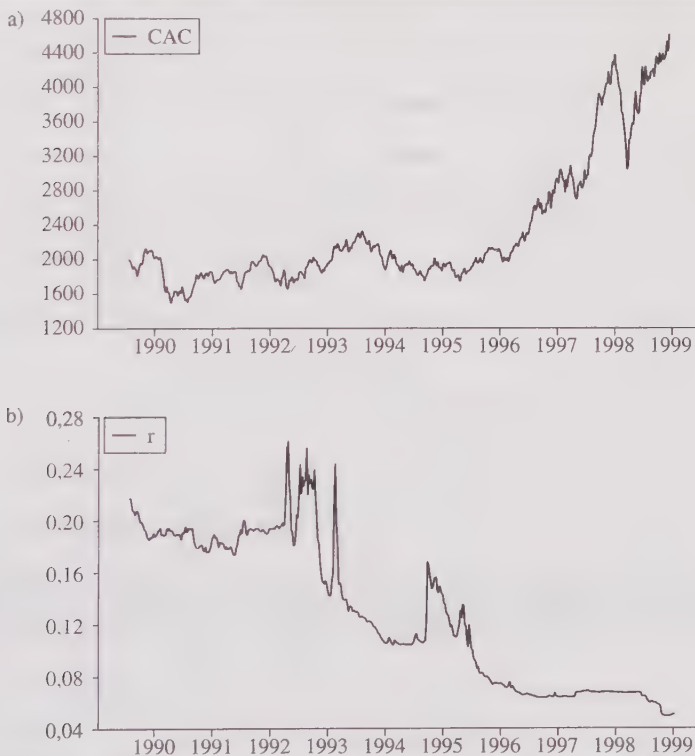


Figure 7.1 Évolution de l'Indice CAC 40 et du taux sans risque. a) indice CAC 40, b) taux sans risque

La figure 7.2 donne l'évolution des séries de Price Index et de Return Index des différentes sociétés. Comme la série de l'indice CAC 40, les séries affichent une croissance tendancielle. La comparaison des profils d'évolution entre sociétés n'a pas ici de réel objet. On pourrait vérifier que les séries de Return Index ont un profil très proche des séries de Price Index. Le décalage entre les deux variables s'explique par les versements ponctuels de dividendes.

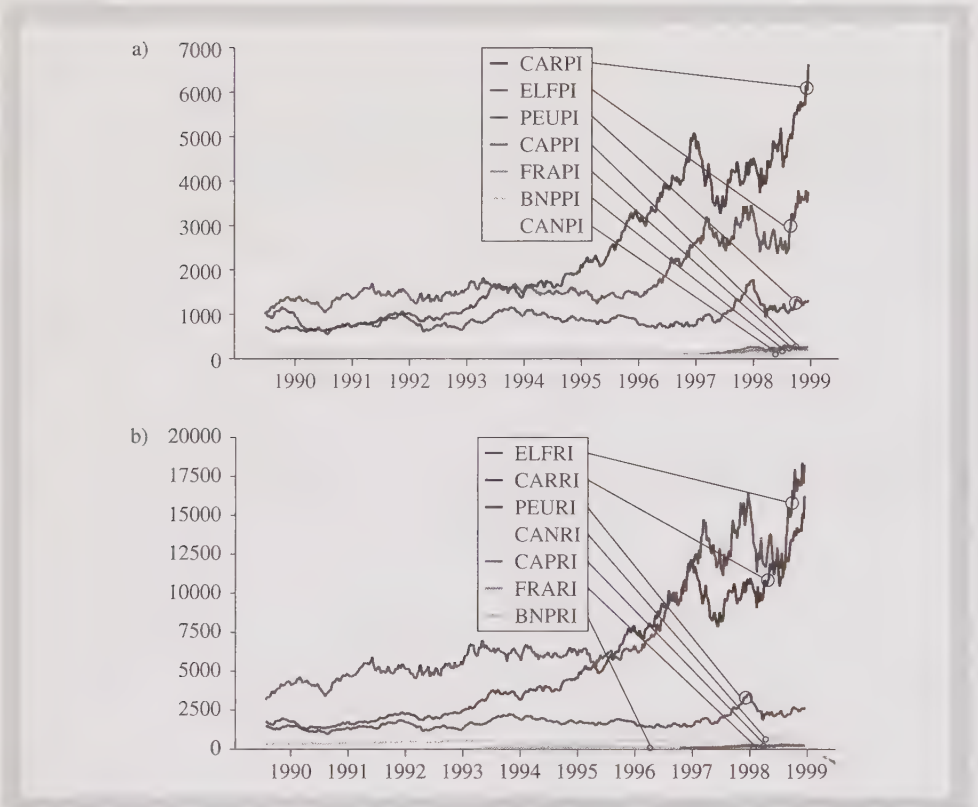


Figure 7.2 Évolution des séries de Price Index et de Return Index. a) Série de Price Index, b) Série de Return Index

1-b) La figure 7.3 représente, pour les sociétés Elf et Canal +, l'évolution des séries d'excès de rendement comparés aux excès de rendement du marché. Ces séries sont très volatiles car construites à partir des taux de croissance hebdomadaires des séries de prix. Elles évoluent autour d'une moyenne légèrement positive (série stationnaire).

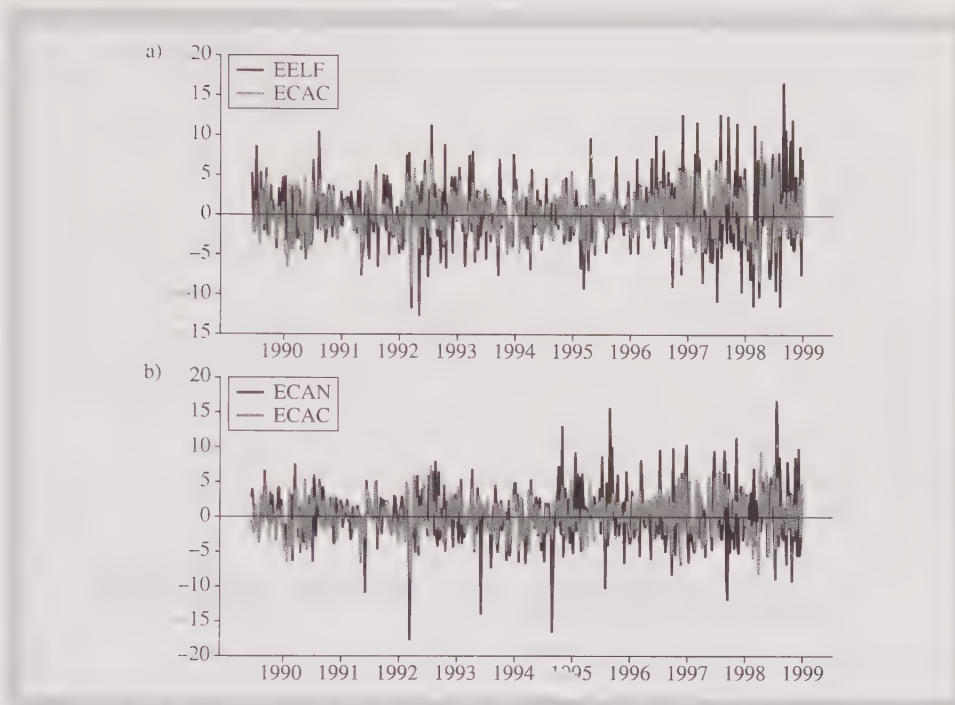


Figure 7.3 Comparaison des séries d'excès de rendement. a) ELF — CAC 40, b) Canal+ — CAC 40

2. Les résultats d'estimation par les MCO de l'équation (7.5) sont résumés dans le tableau 7.1.

Tableau 7.1 Résultats d'estimation par les MCO

Nom	code	β	T-stat	α	T-stat	SEE	$\overline{R^2}$	DW
BNP	BNP	1.28	15.66	-0.03	-0.16	3.83	0.45	2.28
Canal +	CAN	0.55	8.54	0.10	0.60	3.82	0.12	2.09
Cap gémini	CAP	0.78	4.79	0.61	1.11	5.84	0.16	2.45
Carrefour	CAR	0.84	18.38	0.30	2.52	2.70	0.40	1.99
Elf	ELF	1.00	18.48	0.17	1.23	3.22	0.40	2.17
France telecom	FRA	0.71	5.26	0.62	1.40	4.14	0.23	1.94
Peugeot	PEU	1.03	17.49	-0.08	-0.54	3.48	0.38	2.12

Ces estimations laissent apparaître de fortes disparités entre les sociétés cotées quant à la valeur estimée du coefficient β . Il varie de 0.55 pour Canal + à 1.28 pour la BNP. Le coefficient de détermination $\overline{R^2}$ affiche un niveau d'ensemble plutôt bas (inférieur à 0.50) : il varie de 0.12 pour le titre Canal+ à 0.45 pour

la BNP. La variance des rendements hebdomadaires expliquée par le modèle, c'est-à-dire les rendements du marché, est donc faible dans l'ensemble. La variance non expliquée des rendements, c'est-à-dire en fait le risque intrinsèque des titres, est en conséquence élevée (supérieure à 0.50).

Les graphiques sur les résidus estimés donnent un premier aperçu de l'autocorrélation des aléas. Il faut ici opposer les résultats sur BNP et Cap Gemini à ceux obtenus sur les 5 autres sociétés. Même si l'exercice visuel est difficile, on peut déceler pour les deux sociétés citées une autocorrélation négative des résidus où une phase de baisse est dans la plupart des cas immédiatement suivie d'une phase de hausse et réciproquement. Cette propriété est particulièrement visible pour la société Cap Gemini au cours de l'année 1997. Pour les 5 autres sociétés en revanche, les résidus du modèle présentent un profil d'évolution sans autocorrélation apparente. On observe une alternance aléatoire de résidus positifs et négatifs de niveaux assez comparables. Ceci est particulièrement visible sur la société Carrefour.

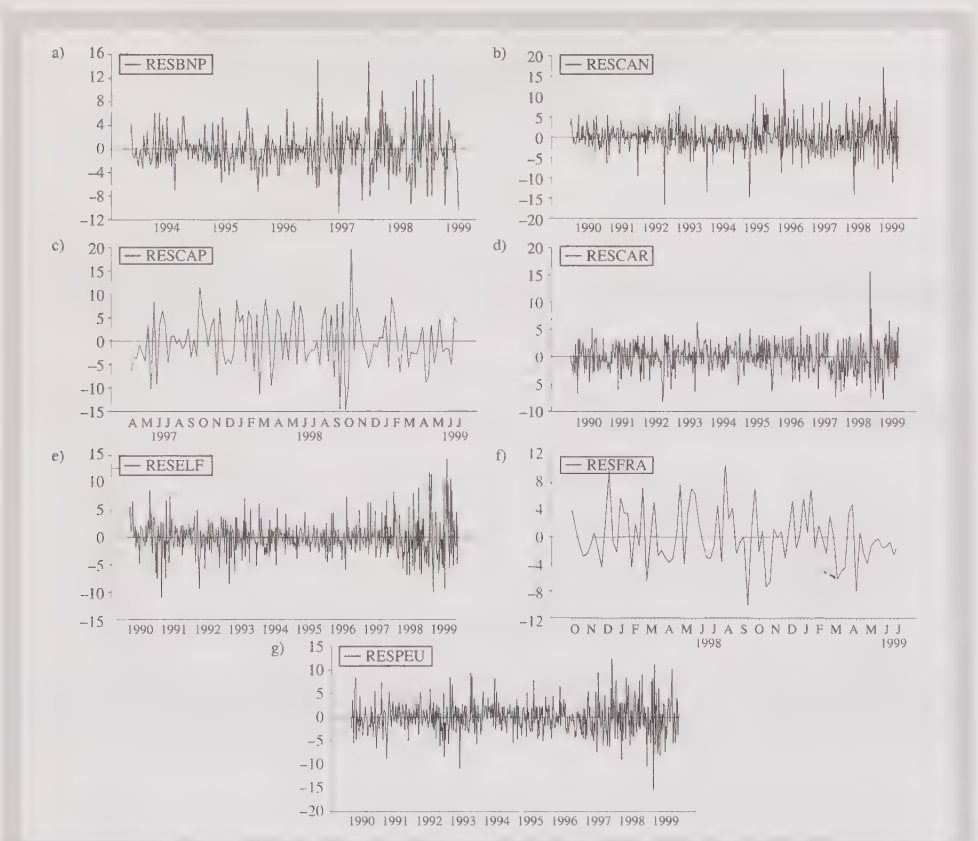


Figure 7.4 Résidus estimés pour les 7 sociétés. a) BNP, b) Canal +, c) Cap Gemini, d) Carrefour, e) ELF, f) France Télécom, g) Peugeot

3. Le test de Fisher de significativité globale du modèle consiste à tester la nullité des coefficients du modèle excepté celui de la constante. Il revient donc ici,

puisque'il s'agit d'un modèle linéaire simple à tester la nullité du coefficient β_i et est donc équivalent au test de Student sur ce paramètre. On vérifie du reste que les risques de rejeter à tort l'hypothèse nulle pour ces deux tests (Statistiques Significance Level) sont rigoureusement identiques pour toutes les sociétés.

Problème 7.2

Test et correction de l'autocorrélation

1. Le test d'autocorrélation de Durbin-Watson teste l'hypothèse nulle $\rho = 0$ où ρ est le coefficient d'autocorrélation des aléas à l'ordre 1. La tabulation des valeurs critiques pour la statistique DW dépend de T , le nombre d'observations et de $K' = K - 1$ avec K , nombre de variables explicatives. On a $K' = 2 - 1 = 1$. En revanche T est variable selon les sociétés, car dépendant de leur date d'introduction dans l'indice CAC 40. On a $T > 200$ pour toutes les sociétés sauf Cap Gemini ($T = 115$) et France Télécom ($T = 88$). La table de Durbin-Watson, pour un seuil $\alpha = 0.05$, fournit trois jeux de valeurs critiques.
 - $T > 200, D_l = 1.76, D_u = 1.78$ et $4 - D_u = 2.22, 4 - D_l = 2.24$
 - $T = 115, D_l = 1.65, D_u = 1.69$ et $4 - D_u = 2.31, 4 - D_l = 2.35$
 - $T = 88, D_l = 1.63, D_u = 1.68$ et $4 - D_u = 2.32, 4 - D_l = 2.37$

On a deux types de conclusions

- On ne peut rejeter H_0 pour Canal +, Carrefour, Elf, France Telecom et Peugeot. On conclut donc à l'absence d'autocorrélation des résidus.
 - On rejette H_0 pour BNP et Cap Gemini. On conclut donc à l'existence d'une autocorrélation des résidus. Les valeurs DW_{obs} supérieures à $4 - D_l$ suggèrent une autocorrélation négative qui pourra être confirmée par l'estimation d'un modèle quasi-différencié.
2. L'autocorrélation des résidus doit donc être corrigée pour BNP et Cap Gemini. Ce nouveau jeu d'estimations doit permettre, selon le théorème de Gauss-Markov, d'obtenir des estimateurs BLUE des paramètres α et β . Le tableau 7.2 présente les résultats complets d'estimation pour la société BNP.

La présentation des résultats pour l'estimation du modèle quasi-différencié, selon la procédure de Cochrane-Orcutt est très proche de celle de l'estimation du modèle de départ par les MCO. Le tableau sur les paramètres du modèle intègre en plus les informations sur le coefficient d'autocorrélation ρ (estimateur, écart type, t-stat et risque de première espèce). Il faut bien remarquer en fait qu'à l'exception des statistiques sur l'autocorrélation des résidus dans le modèle quasi-différencié (Durbin-Watson et statistique Q de Box-Pierce), les informations fournies par le logiciel concerne la relation d'équilibre du ME-DAF dans sa version initiale, c'est-à-dire non quasi-différenciée. Elles sont donc directement comparables à l'estimation précédente par les MCO.

Tableau 7.2 Estimation de la relation d'équilibre — BNP*

Regression with AR1 — Estimation by Cochrane-Orcutt					
Dependent Variable EBNP					
Weekly Data From 1993:11:05 To 1999:07:02					
Usable Observations	296	Degrees of Freedom	293		
Centered R**2	0.469013	R Bar **2	0.465388		
Uncentered R**2	0.470584	T x R**2	139.293		
Mean of Dependent Variable	0.2825988522				
Std Error of Dependent Variable	5.1966018534				
Standard Error of Estimate	3.7996075036				
Sum of Squared Residuals	4230.0460341				
Regression F(2,293)	129.4011				
Significance Level of F	0.00000000				
Log Likelihood	-813.62789				
Durbin-Watson Statistic	1.970841				
Q(36-1)	38.904459				
Significance Level of Q	0.29824879				
	Variable	Coeff	Std Error	T-Stat	Signif
1.	ECAC	1.301586811	0.080615734	16.14557	0.00000000
2.	Constant	-0.046202022	0.191001719	-0.24189	0.80903211
3.	RHO	-0.162967372	0.058389696	-2.79103	0.00559907

*avec correction de l'autocorrélation

Les résultats d'estimation pour les deux sociétés sont synthétisés par le tableau 7.3.

Tableau 7.3 Résultats d'estimation avec correction de l'autocorrélation des résidus par la méthode de Cochrane-Orcutt

Nom	β	Std error	T-stat	α	T-stat	$\overline{R^2}$	DW	ρ	T-stat
BNP	1.30	0.08	16.14	-0.04	-0.24	0.46	1.97	-0.16	-2.79
Cap Gémini	0.83	0.15	5.32	0.62	1.43	0.20	2.01	-0.24	-2.64

On remarque ici que les coefficients estimés ne varient quasiment pas d'une estimation à l'autre (MCO contre modèle quasi-différencié). Ceci illustre le fait qu'en cas d'autocorrélation des résidus, l'estimateur des MCO reste sans biais. Les différences observées, très faibles, s'expliquent par la perte d'un point d'échantillon en début de période lors de l'estimation du modèle quasi-différencié. En revanche, des différences sensibles apparaissent sur les écarts types estimés des estimateurs. Ils sont plus élevés dans le deuxième cas, ce qui signifie que l'estimation initiale par les MCO conduit à surestimer la précision des estimateurs. L'incidence se situe donc au niveau de la fiabilité des tests utilisant ces écarts types.

Le tableau 7.4 présente, pour le cas de Cap Gémini, les résultats d'estimation après correction de l'autocorrélation selon les méthodes de Cochrane-Orcutt et Hildreth-Lu. On constate que les résultats sont identiques.

Tableau 7.4 Comparaison des résultats selon la méthode de correction de l'autocorrélation – Cap Gémini

Regression with AR1 — Estimation by Cochrane-Orcutt					
Dependent Variable ECAP					
Weekly Data From 1997:05:02 To 1999:07:02					
Usable Observations	114		Degrees of Freedom	111	
Centered R**2	0.216880		R Bar **2	0.202770	
Uncentered R**2	0.240695		T x R**2	27.439	
Mean of Dependent Variable	1.1232296265				
Std Error of Dependent Variable	6.3704503785				
Standard Error of Estimate	5.6880307769				
Sum of Squared Residuals	3591.2600471				
Regression F(2,111)	15.3704				
Significance Level of F	0.00000128				
Log Likelihood	-358.41241				
Durbin-Watson Statistic	2.016523				
Q(28-1)	22.761778				
Significance Level of Q	0.69774189				
	Variable	Coeff	Std Error	T-Stat	Signif
1.	ECAC	0.833599012	0.156582419	5.32371	0.00000054
2.	Constant	0.626163573	0.437727489	1.43049	0.15538629
3.	RHO	-0.242782597	0.091740252	-2.64641	0.00931680
Regression with AR1 — Estimation by Hildreth-Lu Search					
Dependent Variable ECAP					
Weekly Data From 1997:05:02 To 1999:07:02					
Usable Observations	114		Degrees of Freedom	111	
Centered R**2	0.216880		R Bar **2	0.202770	
Uncentered R**2	0.240695		T x R**2	27.439	
Mean of Dependent Variable	1.1232296265				
Std Error of Dependent Variable	6.3704503785				
Standard Error of Estimate	5.6880307769				
Sum of Squared Residuals	3591.2600472				
Regression F(2,111)	15.3704				
Significance Level of F	0.00000128				
Log Likelihood	-358.41241				
Durbin-Watson Statistic	2.016532				
Q(28-1)	22.761725				
Significance Level of Q	0.69774473				
	Variable	Coeff	Std Error	T-Stat	Signif
1.	ECAC	0.833597970	0.156582541	5.32370	0.00000054
2.	Constant	0.626164437	0.437729090	1.43048	0.15538722
3.	RHO	-0.242777910	0.091740259	-2.64636	0.00931812

Problème 7.3

Tests sur les coefficients α et β

Les tests de Student sur les coefficients α_i et β_i permettent de détecter d'éventuels rendements anormaux sur les titres et leur caractère défensif ou agressif par rapport au marché.

1. On teste successivement les hypothèses nulles $\alpha_i = 0$ et $\beta_i = 0$. On utilise les résultats du tableau 7.1 (tableau 7.3 pour BNP et Cap Gémini).

Pour la constante, l'hypothèse nulle est acceptée pour toutes les sociétés (conformément au modèle théorique), sauf pour la société Carrefour. On conclut donc à l'existence d'un rendement anormal ou d'une sur-rémunération du titre Carrefour. Par ailleurs, pour le paramètre β_i , l'hypothèse nulle est refusée pour toutes les sociétés.

2. Le test d'égalité à 1 du paramètre β_i s'effectue en calculant la statistique

$$tstatb = \frac{\hat{\beta}_i - 1}{\hat{\sigma}_{\beta_i}}$$

Les valeurs sont reportées dans le tableau 7.5.

Tableau 7.5 Test d'égalité à 1 du coefficient β_i — Statistique *Tstatb*

Cap Gémini	t(111)= -1.062705	with Significance Level 0.29022211
Canal+	t(494)= -6.869886	with Significance Level 0.00000000
BNP	t(293)= 3.741042	with Significance Level 0.00022045
Carrefour	t(494)= -3.431865	with Significance Level 0.00064976
ELF	t(494)= 0.177606	with Significance Level 0.85910558
France Telecom	t(86)= -2.096956	with Significance Level 0.03893403
Peugeot	t(494)= 0.549565	with Significance Level 0.58286602

On rejette H_0 pour quatre titres : BNP (agressif), Canal+ (défensif), Carrefour (défensif) et France Telecom (défensif). Le niveau de significativité du test est inférieur à 5 %. Les autres titres ont donc un comportement semblable à celui du marché.

En tant que modèle d'équilibre général, le MEDAF doit s'appliquer à l'ensemble des titres composant l'indice. La présence de rendements anormaux sur le titre Carrefour est en contradiction avec le modèle théorique et ne permet donc pas en toute rigueur de le valider empiriquement.

L'application de ces tests peut être réalisée en utilisant les données sur l'ensemble des titres composant le CAC 40 (cf. données sur le site web).

8

LA FONCTION D'OFFRE DE VIANDE BOVINE

1. Introduction
2. Spécification de la fonction d'offre
3. Analyse empirique
4. Éléments de corrigé

OBJECTIFS DU CHAPITRE

D'un point de vue méthodologique

- Modèle statique/dynamique
- Méthode des Moindres Carrés Ordinaires
- Méthode des Variables Instrumentales
- Test d'autocorrélation, statistique h de Durbin
- Test d'exogénéité d'Hausman

D'un point de vue économique

- Modélisation de l'offre de viande bovine
- Anticipations adaptatives sur les prix
- Ajustement partiel de l'offre
- Élasticités d'offre à court terme et à long terme

NIVEAU

LICENCE, troisième année

CHAMPS ÉCONOMIQUES TRAITÉS

Microéconomie

Comportement des producteurs

1. INTRODUCTION

La Politique Agricole Commune est principalement régie par le biais d'Organisations Communes de Marchés qui regroupent par produit les instruments des interventions communautaires. Ces cadres réglementaires mis en place pour chaque grande branche de l'agriculture fixent les modalités de gestion des marchés agricoles. Les instruments de régulation utilisés isolément ou de manière combinée peuvent être regroupés en trois catégories : les mesures dites de soutien interne visant directement à accroître les revenus des producteurs (soutien des prix à la production), les mesures de protection externe (droits de douane, prélèvements), et les mesures visant à encourager les exportations (subventions, restitutions). Jusqu'en 1992, la production bovine de l'Union Européenne, excédentaire, bénéficie d'une Organisation Commune de Marché reposant principalement sur le soutien des prix communautaires. Les prix communautaires sont stabilisés à un niveau élevé, grâce au rachat des excédents par les organismes d'intervention. Cette politique de soutien a été modifiée en 1992. La réforme de 1992, effective depuis 1993, devait aboutir à une diminution de l'intervention et à une meilleure maîtrise de la production, par une importance accrue des aides directes au producteur, découplées des prix du marché et une diminution du soutien au marché. Elle avait pour objectif principal de réduire la production pour mieux l'ajuster à la demande, au moyen d'une baisse programmée des prix.

Dans ce contexte, on souhaite rechercher les déterminants des comportements des producteurs français de viande bovine et, en particulier, savoir comment la production a pu être affectée par cette réforme. Pour avoir des éléments d'information sur ces comportements, on estime une fonction d'offre. La spécification de cette fonction est issue d'hypothèses particulières sur le comportement des agents participant au marché étudié.

Il s'agit donc, dans une première étape, de décrire le comportement des agents participant au marché pour spécifier l'équation d'offre. Le modèle théorique est, ensuite, appliqué sur des données françaises pour expliquer l'offre des producteurs de viande.

2. SPÉCIFICATION DE LA FONCTION D'OFFRE

2.1 *Dérivation de la fonction d'offre dans un cadre statique*

Un des objectifs de l'analyse du comportement d'offre des producteurs est de déterminer comment les producteurs vont répondre à des changements de prix à la production et des facteurs de production. Étudier l'offre intègre deux éléments : d'une part, l'analyse de la relation technologique qui existe entre le niveau des facteurs de production et le niveau d'output et, d'autre part, le comportement du producteur dans le choix de ses décisions de production, les prix des produits et des facteurs étant donnés.

Dans le cadre analytique le plus simple, on utilise un modèle statique et on considère une technologie de production monoproduit. Le producteur est supposé

maximiser son profit, le prix à la production et le prix des facteurs de production étant donnés. Il détermine ainsi le niveau de sa production (q_t) et les quantités des facteurs de production (x_{1t}, x_{2t}) sous la contrainte technologique décrite par la fonction de production.

Le programme de comportement s'écrit

$$\max_{x_{1t}, x_{2t}, q_t} (p_t q_t - w_{1t} x_{1t} - w_{2t} x_{2t}) \quad (8.1)$$

sous la contrainte

$$q_t = f(x_{1t}, x_{2t})$$

où p_t représente le prix du bien produit en t , q_t la quantité produite en t , x_{1t}, x_{2t} , les niveaux des facteurs de production variables utilisés, w_{1t}, w_{2t} , leurs prix respectifs et f définit la fonction de production. La résolution de cette forme structurelle permet de déterminer les variables endogènes (offre d'output, demande de facteurs) en fonction des variables exogènes (prix des facteurs). Les conditions du premier ordre liées à la maximisation du profit permettent de dériver la fonction d'offre et le système de demande de facteurs de production.

$$q_t^* = q(p_t, w_{1t}, w_{2t}) \quad (8.2)$$

$$x_{1t}^* = x_1(p_t, w_{1t}, w_{2t}) \quad (8.3)$$

$$x_{2t}^* = x_2(p_t, w_{1t}, w_{2t}) \quad (8.4)$$

La première équation définit la fonction d'offre, les deux équations suivantes correspondent aux fonctions de demande de facteurs de production variables.

Dans ce chapitre, on s'intéresse de manière plus précise à l'estimation de la fonction d'offre. Les principes d'estimation des fonctions de demande de facteurs de production sont étudiés dans le chapitre 10 (Modélisation d'un système de demande de facteurs de production). Pour dériver une relation estimable pour la fonction d'offre, il suffit de spécifier une forme fonctionnelle pour la technologie de production et de résoudre le programme (8.1).

Comme il a déjà été précisé, le modèle analytique est très simple. Une estimation directe de la fonction d'offre déduite de ce cadre d'analyse risque ne pas donner de très bons résultats économétriques notamment en raison du caractère statique du modèle. Il existe des possibilités pour améliorer cette spécification. On présente le modèle d'offre de Nerlove qui permet de dynamiser l'équation d'offre.

2.2 Dynamisation de la fonction d'offre : l'exemple du modèle de Nerlove

Deux extensions peuvent être apportées à la spécification (8.2) de la fonction d'offre : la prise en compte d'anticipations et l'introduction de l'existence de coûts d'ajustement.

La première extension concerne la prise en compte d'anticipations sur le niveau du prix à la production des producteurs. Dans l'équation (8.2), le prix à la production apparaît comme argument de la fonction d'offre. Or, au moment où le producteur effectue ses décisions de production, il ne connaît pas le prix du bien offert p_t . Ce prix est de manière générale déterminé par la confrontation entre l'offre et la demande sur le marché du bien considéré. Aussi, les producteurs doivent former des anticipations sur les prix.

Si l'on pense que le producteur raisonne par rapport à des anticipations de prix, l'équation d'offre (8.2) s'écrit

$$q_t^* = q(p_t^a, w_{1t}, w_{2t}) \quad (8.5)$$

où p_t^a désigne le niveau anticipé des prix pour la période t . Les producteurs essaient de prévoir le niveau qu'ils jugent normal pour le niveau des prix.

Par définition, cette variable n'est pas observable. Des hypothèses doivent être faites pour pouvoir exprimer ces anticipations en fonction des variables observées. Il existe différents schémas pour définir ces anticipations.

La seconde extension concerne le fait que les ajustements de la fonction d'offre à des variations de prix peuvent se faire avec lenteur et décalage en raison de l'existence de coûts d'ajustement. Aussi, le niveau de production observé noté q_t est différent du niveau désiré de production noté q_t^* .

Pour prendre en compte ces deux extensions, on présente les hypothèses retenues dans le modèle de Nerlove¹.

2.2.1 Spécification des anticipations

Pour définir les anticipations de prix, Nerlove retient l'hypothèse d'anticipations adaptatives décrite par la relation suivante²

$$p_t^a = p_{t-1}^a + \gamma \underbrace{(p_{t-1} - p_{t-1}^a)}_{\text{erreur}} \quad (8.6)$$

¹ Nerlove M., (1956), « Estimates of supply of selected agricultural commodities », *Journal of Farm Economics*, 38, 496-509.

La description retenue du modèle de Nerlove est adaptée de la présentation effectuée par Sadoulet E. et De Janvry A., (1995), *Quantitative development policy analysis*, The John Hopkins University Press, 397 pp.

² Il existe d'autres schémas possibles pour définir les anticipations, notamment les anticipations rationnelles. Dans ce cadre, les agents utilisent toute l'information disponible pour former leurs anticipations.

où p_t^a est le prix anticipé pour la période t , p_{t-1} est le prix observé à la période $t-1$, γ est le coefficient des anticipations adaptatives supposé constant et compris entre 0 et 1. À chaque période, le producteur révisé ses anticipations en proportion de l'erreur associée au précédent niveau d'anticipations. L'erreur est égale à la différence entre le niveau observé des prix et le niveau anticipé. Deux cas particuliers apparaissent :

- le premier cas quand $\gamma = 0$, les prix effectifs sont complètement déconnectés des anticipations,
- le second cas quand $\gamma = 1$, les producteurs forment des anticipations naïves car ils anticipent que le prix en t sera identique à celui de la période précédente.

On peut montrer que, dans cette spécification, le prix anticipé est une moyenne mobile à pondérations géométriques décroissantes des prix passés. L'équation précédente peut en effet se réécrire

$$p_t^a = \gamma p_{t-1} + (1 - \gamma) p_{t-1}^a \quad (8.7)$$

En décalant l'équation précédente d'une période, on obtient l'équation suivante

$$p_{t-1}^a = \gamma p_{t-2} + (1 - \gamma) p_{t-2}^a \quad (8.8)$$

Les valeurs successives des anticipations décalées peuvent être ainsi remplacées pour définir l'anticipation du prix à la période t en fonction de variables observées

$$\begin{aligned} p_t^a &= \gamma(p_{t-1} + (1 - \gamma)p_{t-2} + (1 - \gamma)^2 p_{t-3} + \dots) \\ &= \gamma \sum_{i=1}^{\infty} (1 - \gamma)^{i-1} p_{t-i} \end{aligned} \quad (8.9)$$

Les anticipations de prix sont définies suivant un modèle à retards échelonnés au sens de Koyck, c'est-à-dire que les coefficients des variables retardées sont géométriquement décroissants.

Cette équation peut se réécrire

$$\begin{aligned} p_t^a &= \gamma(p_{t-1} + (1 - \gamma)L p_{t-1} + (1 - \gamma)^2 L^2 p_{t-1} + \dots) \\ &= \gamma(1 + (1 - \gamma)L + (1 - \gamma)^2 L^2 + \dots) p_{t-1} \end{aligned} \quad (8.10)$$

où L désigne l'opérateur retard c'est-à-dire $L.p_{t-1} = p_{t-2}$, $L^n.p_{t-1} = p_{t-n-1}$.

Le prix anticipé peut donc s'écrire

$$p_t^a = \frac{\gamma}{1 - (1 - \gamma)L} p_{t-1} \quad (8.11)$$

2.2.2 Existence de coûts d'ajustement : le concept d'ajustement partiel

La seconde hypothèse du modèle de Nerlove est l'existence de coûts d'ajustement qui se traduit par le fait qu'une modification des variables exogènes ne provoque pas un ajustement immédiat et total du niveau de production offert. L'introduction de ces coûts permet de différencier le comportement du producteur à court terme (ajustement partiel du niveau de production) du comportement du producteur à long terme (ajustement total du niveau de production).

Ainsi, à court terme, l'ajustement complet du niveau de produit offert n'est pas possible en raison par exemple, de contraintes liées à l'allocation des facteurs de production. L'ajustement réel est simplement une fraction de l'ajustement réalisé. De manière analytique, ce concept d'ajustement partiel s'écrit

$$q_t - q_{t-1} = \delta(q_t^* - q_{t-1}) \quad (8.12)$$

où q_t correspond au niveau observé de produit offert en t (ou le niveau de court terme), q_{t-1} le niveau observé pour la période précédente et q_t^* représente le niveau de produit désiré (ou le niveau de long terme s'il n'y avait pas de coûts d'ajustement). Le coefficient δ définit le coefficient d'ajustement partiel compris entre 0 et 1. Le producteur souhaite produire le niveau q_t^* . Pour atteindre ce niveau, il doit faire varier sa production de $(q_t^* - q_{t-1})$. En raison de l'existence de coûts d'ajustement, seul un pourcentage de cet ajustement sera possible. La variation réelle de production $q_t - q_{t-1}$ est donc inférieure à l'ajustement souhaité.

2.2.3 Le modèle général de Nerlove

Le modèle structurel d'offre de Nerlove comprend donc l'équation décrivant les anticipations (équation 8.11) ainsi que l'équation expliquant l'ajustement partiel (équation 8.12).

Pour compléter le modèle structurel, Nerlove rajoute une troisième relation qui spécifie la quantité désirée comme une relation linéaire des prix anticipés et de variables exogènes z_t ³.

$$q_t^* = \alpha_1 + \alpha_2 p_t^a + \alpha_3 z_t \quad (8.13)$$

Cette équation correspond au comportement à long terme du producteur lorsqu'il n'y a plus de coûts d'ajustement. Le coefficient du prix anticipé définit ainsi la réponse à long terme.

Le modèle général d'offre de Nerlove s'écrit donc

$$q_t^* = \alpha_1 + \alpha_2 p_t^a + \alpha_3 z_t + u_t \quad (8.14)$$

³ L'introduction de cette variable exogène est nécessaire pour garantir l'identification du modèle. Cette variable correspond à un autre facteur exogène jouant sur l'offre (par exemple, le prix des facteurs de production, les conditions météorologiques, ...)

$$p_t^a = \frac{\gamma}{1 - (1 - \gamma)L} p_{t-1} + \varepsilon_t \quad (8.15)$$

$$q_t - q_{t-1} = \delta(q_t^* - q_{t-1}) + v_t \quad (8.16)$$

où u_t , ε_t et v_t sont les termes d'erreurs associés à chacune des trois équations. Pour obtenir une équation estimable de la fonction d'offre, il faut éliminer de ces équations les variables non observables dans les relations précédentes, c'est-à-dire le niveau de production en l'absence de coûts d'ajustement q_t^* et le niveau des prix anticipés p_t^a .

L'équation (8.15) est introduite dans l'équation (8.14). On obtient

$$q_t^* = \alpha_1 + \alpha_2 \left(\frac{\gamma}{1 - (1 - \gamma)L} p_{t-1} + \varepsilon_t \right) + \alpha_3 z_t + u_t \quad (8.17)$$

Cette nouvelle expression définissant la quantité désirée est reportée dans l'équation (8.16). On obtient

$$q_t - q_{t-1} = \delta \left(\alpha_1 + \alpha_2 \left(\frac{\gamma}{1 - (1 - \gamma)L} p_{t-1} + \varepsilon_t \right) + \alpha_3 z_t + u_t - q_{t-1} \right) + v_t \quad (8.18)$$

Cette équation est multipliée par $(1 - (1 - \gamma)L)$. On obtient

$$\begin{aligned} (q_t - q_{t-1})(1 - L + \gamma L) &= \delta \alpha_1 (1 - L + \gamma L) + \delta \alpha_2 \gamma p_{t-1} \\ &\quad + \delta \alpha_2 \varepsilon_t (1 - L + \gamma L) \\ &\quad + \delta \alpha_3 z_t (1 - L + \gamma L) \\ &\quad + \delta u_t (1 - L + \gamma L) \\ &\quad - \delta q_{t-1} (1 - L + \gamma L) \\ &\quad + v_t (1 - L + \gamma L) \end{aligned} \quad (8.19)$$

En développant les termes on obtient

$$\begin{aligned} q_t - q_{t-1} - q_{t-1} + q_{t-2} + \gamma q_{t-1} - \gamma q_{t-2} &= \delta \alpha_1 - \delta \alpha_1 + \delta \alpha_1 \gamma \\ &\quad + \delta \alpha_2 \gamma p_{t-1} \\ &\quad + \delta \alpha_2 \varepsilon_t - \delta \alpha_2 \varepsilon_{t-1} + \gamma \delta \alpha_2 \varepsilon_{t-1} \\ &\quad + \delta \alpha_3 z_t - \delta \alpha_3 z_{t-1} + \gamma \delta \alpha_3 z_{t-1} \\ &\quad + \delta u_t - \delta u_{t-1} + \gamma \delta u_{t-1} \\ &\quad - \delta (q_{t-1} - q_{t-2} + \gamma q_{t-2}) \\ &\quad + v_t - v_{t-1} + \gamma v_{t-1} \end{aligned} \quad (8.20)$$

En rassemblant les différentes parenthèses, l'équation peut se réécrire

$$\begin{aligned}
 q_t - (2 - \gamma)q_{t-1} + (1 - \gamma)q_{t-2} &= \delta\alpha_1 - \delta\alpha_1 + \delta\alpha_1\gamma + \delta\alpha_2\gamma p_{t-1} \\
 &\quad - \delta q_{t-1} + \delta(1 - \gamma)q_{t-2} \\
 &\quad + \delta\alpha_3 z_t - \alpha_3\delta(1 - \gamma)z_{t-1} \\
 &\quad + \delta\alpha_2\varepsilon_t - \alpha_2\delta(1 - \gamma)\varepsilon_{t-1} \\
 &\quad + \delta u_t - \delta(1 - \gamma)u_{t-1} \\
 &\quad + v_t + v_{t-1}(\gamma - 1)
 \end{aligned} \tag{8.21}$$

Cette équation est réaménagée pour obtenir une relation définissant la quantité observée q_t en fonction des variables endogènes décalées, des prix décalés et des autres variables. En regroupant tous les termes et de manière synthétique, on obtient la relation suivante

$$q_t = \pi_1 + \pi_2 p_{t-1} + \pi_3 q_{t-1} + \pi_4 q_{t-2} + \pi_5 z_t + \pi_6 z_{t-1} + \mu_t \tag{8.22}$$

Les coefficients de la forme réduite notés π_i ($i = 1, \dots, 6$) s'expriment en fonction des coefficients de la forme structurelle de la façon suivante

$$\pi_1 = \alpha_1\delta\gamma ; \pi_2 = \alpha_2\delta\gamma ; \pi_3 = (1 - \delta) + (1 - \gamma) \tag{8.23}$$

$$\pi_4 = -(1 - \delta)(1 - \gamma) ; \pi_5 = \alpha_3\delta ; \pi_6 = -\alpha_3\delta(1 - \gamma) \tag{8.24}$$

Le terme d'erreur de l'équation se définit par l'égalité suivante

$$\begin{aligned}
 \mu_t &= \alpha_2\delta\varepsilon_t - \alpha_2\delta(1 - \gamma)\varepsilon_{t-1} \\
 &\quad + \delta u_t - \delta(1 - \gamma)u_{t-1} \\
 &\quad + v_t - (1 - \gamma)v_{t-1}
 \end{aligned} \tag{8.25}$$

On a six paramètres de la forme réduite et cinq paramètres de la forme structurelle. Pour obtenir une solution unique, une contrainte doit être imposée sur les paramètres de la forme réduite. On peut montrer que cette contrainte est donnée par la relation suivante

$$\pi_6^2 - \pi_4\pi_5^2 + \pi_3\pi_5\pi_6 = 0 \tag{8.26}$$

Si les différentes variables sont exprimées en logarithme, on note que le coefficient attaché à la variable prix représente l'élasticité de court terme.

Pour compléter l'interprétation des résultats, il est intéressant de calculer les paramètres de la forme structurelle qui s'expriment en fonction des coefficients π_i notamment le coefficient des anticipations $(1 - \gamma) = -\pi_6/\pi_5$, le coefficient d'ajustement

$(1 - \delta) = \pi_3 + \pi_6/\pi_5$ et le coefficient donnant l'effet marginal de long terme des prix $\alpha_2 = \pi_2/\delta\gamma$.

Le modèle de Nerlove est retenu pour spécifier la fonction d'offre de viande bovine en France. Toutefois, les caractéristiques du produit étudié conduisent à une adaptation de la spécification générale de la fonction d'offre dérivée du modèle de Nerlove.

2.3 *Prise en compte des caractéristiques de la production de bœuf*

Les spécificités du produit étudié, «la viande bovine», nécessitent d'adopter des hypothèses particulières sur la fonction d'offre et compliquent la modélisation du comportement des producteurs. Ces caractéristiques correspondent notamment à la politique d'intervention existant pour ce produit et à l'existence d'un délai de production.

2.3.1 Existence d'une politique d'intervention

Dans le contexte de soutien des prix communautaires en vigueur sur le marché de la viande bovine, on va supposer que les anticipations de prix des producteurs sont réalisées de manière naïve c'est-à-dire, $p_t^a = p_{t-1}$. Autrement dit, le coefficient défini dans l'équation (8.6) des anticipations adaptatives est égal à 1 ($\gamma = 1$). Dans ce cas, la forme réduite (équation 8.22) définissant l'offre en fonction des variables observées peut se simplifier car certains coefficients de la forme réduite disparaissent.

$$\pi_1 = \alpha_1\delta ; \pi_2 = \alpha_2\delta ; \pi_3 = (1 - \delta) ; \pi_4 = 0 ; \pi_5 = +\alpha_3\delta ; \pi_6 = 0 \quad (8.27)$$

L'équation d'offre devient

$$q_t = \pi_1 + \pi_2 p_{t-1} + \pi_3 q_{t-1} + \pi_5 z_t + \mu_t \quad (8.28)$$

Le délai d'ajustement de la production est mesuré par le coefficient associé à la variable endogène décalée. L'ajustement est d'autant plus important que ce coefficient est faible (c'est-à-dire le coefficient δ est élevé).

2.3.2 Caractère dynamique de la production

Une spécificité importante de la production de viande bovine réside dans son caractère dynamique. Ainsi de manière schématique, la viande provient de jeunes bovins abattus à 15 ou 18 mois, de bœufs abattus à 3 ou 4 ans et enfin de vaches réformées vers 7 ou 8 ans (vaches allaitantes et vaches laitières). Le renouvellement du cheptel nécessite donc un délai. L'effectif du cheptel bovin présent sur l'exploitation agricole une année donnée, ainsi que ses caractéristiques démographiques, conditionnent le niveau de production bovine des mois suivants, voire des années suivantes : la

contrainte biologique crée un délai de production. Par conséquent, le niveau d'offre en t dépend des prix passés. Comme la viande provient de plusieurs types d'animaux différenciés suivant l'âge, différents décalages de prix doivent être introduits.

3. ANALYSE EMPIRIQUE

Description des données

Les données sont des données annuelles françaises sur la période 1953-1998. Elles ont été fournies par le SCEES (Service Central des Enquêtes et Études Statistiques). En France, il n'existe pas de mesure directe de la production bovine dans les élevages. Pour mesurer le niveau de la production bovine française notée q dans l'exercice, on utilise la notion de production indigène contrôlée. La production indigène contrôlée correspond aux abattages contrôlés par les services vétérinaires dans les abattoirs, auxquels sont retirés les animaux vivants importés pour être abattus en France et ajoutés les animaux vivants exportés produits en France. La production indigène contrôlée est exprimée en milliers de tonnes équivalent carcasse (tec)⁴. Par ailleurs, on dispose de deux séries d'indices de prix : l'indice du prix du bœuf à la production (p) exprimé en base 1970 et l'indice du prix des aliments du bétail (w) exprimé en base 1980.

Problème 8.1

Construction des variables et analyse descriptive

1. Transformer les variables de prix de façon à retenir l'année 1990 comme année de référence.
2. Analyser graphiquement les différentes variables.

Problème 8.2

Estimation de la fonction d'offre en utilisant un modèle de Nerlove

Dans un premier temps, on souhaite estimer une fonction d'offre dérivée d'un modèle de Nerlove. La spécification est basée sur des anticipations naïves pour le prix du bœuf à la production et sur un ajustement partiel de la quantité désirée à la quantité observée. De plus, on suppose que l'offre dépend du prix passé des aliments du bétail⁵.

1. Écrire dans le cadre d'analyse retenu la spécification de la fonction d'offre en rappelant l'interprétation des paramètres.

⁴ Les quantités de viande échangées sous formes d'animaux vivants, de carcasses, de pièces de découpe, de conserves sont exprimées dans leur équivalent en poids de carcasse par des coefficients prenant en compte le rendement de ces diverses transformations. Par exemple, pour la viande bovine, 100 kg de poids vif équivalent à 55 kg carcasse, 100 kg de viande sans os équivalent à 140 kg carcasse.

⁵ On suppose que les producteurs réalisent des anticipations naïves sur le prix d'un des principaux facteurs de production, les aliments du bétail.

2. Estimer par les MCO l'équation d'offre. Rappeler pourquoi l'hypothèse d'autocorrélation d'ordre 1 des résidus est essentielle dans le choix de la méthode d'estimation. Tester cette hypothèse.
3. Tester l'hypothèse d'exogénéité de la quantité décalée. En fonction du résultat du test, choisir la méthode d'estimation adaptée.

Problème 8.3

Interprétation des résultats et utilisation des variables indicatrices

1. Interprétation des résultats : calculer les élasticités prix de court et de long terme, calculer la valeur du coefficient d'ajustement partiel et interpréter cette valeur.
2. Sur la période d'estimation, il y a eu des chocs accidentels : le boycott du «veau aux hormones» par les consommateurs durant l'année 1973, la crise de la vache folle en 1996. Montrer que la première année de boycott de la viande de veau a eu un impact sur le niveau des abattages en 1974. La crise de la vache folle en 1996 a-t-elle eu un impact significatif sur le niveau de l'offre ?
3. On veut examiner l'effet de la réforme de la Politique Agricole Commune de 1992, effective depuis 1993, sur les décisions de production. Introduire l'effet de cette réforme en supposant dans un premier temps, que cette modification n'a pas modifié la réponse au prix des producteurs puis en supposant qu'elle a modifié la réponse au prix.

Problème 8.4

Une extension : prise en compte des caractéristiques de la production de bœuf

1. Pour tenir compte du fait que la production de bœuf nécessite plusieurs années, on estime finalement la relation suivante

$$q_t = b_0 + b_1 p_{t-1} + b_2 p_{t-2} + b_3 p_{t-3} + b_4 q_{t-1} + b_5 w_{t-1} + b_6 D_{74} + \varepsilon_t$$

où D_{74} définit une variable indicatrice, elle est égale à 1 pour l'année 1974, 0 sinon.

Commenter l'effet des variables prix.

4. ÉLÉMENTS DE CORRIGÉ

Problème 8.1

Construction des variables et analyse descriptive

1. Pour obtenir des variables d'indice en base 1990, on divise chaque série d'indice de prix par sa valeur observée en 1990 et on la multiplie par 100. Les nouvelles séries d'indice de prix sont ainsi toutes égales à 100 en 1990.
2. On représente le niveau de production et les deux séries de prix.

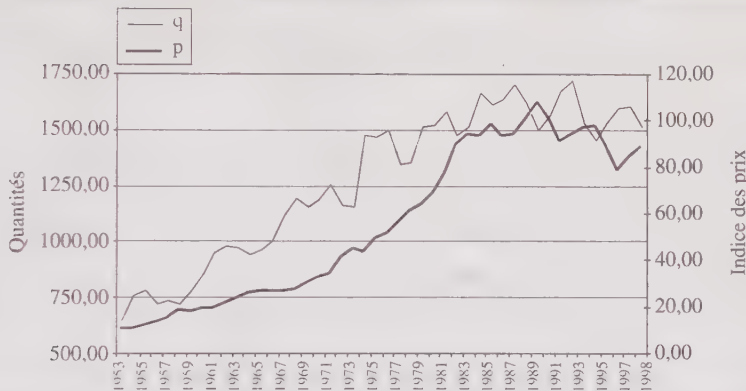


Figure 8.1 Niveau de production (en milliers de TEC) et prix à la production (indice base 1990)

Sur la figure 8.1 sont représentés le niveau de production et le niveau des prix à la production. Cette figure montre une évolution cyclique de la production de bœuf (alternance de phases de pics et de phases de creux). On remarque que certains chocs sont venus perturber ce cycle et ont pu amplifier les écarts. Ainsi, on peut penser que la réforme de la politique agricole commune en 1992 et la crise de la vache folle en 1996 ont eu des effets sur l'évolution de cette production. On observe ainsi une diminution de la production en début de réforme traduisant une modification du comportement de production des éleveurs.

La seconde figure représente l'évolution du prix à la production et du prix d'un des principaux intrants (les aliments du bétail). Les deux séries d'indices de prix ont augmenté de manière assez continue et de façon plutôt linéaire jusqu'au début des années 80. Depuis 1989, on constate une rupture pour l'évolution des deux variables prix, elles suivent une tendance décroissante. Cette évolution est vraisemblablement à mettre en parallèle avec la diminution de la consommation de bœuf observée depuis 1980 et avec les effets de la PAC (baisse programmée des prix en 1994, 1995, 1996).

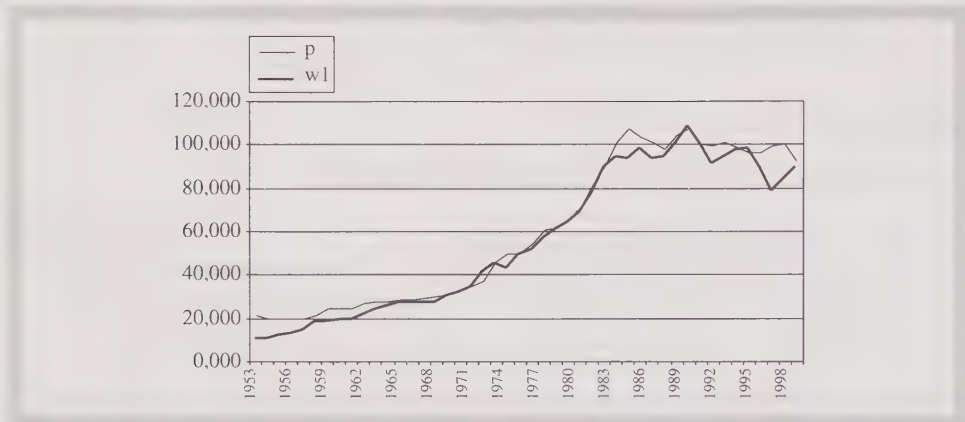


Figure 8.2 Prix à la production (francs 1990) et prix des aliments du bétail (francs 1990)

Problème 8.2

Estimation de la fonction d'offre en utilisant un modèle de Nerlove

1. Le modèle structurel de Nerlove sous l'hypothèse que les producteurs forment des anticipations naïves sur le niveau des prix en début de période s'écrit

$$q_t^* = \alpha_1 + \alpha_2 p_t^a + \alpha_3 w_{t-1} + u_t \quad (8.29)$$

$$p_t^a = p_{t-1} \quad (8.30)$$

$$q_t - q_{t-1} = \delta(q_t^* - q_{t-1}) + v_t \quad (8.31)$$

où w_{t-1} représente le prix des aliments du bétail en $t - 1$.

- Quand on élimine les variables non observées (q_t^* et p_t^a) on obtient l'équation

$$q_t - q_{t-1} = \delta[(\alpha_1 + \alpha_2 p_{t-1} + \alpha_3 w_{t-1} + u_t) - q_{t-1}] + v_t \quad (8.32)$$

Cette équation peut se réécrire

$$q_t = \delta\alpha_1 + \delta\alpha_2 p_{t-1} + (1 - \delta)q_{t-1} + \delta\alpha_3 w_{t-1} + \delta u_t + v_t \quad (8.33)$$

On va donc estimer la relation suivante

$$q_t = \pi_1 + \pi_2 p_{t-1} + \pi_3 q_{t-1} + \pi_4 w_{t-1} + \mu_t \quad (8.34)$$

avec $\pi_1 = \delta\alpha_1$; $\pi_2 = \delta\alpha_2$; $\pi_3 = (1 - \delta)$; $\pi_4 = \delta\alpha_3$; $w_t = \delta u_t + v_t$

Le coefficient de la variable prix π_2 mesure l'effet marginal à court terme sur la production d'une variation d'une unité des prix (ce coefficient est utilisé pour calculer l'élasticité de court terme). L'impact de long terme de la variation des prix est donné par le coefficient α_2 c'est-à-dire par $(\pi_2/1 - \pi_3)$.

Le coefficient de la variable endogène décalée (q_{t-1}) permet de calculer le coefficient d'ajustement partiel ($\delta = 1 - \pi_3$).

Enfin, le coefficient π_4 mesure l'impact marginal d'une variation du prix du facteur de production sur l'offre.

- Les résultats d'estimation du modèle par les MCO sont présentés dans le tableau 8.1.

Tableau 8.1 Estimation de l'équation d'offre dans le cadre d'un modèle à la Nerlove

Dependent Variable Q — Estimation by Least Squares					
Annual Data From 1954 :01 To 1998 :01					
Usable Observations	45	Degrees of Freedom	41		
Centered R**2	0.935333	R Bar **2	0.930601		
Uncentered R**2	0.996299	T x R**2	44.833		
Mean of Dependent Variable	1290.1282222				
Std Error of Dependent Variable	321.4472327				
Standard Error of Estimate	84.6809054				
Sum of Squared Residuals	294005.08551				
Regression F(3,41)	197.6724				
Significance Level of F	0.00000000				
Durbin-Watson Statistic	1.804295				
Q(11-0)	14.180192				
Significance Level of Q	0.22318264				

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	223.8038469	77.1818075	2.89970	0.00597740
2.	P{1}	5.0270062	2.8832498	1.74352	0.08873783
3.	Q{1}	0.7792046	0.0946070	8.23623	0.00000000
4.	W{1}	-3.3492132	2.7303938	-1.22664	0.22695541

Avant de commenter ces résultats, il est nécessaire de vérifier la validité de l'hypothèse de non corrélation des aléas. Ce modèle est autorégressif, ou encore dynamique, c'est-à-dire la variable endogène décalée figure parmi les variables explicatives. Si les perturbations de l'équation sont corrélées entre elles, la variable q_{t-1} n'est plus exogène. Les estimateurs fournis par les MCO sont alors biaisés et non convergents.

Puisqu'une variable dépendante retardée apparaît parmi les variables explicatives, on ne peut pas utiliser le test de Durbin-Watson pour tester l'autocorrélation des aléas. On utilise la statistique du h de Durbin

$$h = (1 - d/2) \sqrt{\frac{T}{1 - T [\widehat{\text{var}}(\widehat{\Pi}_3)]}}$$

où T , d et $\widehat{\text{var}}(\widehat{\pi}_3)$ correspondent respectivement au nombre d'observations, à la statistique de Durbin-Watson et à la variance estimée de l'estimateur par les MCO du paramètre de la variable endogène décalée. Sous l'hypothèse nulle d'absence d'autocorrélation des perturbations, cette statistique suit une loi normale centrée réduite.

**Tableau 8.2 Test hypothèse d'autocorrélation d'ordre 1 des résidus :
Statistique h de Durbin Watson**

statistique h de Durbin	0.84939
---------------------------	---------

La statistique du h est égale à 0.849, au seuil de 5 %. Cette valeur est inférieure à 1.64, on ne rejette donc pas l'hypothèse nulle. Les aléas ne sont pas autocorrélés.

3. Pour tester l'hypothèse d'exogénéité de la quantité décalée, on peut aussi utiliser le test d'Hausman. Dans une première étape (tableau 8.3), on régresse la variable endogène décalée q_{t-1} sur les instruments. Ces instruments correspondent aux variables exogènes du modèle et à leurs valeurs décalées. De cette régression, on calcule le résidu estimé. Dans une seconde étape, on estime une nouvelle équation comprenant la variable endogène décalée, le prix et le résidu estimé (tableau 8.4). Le test d'exogénéité de la variable endogène décalée revient à tester la significativité du coefficient du résidu estimé. Si ce coefficient n'est pas significatif, on peut considérer que la variable endogène décalée est exogène.

L'équation suivante représente l'équation d'instrumentation, on régresse q_{t-1} sur p_{t-1} , w_{t-1} et w_{t-2} et le résidu ($RESIDS$) est calculé.

**Tableau 8.3 Test d'Hausman - 1^{re} étape :
test exogénéité de la quantité décalée sur l'équation d'offre**

Dependent Variable Q1 — Estimation by Least Squares			
Annual Data From 1955 :01 To 1998 :01			
Usable Observations	44	Degrees of Freedom	40
Centered R**2	0.838773	R Bar **2	0.826681
Uncentered R**2	0.990608	T x R**2	43.587
Mean of Dependent Variable	1284.9680455		
Std Error of Dependent Variable	323.2726522		
Standard Error of Estimate	134.5836995		
Sum of Squared Residuals	724510.88670		
Regression F(3,40)	69.3657		
Significance Level of F	0.00000000		
Durbin-Watson Statistic	0.733095		
Q(11-0)	39.558416		
Significance Level of Q	0.00004256		

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	767.8573319	44.0314658	17.43883	0.00000000
2.	P{1}	6.3853811	4.5470360	1.40430	0.16794848
3.	W{1}	8.3868599	6.9176881	1.21238	0.23247974
4.	W{2}	-5.8904868	5.4961206	-1.07175	0.29025479

Tableau 8.4 Test d'Hausman - 2^e étape

Dependent Variable Q — Estimation by Least Squares			
Annual Data From 1955 :01 To 1998 :01			
Usable Observations	44	Degrees of Freedom	39
Centered R**2	0.932746	R Bar **2	0.925848
Uncentered R**2	0.996369	T x R**2	43.840
Mean of Dependent Variable	1302.2061364		
Std Error of Dependent Variable	314.6652607		
Standard Error of Estimate	85.6862451		
Sum of Squared Residuals	286343.17123		
Regression F(4,39)	135.2219		
Significance Level of F	0.00000000		
Durbin-Watson Statistic	1.852089		
Q(11-0)	17.404395		
Significance Level of Q	0.09646980		

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	-156.1755721	461.2995757	-0.33856	0.73676051
2.	P{1}	2.4072714	4.6973909	0.51247	0.61121253
3.	Q{1}	1.2662359	0.5940510	2.13153	0.03940235
4.	W{1}	-5.0232464	3.2466121	-1.54723	0.12988581
5.	RESIDS	-0.4840039	0.6025201	-0.80330	0.42667076

Le coefficient du résidu estimé n'est pas significatif, on peut donc considérer que la quantité décalée est exogène. Une procédure alternative à la précédente consiste à calculer la statistique de Wald (cf. chapitre 6, Les hypothèses du modèle de régression linéaire multiple). On obtient une valeur de la statistique de Wald égale à 0.49 avec un niveau de significativité de 48 %.

Ce résultat est cohérent avec le résultat du test de l'autocorrélation. On peut donc retenir la méthode des MCO pour estimer les paramètres du modèle.

Problème 8.3

Interprétation des résultats et utilisation des variables indicatrices

1. Les signes des coefficients sont cohérents avec la théorie microéconomique : impact positif du prix du produit sur le niveau de production, impact négatif du prix des intrants. Le coefficient du prix des aliments n'est pas significativement différent de zéro à 5 %. Pour autant, on ne l'omet pas de la spécification car cette variable est un argument de base d'une fonction d'offre.

Les coefficients estimés mesurant l'impact d'une variation d'une unité de chaque variable considérée, la hiérarchisation des effets se heurte à des problèmes d'unités. Pour comparer l'effet des variables explicatives, il est nécessaire de calculer des élasticités.

Le coefficient de la variable prix donne directement l'effet marginal d'une variation des prix à court terme (impact d'une variation d'une unité du prix en $t - 1$ sur le niveau de production en t).

Pour calculer l'élasticité de court terme, on retient l'année 1990 comme point de référence.

On multiplie le coefficient (π_2) par le rapport entre le niveau des prix et le niveau de la production en 1990.

$$\varepsilon_{ct}q/p = \left(\frac{\partial q_t}{\partial p_{t-1}} \frac{p_{90}}{q_{90}} \right) = 5.02 \times \frac{100}{1565} = 0.32 \quad (8.35)$$

Sachant que le niveau de production est égal à 1565 (valeur de 1990), si le prix à la production augmente de 1 % en $t - 1$, la production augmente de 0.32 % en t .

Pour calculer l'effet marginal de long terme, il faut calculer le rapport ($\pi_2/1 - \pi_3$). Le calcul suivant montre que l'élasticité de long terme est égale à ce coefficient multiplié par le rapport entre le niveau des prix et le niveau de la production en 1990.

$$\begin{aligned} \varepsilon_{Lt}q/p &= \left(\frac{\partial q_t}{\partial p_{t-1}} + \frac{\partial q_t}{\partial p_{t-1}} \frac{\partial q_{t-1}}{\partial p_{t-2}} + \frac{\partial q_{t-1}}{\partial p_{t-2}} \frac{\partial q_{t-2}}{\partial p_{t-3}} + \dots \right) \frac{p_{90}}{q_{90}} \\ &= (\pi_2 + \pi_2\pi_3 + \pi_2\pi_3^2 + \dots) \frac{p_{90}}{q_{90}} \\ &= \left(\frac{\pi_2}{1 - \pi_3} \right) \frac{p_{90}}{q_{90}} = \frac{5.02}{1 - 0.77} \times \frac{100}{1565} = 1.454 \end{aligned}$$

Le même calcul peut être fait pour interpréter l'effet du prix des aliments du bétail.

Tableau 8.5 Calcul des élasticités avec 1990 comme année de référence

Élasticité court terme par rapport au prix du boeuf =	0.32119
Élasticité long terme par rapport au prix du boeuf =	1.45472
Élasticité court terme par rapport au prix des aliments du bétail =	-0.21399
Élasticité long terme par rapport au prix des aliments du bétail =	-0.96920
Calcul du coefficient d'ajustement partiel	
coef. d'ajustement partiel =	0.22080

Le coefficient d'ajustement partiel δ est égal à $1 - \pi_3$. Il est égal à $1 - 0.779 = 0.22$, ce qui signifie que seulement 22 % de la différence entre la quantité désirée de production et la quantité observée est éliminée en 1 an.

2. Influence du boycott du «veau aux hormones» en 1973 et de la crise de la vache folle en 1996.

Pour mesurer si ces chocs exogènes ont eu un impact significatif sur le niveau de l'offre, on construit deux variables indicatrices notées respectivement D74 et D96 qui vont permettre de marquer ces deux années : D74 est égale à 1 pour l'année 1974, 0 sinon et D96 est égale à 1 en 1996, 0 sinon. Ces deux variables sont introduites de manière linéaire dans la spécification.

Tableau 8.6 Impact du boycott de la viande de veau après 1973

Dependent Variable Q — Estimation by Least Squares					
Annual Data From 1954 :01 To 1998 :01					
Usable Observations	45	Degrees of Freedom	40		
Centered R**2	0.952919	R Bar **2	0.948211		
Uncentered R**2	0.997306	T x R**2	44.879		
Mean of Dependent Variable	1290.1282222				
Std Error of Dependent Variable	321.4472327				
Standard Error of Estimate	73.1523335				
Sum of Squared Residuals	214050.55606				
Regression F(4,40)	202.4006				
Significance Level of F	0.00000000				
Durbin-Watson Statistic	1.743207				
Q(11-0)	20.608445				
Significance Level of Q	0.03766214				

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	205.0835061	66.8498363	3.06782	0.00385887
2.	P{1}	4.0177652	2.5043681	1.60430	0.11651504
3.	Q{1}	0.7856126	0.0817439	9.61066	0.00000000
4.	W{1}	-2.3493143	2.3728169	-0.99010	0.32807937
5.	D74	288.0497151	74.5202149	3.86539	0.00039793

La variable D74 a un impact significatif et positif sur le niveau de production.

Tableau 8.7 Impact de la crise de la vache folle en 1996

Dependent Variable Q — Estimation by Least Squares			
Annual Data From 1954 :01 To 1998 :01			
Usable Observations	45	Degrees of Freedom	40
Centered R**2	0.936001	R Bar **2	0.929601
Uncentered R**2	0.996338	T x R**2	44.835
Mean of Dependent Variable	1290.1282222		
Std Error of Dependent Variable	321.4472327		
Standard Error of Estimate	85.2888805		
Sum of Squared Residuals	290967.72554		
Regression F(4,40)	146.2526		
Significance Level of F	0.00000000		
Durbin-Watson Statistic	1.843285		
Q(11-0)	13.042680		
Significance Level of Q	0.29054505		

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	221.9954924	77.7862994	2.85392	0.00680882
2.	P{1}	5.0692764	2.9046871	1.74521	0.08862794
3.	Q{1}	0.7834817	0.0955158	8.20264	0.00000000
4.	W{1}	-3.4703992	2.7563844	-1.25904	0.21531394
5.	D96	56.7200953	87.7771010	0.64618	0.52185102

La variable D96 n'a pas d'impact significatif sur le niveau de production. On ne conserve donc pas par la suite cette variable.

3. Impact de la réforme de 1992 sur les décisions de production.

On veut examiner l'effet de la réforme de 1992, effective depuis 1993 sur les décisions de production. Comme on pense que la réforme n'a eu qu'un impact depuis 1993, on crée une variable indicatrice notée D93 qui vaut 1 à partir de 1993, 0 sinon. Dans un premier temps, on pense qu'elle ne modifie pas la réponse des producteurs par rapport aux variations de prix, on introduit cette variable muette de manière additive. Dans ce cas, l'effet marginal d'une variation des prix à court terme est constant sur la période étudiée. Dans un second temps, on fait l'hypothèse que la réforme modifie la réponse des producteurs. Dans ce cas, on souhaite que l'effet marginal d'une variation des prix à court terme soit différent avant la réforme et après la réforme. On introduit alors la variable indicatrice de manière multiplicative. Dans ce cas l'équation estimée s'écrit

$$q_t = \pi_1 + \pi_2 p_{t-1} + \pi_3 q_{t-1} + \pi_5 w_{t-1} + \pi_d D93 * p_{t-1} + \mu_t$$

où π_d représente le coefficient estimé de la variable $D93 * p_{t-1}$.

Les valeurs des effets marginaux d'une variation des prix sont différentes avant 92 et après 92

$$\frac{\partial q_t}{\partial p_{t-1}} = \pi_2 \quad \text{pour } t=1954,1992 \quad \frac{\partial q_t}{\partial p_{t-1}} = \pi_2 + \pi_d \quad \text{pour } t=1993,1998$$

Tableau 8.8 Effet de la réforme 92 sur les décisions de production en supposant qu'elle ne modifie pas la réponse au prix

Dependent Variable Q — Estimation by Least Squares			
Annual Data From 1954 :01 To 1998 :01			
Usable Observations	45	Degrees of Freedom	40
Centered R**2	0.937456	R Bar **2	0.931202
Uncentered R**2	0.996421	T x R**2	44.839
Mean of Dependent Variable	1290.1282222		
Std Error of Dependent Variable	321.4472327		
Standard Error of Estimate	84.3137364		
Sum of Squared Residuals	284352.24572		
Regression F(4,40)	149.8878		
Significance Level of F	0.00000000		
Durbin-Watson Statistic	1.779711		
Q(11-0)	12.045615		
Significance Level of Q	0.36022425		

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	220.9802289	76.8853472	2.87415	0.00645831
2.	P{1}	4.0856862	2.9822390	1.37001	0.17832627
3.	Q{1}	0.7690445	0.0945994	8.12948	0.00000000
4.	W{1}	-2.1024865	2.9215101	-0.71966	0.47591999
5.	D93	-51.1889893	43.9285979	-1.16528	0.25080894

Le tableau 8.8 indique que la variable D93 n'est pas significative.

Tableau 8.9 Effet de la réforme 92 sur les décisions de production en supposant qu'elle modifie la réponse au prix

Dependent Variable Q — Estimation by Least Squares			
Annual Data From 1954 :01 To 1998 :01			
Usable Observations	45	Degrees of Freedom	40
Centered R**2	0.937874	R Bar **2	0.931662
Uncentered R**2	0.996445	T x R**2	44.840
Mean of Dependent Variable	1290.1282222		
Std Error of Dependent Variable	321.4472327		
Standard Error of Estimate	84.0314967		
Sum of Squared Residuals	282451.69728		
Regression F(4,40)	150.9637		
Significance Level of F	0.00000000		
Durbin-Watson Statistic	1.774944		
Q(11-0)	11.721661		
Significance Level of Q	0.38492461		

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	223.4335477	76.5904556	2.91725	0.00576727
2.	P{1}	4.2428297	2.9260815	1.45000	0.15485376
3.	Q{1}	0.7656547	0.0944772	8.10412	0.00000000
4.	W{1}	-2.2076468	2.8526530	-0.77389	0.44354444
5.	D93P	-0.6101457	0.4770028	-1.27912	0.20822431

Les résultats du tableau 8.9 montrent que $D93 * p_{t-1}$ notée $D93P$ n'est pas significative.

Dans les deux cas, la variable muette n'est pas significative.

Problème 8.4

Une extension : prise en compte des caractéristiques de la production de bœuf

1. On estime finalement la relation suivante

$$q_t = b_0 + b_1 p_{t-1} + b_2 p_{t-2} + b_3 p_{t-3} + b_4 q_{t-1} + b_5 w_{t-1} + b_6 D74 + \varepsilon_t$$

Tableau 8.10 *Extension*

Dependent Variable Q — Estimation by Least Squares					
Annual Data From 1956 :01 To 1998 :01					
Usable Observations	43	Degrees of Freedom	36		
Centered R**2	0.960404	R Bar **2	0.953805		
Uncentered R**2	0.997985	T x R**2	42.913		
Mean of Dependent Variable	1314.2744419				
Std Error of Dependent Variable	307.9136121				
Standard Error of Estimate	66.1798256				
Sum of Squared Residuals	157671.69527				
Regression F(6,36)	145.5321				
Significance Level of F	0.00000000				
Durbin-Watson Statistic	2.082191				
Q(10-0)	15.157917				
Significance Level of Q	0.12640966				

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	169.1302969	69.3388934	2.43918	0.01977951
2.	P{1}	-2.5673326	3.0856066	-0.83203	0.41087465
3.	P{2}	12.9145092	3.7553866	3.43893	0.00149212
4.	P{3}	-6.0313779	2.7078388	-2.22738	0.03226369
5.	Q{1}	0.8357561	0.0798956	10.46061	0.00000000
6.	W{1}	-3.0716424	2.8279176	-1.08619	0.28461682
7.	D74	265.9697444	68.4383759	3.88627	0.00041959

L'élasticité de l'offre est négative par rapport aux prix à court terme. Cet effet correspond au comportement de capitalisation des éleveurs. Quand les producteurs anticipent les cours à la baisse ou à la hausse, ils ont le choix d'abattre ou de conserver leur cheptel.

Ainsi, à court terme, une baisse anticipée des prix conduit les producteurs à arbitrer en faveur des abattages. Cette baisse anticipée entraîne une augmentation de la production. Inversement, une hausse des prix les incite à conserver les animaux pour bénéficier de prix plus élevés plus tard, donc une hausse des prix provoque une diminution des abattages c'est-à-dire, à une diminution de la production.

9

LES MODÈLES À PLUSIEURS ÉQUATIONS

1. Les modèles à équations apparemment non reliées
2. Présentation des modèles à équations simultanées
3. Le problème de l'identification
4. Estimation d'un modèle à équations simultanées
5. Tests de spécification

OBJECTIFS DU CHAPITRE

Présenter l'estimation d'un modèle composé d'équations apparemment non reliées

- Méthode SUR (Zellner)
- Méthode SUR itérative

Présenter les modèles à équations simultanées

- Identification
- Estimation
- Tests de spécification

NIVEAU

MASTER, première et deuxième années

L'objectif de ce chapitre est de présenter les méthodes d'estimation lorsqu'un modèle comporte plusieurs équations. On distingue les modèles à équations apparemment non reliées et les modèles à équations simultanées.

1. LES MODÈLES À ÉQUATIONS APPAREMMENT NON RELIÉES

Dans les modèles à équations apparemment non reliées, on explique dans chaque équation une variable endogène en fonction de variables explicatives exogènes. La forme structurelle du modèle est identique à la forme réduite du modèle, contrairement à un modèle à équations simultanées.

Le problème consiste à estimer un système composé de N équations. Pour chaque équation on dispose de T observations.

$$\underbrace{Y_i}_{(T,1)} = \underbrace{X_i}_{(T,K_i)} \underbrace{\beta_i}_{(K_i,1)} + \underbrace{\varepsilon_i}_{(T,1)} \quad i = 1, \dots, N \quad (9.1)$$

Pour une observation it le modèle s'écrit

$$y_{it} = \beta_i' x_{it} + \varepsilon_{it}$$

L'estimateur des MCO équation par équation est BLUE lorsque les hypothèses suivantes sur les aléas sont respectées

$$\begin{aligned} E(\varepsilon_{it}) &= 0 \\ \text{Cov}(\varepsilon_{it}, \varepsilon_{it'}) &= E(\varepsilon_{it} \cdot \varepsilon_{it'}) = 0 \quad t \neq t' \\ \text{var}(\varepsilon_{it}) &= E(\varepsilon_{it}^2) = \sigma_i^2 \\ \text{Cov}(\varepsilon_{it}, \varepsilon_{jt'}) &= E(\varepsilon_{it} \cdot \varepsilon_{jt'}) = 0 \quad i \neq j, \quad \forall t, t' \end{aligned}$$

Sous ces hypothèses, la matrice de variance-covariance des aléas du système s'écrit

$$\begin{aligned} \text{Var}(\varepsilon) &= E(\varepsilon \varepsilon') = \begin{pmatrix} E(\varepsilon_1 \varepsilon_1') & E(\varepsilon_1 \varepsilon_2') & \cdot & \cdot & E(\varepsilon_1 \varepsilon_N') \\ E(\varepsilon_2 \varepsilon_1') & E(\varepsilon_2 \varepsilon_2') & \cdot & \cdot & E(\varepsilon_2 \varepsilon_N') \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ E(\varepsilon_N \varepsilon_1') & E(\varepsilon_N \varepsilon_2') & \cdot & \cdot & E(\varepsilon_N \varepsilon_N') \end{pmatrix} \\ (NT, NT) & \\ &= \begin{pmatrix} \sigma_1^2 I & 0 & \cdot & \cdot & 0 \\ 0 & \sigma_2^2 I & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \sigma_N^2 I \end{pmatrix} \end{aligned} \quad (9.2)$$

avec ε' la transposée du vecteur ε

$$\varepsilon' = \left(\underbrace{\varepsilon_{11}, \dots, \varepsilon_{1T}}_{\varepsilon'_1}, \dots, \underbrace{\varepsilon_{N1}, \dots, \varepsilon_{NT}}_{\varepsilon'_N} \right)$$

Les éléments $E(\varepsilon_i \varepsilon'_i)$ et $E(\varepsilon_i \varepsilon'_j)$ de la matrice $Var(\varepsilon)$ sont des matrices.

L'élément $E(\varepsilon_i \varepsilon'_i)$ correspond à la matrice de variance-covariance des aléas d'une équation

$$E(\varepsilon_i \varepsilon'_i) = \begin{pmatrix} E(\varepsilon_{i1}^2) & E(\varepsilon_{i1}\varepsilon_{i2}) & \cdot & \cdot & E(\varepsilon_{i1}\varepsilon_{iT}) \\ E(\varepsilon_{i2}\varepsilon_{i1}) & E(\varepsilon_{i2}^2) & \cdot & \cdot & E(\varepsilon_{i2}\varepsilon_{iT}) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ E(\varepsilon_{iT}\varepsilon_{i1}) & E(\varepsilon_{iT}\varepsilon_{i2}) & \cdot & \cdot & E(\varepsilon_{iT}^2) \end{pmatrix} = \sigma_i^2 I$$

et l'élément $E(\varepsilon_i \varepsilon'_j)$ donne la matrice des covariances des aléas inter-équations

$$E(\varepsilon_i \varepsilon'_j) = \begin{pmatrix} E(\varepsilon_{i1}\varepsilon_{j1}) & E(\varepsilon_{i1}\varepsilon_{j2}) & \cdot & \cdot & E(\varepsilon_{i1}\varepsilon_{jT}) \\ E(\varepsilon_{i2}\varepsilon_{j1}) & E(\varepsilon_{i2}\varepsilon_{j2}) & \cdot & \cdot & E(\varepsilon_{i2}\varepsilon_{jT}) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ E(\varepsilon_{iT}\varepsilon_{j1}) & E(\varepsilon_{iT}\varepsilon_{j2}) & \cdot & \cdot & E(\varepsilon_{iT}\varepsilon_{jT}) \end{pmatrix} = 0 \quad i \neq j$$

Lorsque les corrélations des aléas inter-équations sont non nulles l'estimateur des MCO n'est pas efficace. On applique dans ce cas la méthode de Zellner appelée également la méthode SUR (Seemingly Unrelated Regressions). Il s'agit de la méthode des MCG appliquée à un système de N équations.

Le système (9.1) peut s'écrire

$$\underbrace{\begin{pmatrix} Y_1 \\ Y_2 \\ \cdot \\ Y_N \end{pmatrix}}_{\substack{Y \\ (NT,1)}} = \underbrace{\begin{pmatrix} X_1 & 0 & 0 & \cdot & 0 \\ 0 & X_2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & X_N \end{pmatrix}}_{\substack{X \\ (NT,NK)}} \underbrace{\begin{pmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \beta_N \end{pmatrix}}_{\substack{\beta \\ (NK,1)}} + \underbrace{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \varepsilon_N \end{pmatrix}}_{\substack{\varepsilon \\ (NT,1)}} \quad (9.3)$$

avec $K = \sum_{i=1}^N K_i$.

Lorsque les variables explicatives sont toutes strictement exogènes et les aléas inter-équations corrélés à une même période

$$\begin{aligned} cov(\varepsilon_{it}, \varepsilon_{jt}) &= E(\varepsilon_{it}\varepsilon_{jt}) = \sigma_{ij} \quad i \neq j \\ \implies E(\varepsilon_i \varepsilon'_j) &= \sigma_{ij} I \quad i \neq j \end{aligned}$$

et la matrice de variance-covariance des aléas du système est donnée par

$$\text{Var}(\varepsilon) = \begin{pmatrix} \sigma_1^2 I & \sigma_{12} I & \cdot & \cdot & \sigma_{1N} I \\ \sigma_{21} I & \sigma_2^2 I & \cdot & \cdot & \sigma_{2N} I \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \sigma_{N1} I & \sigma_{N2} I & \cdot & \cdot & \sigma_N^2 I \end{pmatrix} = \Psi \quad (9.4)$$

La méthode de Zellner consiste à appliquer les MCG sur le système (9.3)

$$\widehat{\beta}_{SUR} = (X' \Psi^{-1} X)^{-1} (X' \Psi^{-1} Y) \quad (9.5)$$

et

$$\begin{aligned} E(\widehat{\beta}_{SUR}) &= \beta \\ \text{Var}(\widehat{\beta}_{SUR}) &= (X' \Psi^{-1} X)^{-1} \end{aligned} \quad (9.6)$$

Si la matrice de variance-covariance Ψ est inconnue, on doit procéder de la manière suivante

1. estimer chaque équation ($i = 1, \dots, N$) par la méthode des MCO

$$\widehat{\beta}_{MCO,i} = (X_i' X_i)^{-1} (X_i' Y_i)$$

et calculer les vecteurs de résidus e_i

$$e_i = Y_i - X_i \widehat{\beta}_{MCO,i}$$

2. calculer des estimateurs convergent de σ_i^2 et σ_{ij} ($i, j = 1, \dots, N$)

$$\widehat{\sigma}_i^2 = \frac{\sum_{t=1}^T e_{it}^2}{T} \quad \text{et} \quad \widehat{\sigma}_{ij} = \frac{\sum_{t=1}^T e_{it} e_{jt}}{T}$$

3. estimer le système composé des N équations par

$$\widehat{\beta}_{SUR} = (X' \widehat{\Psi}^{-1} X)^{-1} (X' \widehat{\Psi}^{-1} Y) \quad (9.7)$$

On peut poursuivre l'estimation en procédant de manière itérative (méthode de Zellner itérative) : à partir de l'étape 3 il faut calculer les résidus de chaque équation et reprendre à l'étape 2. On arrête les itérations lorsque les valeurs estimées des paramètres ne varient quasiment plus.

Pour terminer, trois remarques importantes s'imposent :

1. Dans un système où les variables endogènes sont des parts, dont la somme est égale à 1, il y a une équation redondante et la matrice de variance-covariance des aléas Ψ est singulière (cf. chapitre 10 Modélisation d'un système de demande). Il faut dans ce cas omettre une équation du système, cette équation pouvant ensuite être déduite de l'estimation des autres équations. La méthode de Zellner itérative est alors privilégiée car les résultats obtenus sont indépendants du choix de l'équation omise.
2. Si dans le système les variables explicatives sont identiques pour chaque équation, c'est-à-dire $X_i = \underline{X} \forall i = 1, \dots, N$, alors l'estimateur des MCO équation par équation est identique à l'estimateur des MCG même si les aléas inter-équations sont corrélés à la même période. Le système cependant ne doit pas comporter de contrainte sur les paramètres.
3. La méthode SUR peut s'appliquer lorsque les aléas sont autocorrélés. Soit le modèle

$$y_{it} = \beta'_i x_{it} + \varepsilon_{it}, \quad i = 1, \dots, N \quad ; \quad t = 1, \dots, T$$

avec $\varepsilon_{it} = \rho_i \varepsilon_{it-1} + u_{it}$. Il faut

- premièrement, estimer chaque équation par les MCO et estimer $\hat{\rho}_i$ à partir des résidus (cf. chapitre 6, Les hypothèses du modèle de régression linéaire multiple),
- deuxièmement, estimer chaque équation du modèle transformé (modèle quasi-différencié si utilisation de la procédure de Cochrane-Orcutt) par les MCO afin d'estimer la matrice de variance-covariance Ψ
- troisièmement estimer

$$\hat{\beta}_{SUR} = \left(X'_* \hat{\Psi}^{-1} X_* \right)^{-1} \left(X'_* \hat{\Psi}^{-1} Y_* \right) \quad (9.8)$$

avec X_* et Y_* les variables du modèle quasi-différencié.

2. PRÉSENTATION DES MODÈLES À ÉQUATIONS SIMULTANÉES

Afin de simplifier la présentation des modèles à équations simultanées, on s'appuiera tout au long de ce chapitre sur l'exemple d'un modèle de marché. On considère un modèle de marché composé de trois équations

$$\begin{aligned} Q^d &= \gamma_{11} + \beta_{11}P + \gamma_{21}R + \gamma_{31}Ps + \varepsilon_1 \\ Q^O &= \gamma_{12} + \beta_{12}P + \gamma_{22}Pf + \varepsilon_2 \\ Q^d &= Q^O \end{aligned} \quad (9.9)$$

L'équation de demande est dérivée d'un problème de maximisation de l'utilité des consommateurs. On note Q^d les quantités demandées, P le prix du bien, R le revenu des ménages et Ps le prix d'un substitut. L'équation d'offre est obtenue

dans le cadre d'un problème de maximisation des profits. On note Q^O les quantités offertes et Pf le prix d'un input. Le marché est supposé être à l'équilibre et le modèle (9.9) peut s'écrire sous la forme d'un modèle à 2 équations, soit $Q = Q^d = Q^O$

$$\begin{cases} Q = \gamma_{11} + \beta_{11}P + \gamma_{21}R + \gamma_{31}Ps + \varepsilon_1 \\ Q = \gamma_{12} + \beta_{12}P + \gamma_{22}Pf + \varepsilon_2 \end{cases} \quad (9.10)$$

ε_1 et ε_2 sont les aléas du modèle. On suppose que

$$\begin{aligned} E(\varepsilon_1) &= E(\varepsilon_2) = 0 \\ E(\varepsilon_1\varepsilon_1') &= \sigma_1^2 I \text{ et } E(\varepsilon_2\varepsilon_2') = \sigma_2^2 I \end{aligned}$$

On dispose d'un nombre d'observations N pour chaque variable.

Les variables endogènes du modèle sont Q et P . Les variables exogènes sont R , Ps et Pf . Ce modèle est dérivé de la théorie économique et il est donc écrit sous une forme structurelle. L'écriture de la forme structurelle du modèle (FSM) sous forme matricielle est

$$\underbrace{\begin{pmatrix} 1 & -\beta_{11} \\ 1 & -\beta_{12} \end{pmatrix}}_B \underbrace{\begin{pmatrix} Q_i \\ P_i \end{pmatrix}}_{Y_i} + \underbrace{\begin{pmatrix} -\gamma_{11} & -\gamma_{21} & -\gamma_{31} & 0 \\ -\gamma_{12} & 0 & 0 & -\gamma_{22} \end{pmatrix}}_\Gamma \underbrace{\begin{pmatrix} 1 \\ R_i \\ Ps_i \\ Pf_i \end{pmatrix}}_{X_i} = \underbrace{\begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix}}_{\varepsilon_i} \quad (9.11)$$

De manière générale la FSM d'un modèle pour une observation i s'écrit

$$\underbrace{B}_{(G,G)} \underbrace{Y_i}_{(G,1)} + \underbrace{\Gamma}_{(G,K)} \underbrace{X_i}_{(K,1)} = \underbrace{\varepsilon_i}_{(G,1)} \quad (9.12)$$

avec G le nombre de variables endogènes, qui correspond au nombre d'équations et K le nombre de variables exogènes. On note Y_i le vecteur des variables endogènes pour l'observation i et X_i le vecteur des variables strictement exogènes¹, y compris la constante. B est la matrice des coefficients des variables endogènes et Γ la matrice des coefficients associée à X_i .

Dans un modèle à équations simultanées, on cherche à estimer les paramètres de la FSM. Cependant le modèle peut s'écrire sous une forme réduite. La forme réduite du modèle (FRM) correspond à une écriture du modèle où apparaissent, à gauche, pour chaque équation une variable endogène, et à droite toutes les variables

¹ Une variable explicative x_{ki} est strictement exogène si $E(x_{ki}\varepsilon_j) = 0, \forall i, j$, elle est exogène si $E(x_{ki}\varepsilon_i) = 0, \forall i$.

exogènes du modèle. Si on prémultiplie la forme (9.12) par la matrice B^{-1} on obtient la FRM

$$\boxed{Y_i = -B^{-1}\Gamma X_i + B^{-1}\varepsilon_i = \Pi X_i + V_i} \quad (9.13)$$

avec $\Pi = -B^{-1}\Gamma$ et $V_i = B^{-1}\varepsilon_i$. Π est la matrice des paramètres de la FRM, elle est d'ordre (G, K) et V_i est le vecteur des aléas de la forme réduite, il est d'ordre $(G, 1)$. Ainsi la FRM du modèle d'offre et de demande s'écrit

$$\begin{pmatrix} Q_i \\ P_i \end{pmatrix} = \underbrace{\begin{pmatrix} \pi_{11} & \pi_{21} & \pi_{31} & \pi_{41} \\ \pi_{12} & \pi_{22} & \pi_{32} & \pi_{42} \end{pmatrix}}_{\Pi} \begin{pmatrix} R_i \\ P_{s_i} \\ P_{f_i} \end{pmatrix} + \underbrace{\begin{pmatrix} v_{1i} \\ v_{2i} \end{pmatrix}}_{V_i} \quad (9.14)$$

Ce passage à la FRM permet d'expliquer pourquoi on ne peut pas estimer chaque équation d'un modèle à équations simultanées avec l'estimateur des MCO. Considérons la fonction de demande dont on cherche à évaluer les paramètres $\gamma_{11}, \beta_{11}, \gamma_{21}$, et γ_{31} . Dans cette équation, la variable endogène P apparaît en variable explicative. Cette variable étant endogène, sa covariance avec l'aléas ε_1 est non nulle. En effet, la FRM (9.14) indique que

$$P_i = \pi_{12} + \pi_{22}R_i + \pi_{32}P_{s_i} + \pi_{42}P_{f_i} + v_{2i} \quad (9.15)$$

or v_{2i} correspond à la seconde ligne de la matrice $V_i = B^{-1}\varepsilon_i$.

$$V_i = \begin{pmatrix} v_{1i} \\ v_{2i} \end{pmatrix} = \underbrace{\frac{1}{-\beta_{12} + \beta_{11}} \begin{pmatrix} -\beta_{12} & \beta_{11} \\ -1 & 1 \end{pmatrix}}_{B^{-1}} \begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \quad (9.16)$$

d'où

$$v_{2i} = \frac{1}{-\beta_{12} + \beta_{11}} (-\varepsilon_{1i} + \varepsilon_{2i}) \quad (9.17)$$

On montre ainsi que la covariance entre la variable P et l'aléa ε_1 est non nulle²

$$\begin{aligned} cov(P_i, \varepsilon_{1i}) &= E(P_i - E(P_i))(\varepsilon_{1i} - E(\varepsilon_{1i})) \\ &= E(v_{2i} \cdot \varepsilon_{1i}) = \frac{-1}{-\beta_{12} + \beta_{11}} E(\varepsilon_{1i}^2) \neq 0 \end{aligned} \quad (9.18)$$

l'estimateur des MCO de la fonction de demande est donc biaisé et non convergent.

² D'après la FRM

$$P_i = \pi_{12} + \pi_{22}R_i + \pi_{32}P_{s_i} + \pi_{42}P_{f_i} + v_{2i}$$

et

$$E(P_i) = \pi_{12} + \pi_{22}R_i + \pi_{32}P_{s_i} + \pi_{42}P_{f_i}$$

d'où

$$P_i - E(P_i) = v_{2i}$$

Plus généralement, dans un modèle à équations simultanées, les variables endogènes explicatives du modèle ont une covariance non nulle avec les aléas de la FSM. Par conséquent, l'estimateur des MCO de la FSM est biaisé et non convergent, il s'agit d'un biais de simultanéité. Le problème consiste donc à définir un estimateur convergent et asymptotiquement efficace des paramètres de la FSM. Auparavant, on doit se poser la question de l'identification des paramètres du modèle : il s'agit de déterminer si on dispose de suffisamment d'informations pour estimer tous les paramètres de la FSM.

3. LE PROBLÈME DE L'IDENTIFICATION

3.1 Présentation du problème de l'identification

Le problème de l'identification peut être examiné à partir de la FRM. Sous cette forme le modèle peut être estimé avec la méthode des MCO car chaque équation comporte une variable endogène et toutes les variables explicatives exogènes du modèle

$$Y = X\Pi' + V \quad (9.19)$$

L'estimateur des MCO des paramètres Π de la FRM est convergent

$$\begin{array}{l} \hat{\Pi}'_{MCO} = (X'X)^{-1} X'Y \\ p \lim \hat{\Pi}'_{MCO} = \Pi \end{array} \quad (9.20)$$

Cependant, la connaissance de la FRM ne permet pas de déterminer la FSM. En effet, à une forme structurelle, sans contrainte sur les paramètres, correspond une forme réduite, mais la réciproque n'est pas vraie. Si la structure initiale (9.12) est multipliée par une matrice non singulière F

$$(FB)Y_i + (FT)X_i = F\varepsilon_i \quad (9.21)$$

alors la FRM n'est pas modifiée

$$Y_i = -B^{-1}F^{-1}FTX_i + B^{-1}F^{-1}F\varepsilon_i = \Pi X_i + V_i \quad (9.22)$$

Il faut des contraintes sur les paramètres de la FSM pour identifier la FSM à partir de la FRM, c'est-à-dire pour obtenir des matrices de transformation admissibles F uniquement diagonales.

3.2 Identification et FRM

Pour identifier la structure du modèle on peut utiliser les paramètres de la FRM, mais il faut également des informations liées à la théorie économique. Par exemple, dans le modèle (9.10), selon la théorie économique, la variable de revenu

R et la variable de prix P s ont un coefficient nul dans l'équation d'offre, ces deux variables n'expliquent pas l'offre de biens, de même le prix de l'input Pf n'explique pas la demande. La théorie économique conduit à imposer des contraintes (contraintes d'exclusions, contraintes linéaires,...) sur les paramètres de la FSM qui réduisent le nombre de paramètres à estimer.

Exemple 9.1

Cas d'un modèle dont les équations ne sont pas identifiées.

Considérons le modèle de marché à l'équilibre dans lequel les paramètres γ_{21} , γ_{31} et γ_{22} sont nuls

$$\begin{aligned} Q &= \gamma_{11} + \beta_{11}P + \varepsilon_1 \\ Q &= \gamma_{12} + \beta_{12}P + \varepsilon_2 \end{aligned} \quad (9.23)$$

Dans ce modèle il n'apparaît que des variables endogènes et la FRM s'écrit

$$\begin{pmatrix} Q_i \\ P_i \end{pmatrix} = \begin{pmatrix} \pi_{11} \\ \pi_{12} \end{pmatrix} ((1)) + \begin{pmatrix} v_{1i} \\ v_{2i} \end{pmatrix} \quad (9.24)$$

La relation entre les paramètres de la FSM et de la FRM est donnée par

$$\Pi = -B^{-1}\Gamma \Leftrightarrow B\Pi = -\Gamma \quad (9.25)$$

Il s'agit d'un système d'équations à résoudre dans lequel on suppose, d'une part, que les paramètres de la matrice Π sont connus, car ils peuvent être estimés par la méthode des MCO et, d'autre part, que les paramètres des matrices B et Γ sont inconnus

$$\begin{pmatrix} 1 & -\beta_{11} \\ 1 & -\beta_{12} \end{pmatrix} \begin{pmatrix} \pi_{11} \\ \pi_{12} \end{pmatrix} = \begin{pmatrix} \gamma_{11} \\ \gamma_{12} \end{pmatrix} \quad (9.26)$$

Il faut résoudre les 2 équations

$$\begin{aligned} \pi_{11} - \beta_{11}\pi_{12} &= \gamma_{11} \\ \pi_{11} - \beta_{12}\pi_{12} &= \gamma_{12} \end{aligned} \quad (9.27)$$

Dans ce système il y a 4 inconnues : β_{11} , γ_{11} , β_{12} et γ_{12} et deux équations. Il n'existe pas de solution. Les équations du modèle ne sont pas identifiées.

Exemple 9.2

Cas d'un modèle où les équations sont exactement identifiées.

Dans le modèle de marché on suppose maintenant que seul le paramètre γ_{31} est nul

$$\begin{aligned} Q &= \gamma_{11} + \beta_{11}P + \gamma_{21}R + \varepsilon_1 \\ Q &= \gamma_{12} + \beta_{12}P + \gamma_{22}Pf + \varepsilon_2 \end{aligned} \quad (9.28)$$

Les deux équations de ce modèle sont exactement identifiées. En effet, la FRM est donnée par

$$\begin{pmatrix} Q_i \\ P_i \end{pmatrix} = \begin{pmatrix} \pi_{11} & \pi_{21} & \pi_{41} \\ \pi_{12} & \pi_{22} & \pi_{42} \end{pmatrix} \begin{pmatrix} 1 \\ R_i \\ Pf_i \end{pmatrix} + \begin{pmatrix} v_{1i} \\ v_{2i} \end{pmatrix} \quad (9.29)$$

et on doit résoudre le système d'équations

$$\begin{pmatrix} 1 & -\beta_{11} \\ 1 & -\beta_{12} \end{pmatrix} \begin{pmatrix} \pi_{11} & \pi_{21} & \pi_{41} \\ \pi_{12} & \pi_{22} & \pi_{42} \end{pmatrix} = \begin{pmatrix} \gamma_{11} & \gamma_{21} & 0 \\ \gamma_{12} & 0 & \gamma_{22} \end{pmatrix} \quad (9.30)$$

Les paramètres de la première équation sont obtenus en résolvant les équations

$$\begin{aligned} \pi_{11} - \beta_{11}\pi_{12} &= \gamma_{11} \\ \pi_{21} - \beta_{11}\pi_{22} &= \gamma_{21} \\ \pi_{41} - \beta_{11}\pi_{42} &= 0 \end{aligned} \quad (9.31)$$

grâce à la dernière équation on obtient $\beta_{11} = \pi_{41}/\pi_{42}$, on peut alors en déduire γ_{21} et γ_{11} .

Pour la deuxième équation on doit résoudre

$$\begin{aligned} \pi_{11} - \beta_{12}\pi_{12} &= \gamma_{12} \\ \pi_{21} - \beta_{12}\pi_{22} &= 0 \\ \pi_{41} - \beta_{12}\pi_{42} &= \gamma_{22} \end{aligned} \quad (9.32)$$

la seconde équation donne $\beta_{12} = \pi_{21}/\pi_{22}$ et on en déduit γ_{12} et γ_{22} .

Par conséquent, les deux équations du modèle d'offre et de demande sont exactement identifiées. À partir de la FRM et des contraintes sur la forme structurelle, on peut trouver une valeur unique pour les paramètres de la forme structurelle. Dans l'équation de demande le coefficient de la variable Pf est nul, et dans l'équation d'offre le coefficient de la variable R est égal à 0. Ces deux contraintes permettent d'identifier la structure du modèle. Ce sont des restrictions d'exclusions.

Exemple 9.3

Cas d'un modèle avec une équation sur-identifiée.

Dans cet exemple on étudie l'identification du modèle dont la FSM est donnée par (9.11) et la FRM par (9.14). Pour déterminer si le modèle est identifié on résout le système d'équations

$$\begin{pmatrix} 1 & -\beta_{11} \\ 1 & -\beta_{12} \end{pmatrix} \begin{pmatrix} \pi_{11} & \pi_{21} & \pi_{31} & \pi_{41} \\ \pi_{12} & \pi_{22} & \pi_{32} & \pi_{42} \end{pmatrix} = \begin{pmatrix} \gamma_{11} & \gamma_{21} & \gamma_{31} & 0 \\ \gamma_{12} & 0 & 0 & \gamma_{22} \end{pmatrix} \quad (9.33)$$

Pour l'équation de demande on doit résoudre

$$\begin{aligned} \pi_{11} - \beta_{11}\pi_{12} &= \gamma_{11} \\ \pi_{21} - \beta_{11}\pi_{22} &= \gamma_{21} \\ \pi_{31} - \beta_{11}\pi_{32} &= \gamma_{31} \\ \pi_{41} - \beta_{11}\pi_{42} &= 0 \end{aligned} \quad (9.34)$$

La dernière équation permet de calculer $\beta_{11} = \pi_{41}/\pi_{42}$ et on en déduit les autres paramètres de la fonction de demande. La fonction de demande est exactement identifiée.

Pour l'équation d'offre on résout

$$\begin{aligned} \pi_{11} - \beta_{12}\pi_{12} &= \gamma_{12} \\ \pi_{21} - \beta_{12}\pi_{22} &= 0 \\ \pi_{31} - \beta_{12}\pi_{32} &= 0 \\ \pi_{41} - \beta_{12}\pi_{42} &= \gamma_{22} \end{aligned} \quad (9.35)$$

Il existe deux structures possibles pour l'équation d'offre. La seconde équation permet de calculer $\beta_{12} = \pi_{21}/\pi_{22}$ et on peut en déduire γ_{12} et γ_{22} . La troisième équation donne aussi une valeur de $\beta_{12} = \pi_{31}/\pi_{32}$, différente de la précédente, et à partir de laquelle on peut calculer les valeurs correspondantes pour les paramètres γ_{12} et γ_{22} . L'équation d'offre est sur-identifiée car il y a deux variables exclues de l'équation dans ce modèle, qui comporte deux équations.

Ces trois exemples mettent en évidence des équations non identifiées, exactement identifiées et sur-identifiées. Dans un modèle, ces trois formes d'équations peuvent se rencontrer. On ne peut estimer que des équations de la FSM exactement identifiées ou sur-identifiées.

3.3 Les conditions d'ordre et de rang

Une manière rapide de déterminer l'identification d'un modèle consiste à appliquer les conditions d'ordre et de rang. Ces conditions cherchent à déterminer si le nombre de contraintes sur les paramètres de la FSM est suffisant pour identifier les équations de la FSM. Pour illustrer l'application de ces conditions, dans le cas de restrictions d'exclusion, on reprend le modèle (9.10). Le tableau suivant donne les paramètres de la FSM.

Tableau 9.1 Coefficients de la FSM

Q	P	(1)	R	Ps	Pf	Nombre de variables exclues
1	β_{11}	γ_{11}	γ_{21}	$-\gamma_{31}$	0	1
1	$-\beta_{12}$	$-\gamma_{11}$	0	0	$-\gamma_{22}$	2

Le nombre de variables exclues dans chaque équation est déterminé par le nombre de 0. Selon la condition d'ordre, si pour une équation le nombre de variables exclues, noté J , est supérieur ou égal au nombre d'équations moins 1 c'est-à-dire $(G-1)^3$, alors l'équation est identifiée. Pour l'équation de demande $J = 1 = G - 1$, pour l'équation d'offre $J = 2 > 1$.

La condition d'ordre est une condition nécessaire, tandis que la condition de rang est nécessaire et suffisante. Pour une équation donnée, la condition de rang consiste à conserver les colonnes du tableau des coefficients de la FSM où apparaissent les restrictions. On construit ainsi une matrice de coefficients, notée Ψ , dont il faut déterminer si le rang est égal à $(G-1)^4$ c'est-à-dire égal à 1 dans cet exemple. Pour l'équation de demande, il s'agit de la matrice

$$\Psi = \begin{pmatrix} 0 \\ \gamma_{22} \end{pmatrix}$$

dans cette matrice il existe un déterminant d'ordre (1,1) égal à $|\gamma_{22}| \neq 0$. On conclut que l'équation de demande est exactement identifiée. Pour l'équation d'offre on doit déterminer le rang de la matrice suivante

$$\Psi = \begin{pmatrix} -\gamma_{21} & -\gamma_{31} \\ 0 & 0 \end{pmatrix}$$

Pour cette matrice il existe deux déterminants d'ordre (1,1) différents de 0, $|\gamma_{21}|$ et $|\gamma_{31}|$, par conséquent l'équation d'offre est sur-identifiée.

4. ESTIMATION D'UN MODÈLE À ÉQUATIONS SIMULTANÉES

On peut estimer les paramètres de la FSM lorsque les équations sont exactement identifiées ou sur-identifiées. On distingue les méthodes à information limitée et les méthodes à information complète. Les premières consistent à estimer équation par équation le modèle, sous l'hypothèse qu'il n'existe pas de corrélations entre les

³ On enlève 1 car le modèle est normalisé, c'est-à-dire que pour chaque équation du modèle une variable a un coefficient égal à 1.

⁴ Le rang de la matrice P est égal à $G-1$ s'il existe au moins un déterminant d'ordre $(G-1, G-1)$ dans P . S'il existe un seul déterminant différent de 0 alors l'équation est exactement identifiée et s'il en existe plusieurs elle est sur-identifiée. S'il n'existe pas de déterminant d'ordre $(G-1, G-1)$ différent de 0 alors l'équation n'est pas identifiée.

aléas des différentes équations. Les secondes considèrent le modèle dans sa globalité et estiment les paramètres sous l'hypothèse qu'il existe des corrélations entre les aléas interéquations. Ici on présente deux méthodes souvent appliquées : une méthode à information limitée, la méthode des Doubles Moindres Carrés (DMC), et une méthode à information complète, la méthode des Triples Moindres Carrés (TMC).

4.1 La méthode des DMC

L'estimateur des DMC est un estimateur des Variables Instrumentales. Il est appliqué équation par équation. La première étape consiste à estimer la FRM afin d'obtenir des variables instrumentales pour les variables endogènes présentes dans l'équation comme variables explicatives. La seconde étape consiste à appliquer l'estimateur des Variables Instrumentales.

Par exemple, considérons l'équation de demande dans le modèle de marché (9.10). Dans cette équation la variable de prix P est une variable explicative endogène et donc $cov(P_i, \varepsilon_{1i}) \neq 0$. Pour obtenir un estimateur convergent des paramètres de l'équation de demande il faut chercher une variable instrumentale, cette variable doit être fortement corrélée à la variable P et non corrélée à l'aléa ε_1 . Dans la méthode des DMC l'instrument choisit est la variable correspondant à la valeur estimée de P dans la FRM, on note \hat{P}_{FRM}^5 cette variable

$$\hat{P}_{FRM i} = \hat{\pi}_{12} + \hat{\pi}_{22}R_i + \hat{\pi}_{32}Ps_i + \hat{\pi}_{42}Pf_i \quad i = 1, \dots, N \quad (9.36)$$

la variable \hat{P}_{FRM} est une variable fortement corrélée à P et non corrélée à ε_1 , car les variables R , Ps et Pf sont strictement exogènes et donc non corrélées aux aléas de la FSM. Cette variable peut donc être utilisée comme variable instrumentale. Les paramètres de l'équation de demande sont obtenus de la manière suivante

$$\hat{\delta}_{DMC,1} = \begin{pmatrix} \widehat{\gamma}_{11} \\ \widehat{\beta}_{11} \\ \widehat{\gamma}_{12} \\ \widehat{\gamma}_{13} \end{pmatrix} = \left(\widehat{X}^{d'} X^d \right)^{-1} \left(\widehat{X}^{d'} Y^d \right) =$$

$$\left[\underbrace{\begin{pmatrix} 1 & 1 & \dots & 1 \\ \hat{P}_{FRM,1} & \hat{P}_{FRM,2} & \dots & \hat{P}_{FRM,N} \\ R_1 & R_2 & \dots & R_N \\ Ps_1 & Ps_2 & \dots & Ps_N \end{pmatrix}}_{\widehat{X}^{d'}} \underbrace{\begin{pmatrix} 1 & P_1 & R_1 & Ps_1 \\ 1 & P_2 & R_2 & Ps_2 \\ \cdot & \cdot & \cdot & \cdot \\ 1 & P_N & R_N & Ps_N \end{pmatrix}}_{X^d} \right]^{-1} \quad (9.37)$$

$$\left[\underbrace{\begin{pmatrix} 1 & 1 & \dots & 1 \\ \hat{P}_{FRM,1} & \hat{P}_{FRM,2} & \dots & \hat{P}_{FRM,N} \\ R_1 & R_2 & \dots & R_N \\ Ps_1 & Ps_2 & \dots & Ps_N \end{pmatrix}}_{\widehat{X}^{d'}} \underbrace{\begin{pmatrix} Q_1 \\ Q_2 \\ \cdot \\ Q_N \end{pmatrix}}_{Y^d} \right]$$

⁵ La méthode d'estimation retenue est la méthode des MCO.

ou de manière équivalente en appliquant les MCO sur l'équation

$$Q_i = \gamma_{11} + \beta_{11}\widehat{P}_{FRM,i} + \gamma_{21}R_i + \gamma_{31}Ps_i + \varepsilon_{1i} \quad i = 1, \dots, N \quad (9.38)$$

$$\widehat{\delta}_{DMC,1} = \left(\widehat{X}^{d'}\widehat{X}^d\right)^{-1}\left(\widehat{X}^{d'}Y^d\right) \quad (9.39)$$

car les matrices symétriques \widehat{X}'^dX^d et $\widehat{X}'^d\widehat{X}^d$ sont égales

$$\widehat{\delta}_{DMC,1} = \left(\widehat{X}^{d'}\widehat{X}^d\right)^{-1}\left(\widehat{X}^{d'}Y^d\right) = \left(\widehat{X}^{d'}X^d\right)^{-1}\left(\widehat{X}^{d'}Y^d\right) \quad (9.40)$$

L'estimateur des DMC est un estimateur convergent et sa matrice de variance covariance asymptotique est donnée par

$$\begin{aligned} Asym.Var\left(\widehat{\delta}_{DMC,1}\right) &= \frac{\sigma_1^2}{N}\left(\frac{\widehat{X}^{d'}X^d}{N}\right)^{-1}\left(\frac{\widehat{X}^{d'}\widehat{X}^d}{N}\right)\left(\frac{X^{d'}\widehat{X}^d}{N}\right)^{-1} \\ &= \sigma_1^2\left(\widehat{X}^{d'}\widehat{X}^d\right)^{-1} \end{aligned} \quad (9.41)$$

la variance des aléas pouvant être estimée par

$$\widehat{\sigma}_1^2 = \frac{\sum_{i=1}^N \left[Q_i - \left(\widehat{\gamma}_{DMC,11} + \widehat{\beta}_{DMC,11}P_i + \widehat{\gamma}_{DMC,12}R_i + \widehat{\gamma}_{DMC,13}Ps_i\right)\right]^2}{N} \quad (9.42)$$

Dans cette procédure d'estimation on ignore les corrélations éventuelles entre les aléas de l'équation d'offre et de l'équation de demande. Par conséquent, s'il existe des corrélations entre ces aléas, l'estimateur des DMC ne sera pas asymptotiquement efficace car il n'utilise pas toutes les informations contenues dans le modèle. Dans ce cas il faut appliquer la méthode des TMC.

4.2 La méthode des TMC

La méthode des TMC estime le système d'équations et prend en compte les corrélations des aléas interéquations à une même période. Le modèle (9.11) peut

s'écrire sous la forme matricielle

$$\underbrace{\begin{pmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_N \\ Q_1 \\ Q_2 \\ \vdots \\ Q_N \end{pmatrix}}_{\tilde{Y}} = \underbrace{\begin{pmatrix} 1 & P_1 & R_1 & P s_1 & 0 & 0 & 0 \\ 1 & P_2 & R_2 & P s_2 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & P_N & R_N & P s_N & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & P_1 & P f_1 \\ 0 & 0 & 0 & 0 & 1 & P_2 & P f_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 & P_N & P f_N \end{pmatrix}}_{\tilde{X}} \underbrace{\begin{pmatrix} \gamma_{11} \\ \beta_{11} \\ \gamma_{21} \\ \gamma_{31} \\ \gamma_{12} \\ \beta_{12} \\ \gamma_{22} \end{pmatrix}}_{\delta} + \underbrace{\begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{1N} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \vdots \\ \varepsilon_{2N} \end{pmatrix}}_{\varepsilon} \quad (9.43)$$

Si dans le modèle d'offre et de demande on suppose que

$$\begin{aligned} cov(\varepsilon_{1i}, \varepsilon_{2i}) &= E(\varepsilon_{1i} \varepsilon_{2i}) = \sigma_{12} \\ cov(\varepsilon_{1i}, \varepsilon_{2j}) &= E(\varepsilon_{1i} \varepsilon_{2j}) = 0 \quad i \neq j \end{aligned}$$

alors la matrice de variance-covariance des aléas de la FSM notée Σ est donnée par $E(\varepsilon\varepsilon')$

$$\begin{aligned} \Sigma &= E \left[\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} (\varepsilon_1' \quad \varepsilon_2') \right] = \begin{pmatrix} E(\varepsilon_1 \varepsilon_1') & E(\varepsilon_1 \varepsilon_2') \\ E(\varepsilon_2 \varepsilon_1') & E(\varepsilon_2 \varepsilon_2') \end{pmatrix} \\ &= \begin{pmatrix} \sigma_1^2 I & \sigma_{12} I \\ \sigma_{12} I & \sigma_2^2 I \end{pmatrix} \end{aligned} \quad (9.44)$$

Le problème consiste donc à estimer le modèle

$$\boxed{\begin{aligned} \tilde{Y} &= \tilde{X} \delta + \varepsilon \quad \text{avec} \\ E(\varepsilon\varepsilon') &= \Sigma \end{aligned}} \quad (9.45)$$

et $p \lim \frac{\tilde{X}'\varepsilon}{N} \neq 0$ car $cov(P_i, \varepsilon_{1i}) \neq 0$ et $cov(P_i, \varepsilon_{2i}) \neq 0$. Ainsi, d'une part, pour corriger le biais de simultanéité il faut utiliser l'estimateur des Variables Instrumentales et, d'autre part, pour tenir compte des corrélations des aléas interéquations il faut appliquer l'estimateur des MCG. La méthode des TMC combine les deux méthodes : elle utilise les mêmes instruments que la méthode des DMC et pondère les observations par $\Sigma^{-1/2}$. L'estimateur des TMC est donné par l'expression

$$\boxed{\hat{\delta}_{TMC} = \left(\hat{\tilde{X}}' \Sigma^{-1} \hat{\tilde{X}} \right)^{-1} \left(\hat{\tilde{X}}' \Sigma^{-1} \tilde{Y} \right)} \quad (9.46)$$

avec

$$\widehat{X} = \begin{pmatrix} 1 & \widehat{P}_{FRM,1} & R_1 & P_{s1} & 0 & 0 & 0 \\ 1 & \widehat{P}_{FRM,2} & R_2 & P_{s2} & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & & \\ 1 & \widehat{P}_{FRM,N} & R_N & P_{sN} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \widehat{P}_{FRM,1} & Pf_1 \\ 0 & 0 & 0 & 0 & 1 & \widehat{P}_{FRM,2} & Pf_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 1 & \widehat{P}_{FRM,N} & Pf_N \end{pmatrix}$$

Ou, de manière équivalente, on estime avec la méthode des MCG le système dans lequel la variable explicative P est remplacée par la variable \widehat{P}_{FRM} , c'est-à-dire le système $\widetilde{Y} = \widehat{X}\delta + \varepsilon$, d'où

$$\widehat{\delta}_{TMC} = \left(\widehat{X}' \Sigma^{-1} \widehat{X} \right)^{-1} \left(\widehat{X}' \Sigma^{-1} \widetilde{Y} \right) \tag{9.47}$$

La matrice de variance-covariance asymptotique de l'estimateur des TMC est donnée par

$$Asym.Var \left(\widehat{\delta}_{TMC} \right) = \frac{1}{N} \left(\frac{\widehat{X}' \Sigma^{-1} \widehat{X}}{N} \right)^{-1} \tag{9.48}$$

En pratique, la matrice de variance-covariance des aléas Σ est généralement inconnue et on procède alors de la manière suivante

1. On estime par les DMC chaque équation du système. On en déduit $\widehat{\sigma}_1^2$, $\widehat{\sigma}_2^2$ et $\widehat{\sigma}_{12}$ afin d'estimer la matrice de variance covariance des aléas $\widehat{\Sigma}$

$$\widehat{\sigma}_1^2 = \frac{\sum_{i=1}^N e_{1i}^2}{N} \quad \widehat{\sigma}_2^2 = \frac{\sum_{i=1}^N e_{2i}^2}{N} \quad \widehat{\sigma}_{12} = \frac{\sum_{i=1}^N e_{1i} \cdot e_{2i}}{N}$$

2. On estime les paramètres du modèle avec

$$\widehat{\delta}_{TMC} = \left(\widehat{X}' \widehat{\Sigma}^{-1} \widehat{X} \right)^{-1} \left(\widehat{X}' \widehat{\Sigma}^{-1} \widetilde{Y} \right)$$

3. On peut poursuivre et appliquer les TMC itératifs. Dans la deuxième étape on réestime la matrice Σ afin d'obtenir une nouvelle estimation des paramètres. On s'arrête lorsqu'entre deux itérations les paramètres estimés ne varient que très faiblement.

Lorsque les covariances des aléas interéquations sont nulles l'estimateur des DMC et des TMC donnent des résultats identiques. Sinon, les TMC donnent des estimateurs plus précis. Cependant avec une telle méthode la moindre erreur de spécification se diffuse dans tout le modèle. Il faut donc être prudent lorsqu'on utilise cette méthode.

Les méthodes DMC et TMC sont des méthodes des variables instrumentales. En appliquant ces méthodes on ne prend pas en compte des problèmes éventuels d'hétéroscédasticité ou d'autocorrélation des aléas des équations.

Comme dans le cas du modèle de régression linéaire généralisé on peut appliquer la méthode des moments généralisés. Dans un modèle à équations simultanées, elle peut s'appliquer comme les DMC, équation par équation, ou comme les TMC, à l'ensemble du système.

5. TESTS DE SPÉCIFICATION

Deux tests de spécification sont présentés : un test lié aux contraintes de sur-identification et un test d'exogénéité des variables explicatives.

5.1 Test sur les contraintes de sur-identification

Dans un modèle dont les équations sont exactement identifiées, le système d'équations obtenu à partir de la relation $B\Pi = -\Gamma$ admet une seule solution (exemple 9.2). Dans un modèle sur-identifié, le système donne plusieurs solutions pour les paramètres des matrices B et Γ (exemple 9.3).

Le test sur les contraintes de sur-identification est un test du ratio de vraisemblance. Si les aléas de la FRM sont distribués selon une loi normale, le logarithme de la fonction de vraisemblance de la FRM (9.13) s'écrit

$$\begin{aligned} \text{Log}L &= -\frac{GN}{2} \text{Log}2\pi - \frac{N}{2} \text{Log}|\Phi| - \frac{1}{2} \sum_{i=1}^N V_i' \Phi^{-1} V_i \\ &= -\frac{GN}{2} \text{Log}2\pi - \frac{N}{2} \text{Log}|\Phi| - \frac{1}{2} \sum_{i=1}^N (Y_i - \Pi X_i)' \Phi^{-1} (Y_i - \Pi X_i) \end{aligned} \quad (9.49)$$

avec

$$\Phi = E(V_i V_i') = \begin{pmatrix} E(v_{1i}^2) & E(v_{1i} v_{2i}) \\ E(v_{2i} v_{1i}) & E(v_{2i}^2) \end{pmatrix}$$

Le test du ratio de vraisemblance consiste à calculer la statistique

$$LR = -2 \left(\text{Log}L_{\hat{\Pi}_{TMC}} - \text{Log}L_{\hat{\Pi}_{MCO}} \right) \sim \chi^2(d) \tag{9.50}$$

avec $\text{Log}L_{\hat{\Pi}_{TMC}}$ la valeur de la fonction de vraisemblance pour une valeur des paramètres égale à $\hat{\Pi}_{TMC}$, $\text{Log}L_{\hat{\Pi}_{MCO}}$ la valeur de la fonction de vraisemblance pour une valeur des paramètres égale à $\hat{\Pi}_{MCO}$ et d le nombre de contraintes de sur-identification. Pour obtenir $\hat{\Pi}_{TMC}$ il faut, dans un premier temps, estimer le modèle avec la méthode des TMC et, dans un second temps, calculer $-\hat{B}_{TMC}^{-1} \hat{\Gamma}_{TMC}$. L'estimateur $\hat{\Pi}_{TMC}$ est l'estimateur obtenu lorsque les contraintes de sur-identification sont imposées, tandis que $\hat{\Pi}_{MCO}$ est l'estimateur sans contrainte.

Si $LR < \chi^2_{\alpha}(d)$ alors on accepte les contraintes de sur-identification sur les paramètres de la FSM.

5.2 Test d'exogénéité des variables explicatives : le test d'Hausman

Dans le modèle, les variables explicatives supposées exogènes sont utilisées pour construire la matrice des variables instrumentales des estimateurs des DMC et des TMC. Ces estimateurs sont convergents sauf lorsqu'on suppose de manière erronée qu'une variable explicative est exogène. La question de l'exogénéité d'une variable est donc très importante car, en cas d'erreur, les estimateurs des DMC et des TMC sont biaisés et non convergents.

Par exemple dans le modèle d'offre et de demande

$$\begin{aligned} Q_i &= \gamma_{11} + \beta_{11}P_i + \beta_{12}Q_{i-1} + \gamma_{21}R_i + \gamma_{31}Ps_i + \varepsilon_{1i} \\ Q_i &= \gamma_{12} + \beta_{12}P_i + \gamma_{22}Pf_i + \varepsilon_{2i} \end{aligned} \tag{9.51}$$

on suppose que les variables endogènes sont les prix P et les quantités Q , que les variables exogènes sont R , Ps , Pf et que la variable endogène décalée Q_{i-1} est prédéterminée, c'est-à-dire que les termes $E(Q_{i-1}, \varepsilon_{1i'})$ et $E(Q_{i-1}, \varepsilon_{2i'})$ sont nuls pour tout $i' \geq i^6$.

Dans le cadre de ce modèle, on souhaite tester l'hypothèse

$$\begin{aligned} H_0 : E(Q_{i-1}, \varepsilon_{1i}) &= 0 \\ H_1 : E(Q_{i-1}, \varepsilon_{1i}) &\neq 0 \end{aligned} \tag{9.52}$$

⁶ Une variable prédéterminée est exogène $cov(Q_{i-1}, \varepsilon_{ji}) = 0 \quad j=1,2$.

La version du test d'Hausman développée par Spencer et Berk passe par les étapes suivantes

1. Appliquer l'estimateur des DMC à l'équation de demande sous l'hypothèse H_0 . Sous cette hypothèse, on suppose que la variable Q_{i-1} est une variable prédéterminée et la FRM s'écrit

$$\underbrace{\begin{pmatrix} Q_i \\ P_i \end{pmatrix}}_{Y_i} = \underbrace{\begin{pmatrix} \pi_{11} & \pi_{21} & \pi_{31} & \pi_{41} & \pi_{51} \\ \pi_{12} & \pi_{22} & \pi_{32} & \pi_{42} & \pi_{52} \end{pmatrix}}_{\Pi} \underbrace{\begin{pmatrix} (1) \\ R_i \\ P_{s_i} \\ P_{f_i} \\ Q_{i-1} \end{pmatrix}}_{X_i} + \underbrace{\begin{pmatrix} v_{1i} \\ v_{2i} \end{pmatrix}}_{V_i} \quad (9.53)$$

On obtient l'estimateur des DMC en utilisant la variable \hat{P}_{FRM} comme variable instrumentale

$$\hat{P}_{FRM} = \hat{\pi}_{12} + \hat{\pi}_{22}R + \hat{\pi}_{32}Ps + \hat{\pi}_{42}Pf + \hat{\pi}_{52}Q_{-1} \quad (9.54)$$

et

$$\hat{\delta}_{DMC,1} = \left(\hat{X}^{d'} X^d \right)^{-1} \left(\hat{X}^{d'} Y^d \right) \quad (9.55)$$

avec

$$X^d = \begin{pmatrix} 1 & P_1 & - & R_1 & P_{s1} \\ 1 & P_2 & Q_1 & R_2 & P_{s2} \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ 1 & P_N & Q_N & R_N & P_{sN} \end{pmatrix} \text{ et } \hat{X}^d = \begin{pmatrix} 1 & \hat{P}_{FRM1} & - & R_1 & P_{s1} \\ 1 & \hat{P}_{FRM2} & Q_1 & R_2 & P_{s2} \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ 1 & \hat{P}_{FRMN} & Q_N & R_N & P_{sN} \end{pmatrix} \quad (9.56)$$

$\hat{\delta}_{DMC,1}$ est convergent sous H_0 .

2. Estimer avec l'estimateur des Variables Instrumentales les paramètres du modèle, en traitant la variable Q_{i-1} comme endogène. Dans ce cas la FRM s'écrit

$$\begin{pmatrix} Q_i \\ P_i \\ Q_{i-1} \end{pmatrix} = \begin{pmatrix} \pi_{11} & \pi_{21} & \pi_{31} & \pi_{41} \\ \pi_{12} & \pi_{22} & \pi_{32} & \pi_{42} \\ \pi_{13} & \pi_{23} & \pi_{33} & \pi_{43} \end{pmatrix} \begin{pmatrix} (1) \\ R_i \\ P_{s_i} \\ P_{f_i} \end{pmatrix} + \begin{pmatrix} v_{1i} \\ v_{2i} \\ v_{3i} \end{pmatrix} \quad (9.57)$$

et l'estimation avec la méthode des MCO de (9.57) donne: $\hat{P}_{FRM,i}$ et $\hat{Q}_{FRM,i-1}$. Ces deux variables sont alors utilisées comme variables instrumentales pour estimer l'équation de demande

$$\hat{\delta}_{VI,1} = \left(\hat{X}^{dv'} X^d \right)^{-1} \left(\hat{X}^{dv'} Y^d \right) \quad (9.58)$$

avec

$$X^d = \begin{pmatrix} 1 & P_1 & - & R_1 & P_{S1} \\ 1 & P_2 & Q_1 & R_2 & P_{S2} \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ 1 & P_N & Q_N & R_N & P_{SN} \end{pmatrix}$$

et

$$\widehat{X}^{dv} = \begin{pmatrix} 1 & \widehat{P}_{FRM1} & - & R_1 & P_{S1} \\ 1 & \widehat{P}_{FRM2} & \widehat{Q}_{FRM1} & R_2 & P_{S2} \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ 1 & \widehat{P}_{FRMN} & \widehat{Q}_{FRMN} & R_N & P_{SN} \end{pmatrix} \tag{9.59}$$

On obtient ainsi un estimateur des Variables Instrumentales des paramètres de la fonction de demande noté $\widehat{\delta}_{VI,1}$.

3. Sous l'hypothèse H_0 les deux estimateurs sont convergents mais $\widehat{\delta}_{DMC,1}$ est plus précis que $\widehat{\delta}_{VI,1}$, par contre si l'hypothèse H_0 est fausse alors $\widehat{\delta}_{DMC,1}$ est biaisé et non convergent tandis que $\widehat{\delta}_{VI,1}$ est convergent. On calcule alors la statistique de Wald

$$W = \left(\widehat{\delta}_{DMC,1} - \widehat{\delta}_{VI,1} \right)' \left(Var \left(\widehat{\delta}_{VI,1} \right) - Var \left(\widehat{\delta}_{DMC,1} \right) \right)^{-1} \left(\widehat{\delta}_{DMC,1} - \widehat{\delta}_{VI,1} \right) \\ \sim \chi^2(1) \tag{9.60}$$

si la statistique W est inférieure au $\chi^2_\alpha(1)$ on accepte l'hypothèse H_0 ⁷. L'écart entre les deux estimateurs est non significativement différent de 0, ils sont donc tous les deux convergents et on choisira l'estimateur des DMC car il est plus précis que l'estimateur $\widehat{\delta}_{VI,1}$. Le degré de liberté du test est égal à 1 car on teste l'exogénéité d'une seule variable. L'extension du test à plusieurs variables est immédiate.

⁷ Dans la pratique on calcule la statistique de Wald ou de manière plus simple, mais équivalente, on procède à un test de significativité du paramètre δ dans l'équation suivante

$$Q_i = \gamma_{11} + \beta_{11} \widehat{P}_{FRMi} + \beta_{12} Q_{i-1} + \gamma_{21} R_i + \gamma_{31} P_{Si} + \delta \widehat{v}_{3i} + \varepsilon_{1i}$$

où \widehat{v}_3 est le vecteur de résidu obtenu lorsqu'on estime avec la méthode des MCO la troisième équation de (9.57).

10

MODÉLISATION D'UN SYSTÈME DE DEMANDE DE FACTEURS DE PRODUCTION

1. Introduction
2. Approche par la fonction de coût
3. Illustration en spécifiant une fonction translog
4. Estimation de la fonction translog
5. Analyse empirique
6. Éléments de corrigé

OBJECTIFS DU CHAPITRE

D'un point de vue méthodologique

- Présenter un système d'équations apparemment non reliées
- Appliquer la méthode de Zellner (SURE) simple et itérative

D'un point de vue économique

- Appliquer la théorie de la dualité
- Analyser les substitutions entre les facteurs de production
- Calculer l'élasticité partielle de substitution d'Allen

NIVEAU

MASTER, première et deuxième années

CHAMPS ÉCONOMIQUES TRAITÉS

Microéconomie

Théorie de la dualité

1. INTRODUCTION

L'objet de ce chapitre est l'étude des relations de substitution-complémentarité entre les facteurs de production pour une branche de l'agro-alimentaire français. En particulier, on cherche à mesurer l'influence de l'évolution des prix relatifs des facteurs de production sur l'utilisation respective des inputs.

La réponse à cette question nécessite de définir un système d'équations spécifiant les niveaux des facteurs de production en fonction des variables considérées comme exogènes, notamment les prix des facteurs.

Pour dériver les fonctions de demandes de facteurs de production on peut utiliser la démarche décrite dans le chapitre sur la fonction d'offre (cf. chapitre 8, La fonction d'offre de viande bovine) : maximiser le profit sous la contrainte d'une technologie de production, les prix du produit offert et des facteurs de production étant donnés. La résolution de ce programme structurel détermine un système formé par la fonction d'offre et les fonctions de demande de facteurs de production. D'un point de vue économétrique, les niveaux de facteurs de production et le niveau de production sont endogènes. Les prix sont exogènes. Pour dériver analytiquement les demandes d'inputs, il suffit de choisir une forme fonctionnelle pour la fonction de production.

Une autre approche possible repose sur l'utilisation d'une fonction de coût. En appliquant les résultats de la dualité, on peut aisément dériver un système de demande de facteurs. Dans ce cadre d'analyse, les quantités de facteurs de production sont endogènes, le niveau de production et les prix des facteurs sont exogènes et, pour dériver analytiquement les demandes d'inputs, il suffit de choisir une forme fonctionnelle pour la fonction de coût.

Dans ce chapitre, on privilégie cette approche. On présente, dans la seconde section, les propriétés de la fonction de coût et la dérivation des fonctions de demande de facteurs. Puis, la troisième section présente une illustration basée sur une fonction de coût translog. La spécification économétrique et les principes d'estimation sont détaillés dans la quatrième section. La cinquième section est consacrée à l'analyse empirique.

2. APPROCHE PAR LA FONCTION DE COÛT

On rappelle, tout d'abord, la définition d'une fonction de coût et les propriétés fondamentales nécessaires de cette fonction. Le second paragraphe détaille certaines propriétés particulières, non nécessaires, d'une fonction de coût. Le troisième paragraphe présente le calcul des élasticités de substitution d'Allen qui permettent d'analyser les relations de substitution entre facteurs de production

2.1 Définition et propriétés fondamentales d'une fonction de coût

La fonction de coût mesure le coût minimal de production d'un niveau déterminé d'output pour des prix des facteurs donnés. De manière analytique, cette fonction est déduite du programme suivant

$$C(y, w) = \min_{x_i} \sum_{i=1}^n w_i x_i \quad (10.1)$$

sous la contrainte

$$y = f(x_1, \dots, x_n)$$

où C représente la fonction de coût, y le niveau de production, w le vecteur prix des n inputs ($w = (w_1, \dots, w_i, \dots, w_n)$), x le vecteur des quantités utilisées des n facteurs de production ($x = (x_1, \dots, x_i, \dots, x_n)$) et f la fonction de production.

Une fonction de coût possède certaines propriétés qui découlent du problème structurel de minimisation du coût : la fonction de coût est non décroissante par rapport aux prix des facteurs, continue et concave en w et homogène de degré 1 par rapport à w . Avant de préciser les restrictions induites par ces différentes propriétés, on présente un résultat issu de la théorie de la dualité (propriété de dérivation de la fonction de coût).

2.1.1 Propriété issue de la théorie de la dualité

La théorie de la dualité fondée principalement par Hotelling¹ et Shepard² permet de montrer l'équivalence entre la représentation de la technologie d'une entreprise (définie par sa fonction de production) et sa fonction de coût. Autrement dit, la fonction de coût synthétise toutes les informations relatives à la technologie de production et inclut par définition une hypothèse comportementale (minimisation du coût).

De cette fonction de coût, il est simple de dériver les demandes conditionnelles de facteurs de production. En effet, d'après le lemme de Shepard, si la fonction de coût est différentiable en (y, w) alors

$$\frac{\partial C(y, w)}{\partial w_i} = x_i(y, w) \quad \text{si } w_i > 0 \quad \forall i = 1, \dots, n \quad (10.2)$$

où $x_i(y, w)$ définit la demande conditionnelle de facteur i de l'entreprise³, ou demande dérivée de facteur. Cette demande représente la quantité d'input nécessaire pour produire l'output y en minimisant le coût, pour des prix des inputs donnés.

¹ Hotelling (1932), voir Diewert, W. E, (1971) *Applications of duality theory*, in *Frontiers of quantitative Economics*, vol. II M. Intriligator and D.A. Kendrick, eds. (Amsterdam North-Holland) 106-206.

² Shepard, R. (1953), *Cost and production functions*, Princeton, NJ, Princeton University Press.

³ Conditionnelle au fait que la firme produise un niveau donné d'output.

Par ailleurs, la fonction de coût vérifie différentes propriétés qui vont impliquer des restrictions au niveau de ses dérivées et des fonctions de demande de facteurs. Les propriétés énumérées précédemment sont successivement présentées.

2.1.2 La contrainte de monotonie et la contrainte de concavité

La fonction de coût est non décroissante par rapport aux prix des facteurs c'est-à-dire

$$\frac{\partial C(y, w)}{\partial w_i} = x_i(y, w) \geq 0 \quad (10.3)$$

La fonction de coût est concave par rapport à w . La matrice des dérivées secondes de la fonction de coût (c'est-à-dire la matrice des dérivées premières des fonctions de demande de facteurs) est une matrice symétrique semi-définie négative. Cette propriété entraîne des restrictions sur les dérivées de $C(y, w)$.

- Les effets prix croisés sont symétriques (symétrie de la matrice des dérivées secondes).

$$\frac{\partial^2 C(y, w)}{\partial w_i \partial w_j} = \frac{\partial^2 C(y, w)}{\partial w_j \partial w_i} = \frac{\partial x_i(y, w)}{\partial w_j} = \frac{\partial x_j(y, w)}{\partial w_i} \quad \forall i, j = 1, \dots, n \quad (10.4)$$

- Les effets prix propres sont non positifs (les termes diagonaux de la matrice des dérivées secondes doivent être non positifs).

$$\frac{\partial^2 C(y, w)}{\partial w_i^2} = \frac{\partial x_i(y, w)}{\partial w_i} \leq 0 \quad \forall i = 1, \dots, n \quad (10.5)$$

2.1.3 La contrainte d'homogénéité de degré 1 par rapport aux prix

La fonction de coût doit être homogène de degré 1 par rapport aux prix. Cette propriété signifie que si tous les prix des facteurs de production sont multipliés par un scalaire λ , alors le coût total pour produire un niveau de production identique y est aussi multiplié par ce coefficient λ .

$$C(y, \lambda w_1, \dots, \lambda w_n) = \lambda C(y, w_1, \dots, w_n) \quad (10.6)$$

Cette propriété implique que les demandes de facteurs sont homogènes de degré zéro par rapport à w^4 .

⁴ Suivant le théorème d'Euler, quand une fonction est homogène de degré k , sa dérivée première est homogène de degré $k-1$.

2.2 Propriétés supplémentaires

À ces différentes propriétés nécessaires pour définir une fonction de coût, des propriétés supplémentaires peuvent être imposées pour réduire le nombre de paramètres à estimer : la propriété d'homothétie et la propriété de rendements d'échelle constants. Ces propriétés sont successivement examinées.

2.2.1 Hypothèse d'homothétie de la fonction de coût

La fonction de coût homothétique s'écrit comme une fonction séparable de l'output et des prix des facteurs de production et vérifie donc l'égalité suivante

$$C(y, w) = h(y)g(w) \quad (10.7)$$

où h est une fonction qui dépend uniquement du niveau de production et g est une fonction qui dépend uniquement des prix des inputs.

Si l'on applique le lemme de Shepard (équation 10.2) à l'égalité précédente, la demande conditionnelle de facteur i s'écrit

$$x_i(y, w) = h(y) \frac{\partial g(w)}{\partial w_i} \quad (10.8)$$

la demande conditionnelle de facteur j s'écrit

$$x_j(y, w) = h(y) \frac{\partial g(w)}{\partial w_j}$$

Ainsi, le rapport entre deux demandes de facteurs de production ne dépend pas du niveau de production (le sentier d'expansion est linéaire).

2.2.2 Hypothèse de rendements d'échelle constants

Un cas particulier de la propriété précédente apparaît lorsque $h(y) = y$. Sous cette hypothèse, la fonction de coût s'écrit

$$C(y, w) = yg(w) \quad (10.9)$$

Dans ce cadre d'analyse, la fonction de production admet des rendements d'échelle constants. En particulier, les fonctions de coût marginal et de coût moyen sont identiques et indépendantes du niveau de production. En effet

$$\frac{C(y, w)}{y} = \frac{\partial C(y, w)}{\partial y} = g(w) \quad (10.10)$$

Dans ce cas, la fonction $g(w)$ s'interprète comme le coût unitaire de production (qui est constant quel que soit le niveau de production) c'est-à-dire $g(w) =$

$C(1, w)$. De plus, si on applique le lemme de Shepard (équation 10.2) on déduit la demande du i ème facteur

$$x_i(y, w) = y \frac{\partial g(w)}{\partial w_i} \tag{10.11}$$

$\frac{\partial g(w)}{\partial w_i}$ représente la demande de facteur par unité de production.

Pour traduire cette propriété, il est équivalent de dire que la fonction de coût est homogène de degré 1 par rapport à la production, c'est-à-dire qu'elle vérifie l'égalité suivante

$$C(\lambda y, w) = \lambda C(y, w_1, \dots, w_n) \tag{10.12}$$

Si le niveau de production est multiplié par λ alors le coût est aussi multiplié par λ .

2.3 Interprétation des résultats

L'objectif de ce chapitre est d'analyser la substitution entre différents facteurs de production. Une notion couramment utilisée est l'élasticité partielle de substitution d'Allen. L'élasticité partielle substitution entre deux facteurs de production i et j s'écrit

$$\sigma_{ij} = \frac{C C_{ij}}{C_i C_j} \quad i \neq j \tag{10.13}$$

où C représente la fonction de coût, C_{ij} sa dérivée seconde par rapport aux prix des facteurs de production i et j ($C_{ij} = \partial^2 C / \partial w_i \partial w_j$), C_i et C_j ses dérivées premières ($C_i = \partial C / \partial w_i$, $C_j = \partial C / \partial w_j$). Ces élasticités sont symétriques. Le signe de cette élasticité éclaire sur les relations de substitution entre les facteurs de production.

- Si σ_{ij} est négatif, les deux facteurs i et j sont complémentaires.
- Si σ_{ij} est positif, les deux facteurs i et j sont substituables.
- Si σ_{ij} est égal à zéro, les facteurs sont parfaitement complémentaires.
- Si σ_{ij} tend vers l'infini, les facteurs sont parfaitement substituables.

Cette élasticité peut être reliée à l'élasticité partielle de la demande du facteur i par rapport au prix du facteur j le niveau de production et les prix des autres facteurs étant donnés. Cette élasticité est notée η_{ij}

$$\eta_{ij} = \frac{\partial \ln x_i}{\partial \ln w_j} \tag{10.14}$$

Cette élasticité peut se réécrire un utilisant le lemme de Shepard

$$\begin{aligned} \eta_{ij} &= \frac{\partial x_i}{\partial w_j} \frac{w_j}{x_i} \\ &= C_{ij} \frac{w_j}{C_i} \end{aligned}$$

En divisant cette équation par la part de la dépense en facteur j dans le coût total ($S_j = w_j x_j / C$), on fait apparaître l'élasticité partielle de substitution d'Allen

$$\begin{aligned}\eta_{ij} &= C_{ij} \frac{w_j}{C_i} \frac{C}{C_j} \frac{C_j}{C} \\ &= \frac{C C_{ij}}{C_i C_j} \frac{w_j C_j}{C} \\ &= \sigma_{ij} S_j\end{aligned}$$

L'élasticité partielle entre deux facteurs de production i et j s'écrit

$$\sigma_{ij} = \frac{\eta_{ij}}{S_j} \quad (10.15)$$

Dans ce chapitre, on choisit une forme fonctionnelle pour la fonction de coût et on déduit par dérivation les demandes conditionnelles de facteurs. L'approche par une fonction de coût présente plusieurs avantages. Tout d'abord, elle permet une représentation plus simple d'une technologie multi-facteurs (les calculs analytiques sont plus aisés). De plus, d'un point de vue économétrique les variables prix qui entrent dans la spécification de la fonction de coût sont considérées comme exogènes dans le comportement de l'entreprise (contrairement aux niveaux des facteurs de production dans la fonction de production).

La forme fonctionnelle la plus simple est la fonction Cobb-Douglas, mais la limite majeure de cette forme fonctionnelle est d'imposer que toutes les élasticités de substitution sont constantes et égales à 1.

Aussi, on utilise une forme fonctionnelle dite flexible c'est-à-dire qui ne pose aucune hypothèse a priori sur les relations de substitution entre les facteurs. On travaille à partir d'une fonction de coût de type translog.

3. ILLUSTRATION EN SPÉCIFIANT UNE FONCTION TRANSLOG

De manière générale, une fonction translog ayant comme arguments y (le niveau de production) et w le vecteur prix des n facteurs de production ($w = (w_1, \dots, w_i, \dots, w_n)$) s'écrit

$$\begin{aligned}\ln C(y, w) &= \ln \alpha_0 + \sum_{i=1}^n \alpha_i \ln w_i + 1/2 \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij} \ln w_i \ln w_j + \\ &\quad \alpha_y \ln y + 1/2 \alpha_{yy} (\ln y)^2 + \sum_{i=1}^n \alpha_{iy} \ln w_i \ln y\end{aligned} \quad (10.16)$$

où $\alpha_0, \alpha_i, \alpha_{ij}, \alpha_y, \alpha_{yy}$ et α_{iy} sont les paramètres à estimer. Cette équation comporte donc $3 + n^2 + 2n$ paramètres à estimer.

3.1 Propriétés de la fonction de coût

Pour que cette fonction définisse une fonction de coût, elle doit respecter les différentes propriétés fondamentales précédemment décrites. La prise en compte de ces propriétés impose des contraintes sur les coefficients à estimer.

3.1.1 Les contraintes de symétrie

Le respect de l'égalité (10.4) implique la symétrie de certains coefficients, c'est-à-dire implique l'égalité suivante $\alpha_{ij} = \alpha_{ji}$. Il y a $n(n-1)/2$ contraintes de symétrie.

3.1.2 La contrainte d'homogénéité de degré 1 par rapport aux prix

Écrite sous forme logarithmique, la condition (10.6) devient

$$\ln C(y, \lambda w_1, \dots, \lambda w_n) = \ln \lambda + \ln C(y, w_1, \dots, w_n) \quad (10.17)$$

L'égalité précédente est imposée sur la fonction translog (10.16). Le respect de cette propriété détermine des contraintes sur les coefficients. On doit avoir

$$\begin{aligned} \ln \alpha_0 + \sum_{i=1}^n \alpha_i \ln \lambda w_i + 1/2 \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij} \ln \lambda w_i \ln \lambda w_j + \\ \alpha_y \ln y + 1/2 \alpha_{yy} (\ln y)^2 + \sum_{i=1}^n \alpha_{iy} \ln \lambda w_i \ln y \\ = \ln \lambda + \ln C(y, w_1, \dots, w_n) \end{aligned} \quad (10.18)$$

En développant l'expression à gauche

$$\begin{aligned} \ln \alpha_0 + \sum_{i=1}^n \alpha_i \ln \lambda + \sum_{i=1}^n \alpha_i \ln w_i + 1/2 \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij} (\ln \lambda + \ln w_i) (\ln \lambda + \ln w_j) \\ + \alpha_y \ln y + 1/2 \alpha_{yy} (\ln y)^2 + \sum_{i=1}^n \alpha_{iy} \ln w_i \ln y + \sum_{i=1}^n \alpha_{iy} \ln \lambda \ln y \\ = \ln \lambda + \ln C(y, w_1, \dots, w_n) \end{aligned} \quad (10.19)$$

À gauche, on retrouve l'expression de la fonction translog et des termes additionnels

$$\begin{aligned} \ln C(y, w) + \sum_{i=1}^n \alpha_i \ln \lambda + \ln \lambda \sum_{i=1}^n \alpha_{iy} \ln y + \\ 1/2 \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij} (\ln \lambda \ln \lambda + \ln \lambda \ln w_j + \ln w_i \ln \lambda) \\ = \ln \lambda + \ln C(w_1, \dots, w_n, y) \end{aligned} \quad (10.20)$$

La double somme dans l'expression à gauche est développée

$$\begin{aligned} \ln C(y, w) + \sum_{i=1}^n \alpha_i \ln \lambda + \ln \lambda \sum_{i=1}^n \alpha_{iy} \ln y \\ + (\ln \lambda \ln \lambda) 1/2 \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij} + \ln \lambda \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij} \ln w_i \\ = \ln \lambda + \ln C(y, w) \end{aligned} \quad (10.21)$$

Pour que l'égalité précédente soit vérifiée, les contraintes suivantes doivent être imposées sur les paramètres

$$\begin{aligned} \sum_{i=1}^n \alpha_i &= 1 \\ \sum_{i=1}^n \alpha_{iy} &= 0 \\ \sum_{i=1}^n \alpha_{ij} &= 0 \quad \forall j = 1, \dots, n \\ \sum_{j=1}^n \alpha_{ij} &= 0 \quad \forall i = 1, \dots, n \end{aligned}$$

3.2 Propriétés particulières

Des propriétés particulières peuvent être imposées pour réduire le nombre de paramètres à estimer : la propriété d'homothétie et la propriété de rendements d'échelle constants. Ces propriétés sont successivement examinées dans le cas de la fonction de coût Translog.

3.2.1 Hypothèse d'homothétie de la fonction de coût

Sous forme logarithmique, cette propriété s'écrit : $\ln C(y, w) = \ln h(y) + \ln g(w)$

Aussi, la fonction de coût est homothétique si

$$\frac{\partial^2 \ln C}{\partial \ln w_i \partial \ln y} = 0 \quad (10.22)$$

Dans le cas translog $\frac{\partial^2 \ln C}{\partial \ln w_i \partial \ln y} = \alpha_{iy}$

La propriété d'homothétie est donc vérifiée si $\alpha_{iy} = 0 \forall i = 1, \dots, n$.

3.2.2 Hypothèse de rendements d'échelle constants

Un cas particulier de la propriété précédente apparaît lorsque $h(y) = y$. Dans ce cas, on a la propriété de rendements d'échelle constants et on a l'égalité suivante

$$\ln C(y, w) = \ln y + \ln C(1, w) \quad (10.23)$$

Cette relation implique que

$$\frac{\partial \ln C}{\partial \ln y} = 1 \quad (10.24)$$

Dans le cas translog

$$\frac{\partial \ln C}{\partial \ln y} = \alpha_y + \alpha_{yy} \ln y + \sum_{i=1}^n \alpha_{iy} \ln w_i \quad (10.25)$$

Si on veut que la condition (10.24) soit vérifiée il faut que $\alpha_y = 1$, $\alpha_{yy} = 0$ et $\alpha_{iy} = 0 \forall i = 1, \dots, n$ (la dernière condition étant déjà assurée avec la propriété d'homothétie⁵).

4. ESTIMATION DE LA FONCTION TRANSLOG

Cette section présente les équations qui vont être estimées dans la partie empirique.

4.1 Dérivation du système de parts de dépense

Pour dériver l'équation de demande du i ème facteur de production, il faut dériver la fonction de coût par rapport au prix du i ème facteur de production. En effet, par application du lemme de Shepard, on a

⁵ De manière équivalente, on peut retrouver ces trois contraintes en cherchant les restrictions imposées par la propriété d'homogénéité de degré 1 de la fonction de coût par rapport à y , c'est-à-dire en étudiant les restrictions impliquées par l'égalité suivante $\ln C(\lambda y, w) = \ln \lambda + \ln C(y, w)$.

$$\frac{\partial C(y, w)}{\partial w_i} = x_i(y, w) \quad (10.26)$$

Dans le cas de la fonction de coût translog on a facilement $\frac{\partial \ln C}{\partial \ln w_i}$

$$\frac{\partial \ln C}{\partial \ln w_i} = \alpha_i + \sum_{j=1}^n \alpha_{ij} \ln w_j + \alpha_{iy} \ln y \quad (10.27)$$

Par définition,

$$\frac{\partial \ln C}{\partial \ln w_i} = \frac{\partial C}{\partial w_i} \frac{w_i}{C} = \frac{x_i w_i}{C} \quad (10.28)$$

avec $\sum_{i=1}^n w_i x_i = C$

À partir du lemme de Shepard, on détermine la part optimale de dépense notée S_i , de chaque facteur de production dans le coût total de production. On obtient un système de n équations

$$S_i = \frac{w_i x_i}{C} = \alpha_i + \sum_{j=1}^n \alpha_{ij} \ln w_j + \alpha_{iy} \ln y \quad i = 1, \dots, n \quad (10.29)$$

avec $\sum_{i=1}^n S_i = 1$.

Par construction, la somme des parts est égale à un

$$\sum_{i=1}^n S_i = \sum_{i=1}^n (\alpha_i + \sum_{j=1}^n \alpha_{ij} \ln w_j + \alpha_{iy} \ln y) \quad (10.30)$$

$$= \sum_{i=1}^n \alpha_i + \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij} \ln w_j + \sum_{i=1}^n \alpha_{iy} \ln y \quad (10.31)$$

$$= 1 \quad \forall i \quad (10.32)$$

Cette condition implique des restrictions sur les paramètres à estimer

$$\sum_{i=1}^n \alpha_i = 1 \quad ; \quad \sum_{i=1}^n \alpha_{iy} = 0 \quad ; \quad \sum_{i=1}^n \alpha_{ij} = 0 \quad \forall j = 1, \dots, n$$

Ces restrictions définissent les contraintes d'additivité.

De manière synthétique, les tableaux suivants résument les principales restrictions sur les paramètres à estimer.

Tableau 10.1 Propriétés fondamentales et contrainte d'additivité

Contraintes	Symétrie	Homogénéité par rapport aux prix	Additivité
$\alpha_{ij} = \alpha_{ji}$	à imposer		
$\sum_{i=1}^n \alpha_i = 1$		à imposer	à imposer
$\sum_{i=1}^n \alpha_{iy} = 0$		à imposer	à imposer
$\sum_{i=1}^n \alpha_{ij} = 0$		à imposer	à imposer
$\sum_{j=1}^n \alpha_{ij} = 0$		à imposer	

Ce tableau montre que si les restrictions liées aux contraintes d'additivité et les restrictions de symétrie sont imposées alors les restrictions liées aux contraintes d'homogénéité par rapport aux prix des facteurs sont vérifiées. De même, si les conditions d'homogénéité par rapport aux prix sont imposées alors les conditions d'additivité sont respectées.

Tableau 10.2 Propriétés particulières

Contraintes	Homothétie	Rendements d'échelle constants
$\alpha_{iy} = 0$	X	X
$\alpha_y = 1$		X
$\alpha_{yy} = 0$		X

Si les restrictions liées aux contraintes d'homogénéité par rapport aux prix des facteurs sont vérifiées, alors il n'y a que deux conditions à rajouter pour avoir la propriété de rendements d'échelle constants⁶.

4.2 Illustration dans le cas de trois facteurs

On suppose qu'il existe trois facteurs de production : le travail (L), le capital (K) et les consommations intermédiaires (CI).

⁶ Ou de manière équivalente la propriété d'homogénéité de degré 1 par rapport à la production.

De manière générale (sans imposer les restrictions déduites des propriétés précédemment décrites), les équations de parts s'écrivent

$$Sk = \alpha_K + \alpha_{KL} \ln w_L + \alpha_{KK} \ln w_K + \alpha_{KCI} \ln w_{CI} + \alpha_{Ky} \ln y \quad (10.33)$$

$$Sl = \alpha_L + \alpha_{LL} \ln w_L + \alpha_{LK} \ln w_K + \alpha_{LCI} \ln w_{CI} + \alpha_{Ly} \ln y \quad (10.34)$$

$$Sci = \alpha_{CI} + \alpha_{CIL} \ln w_L + \alpha_{CIK} \ln w_K + \alpha_{CICI} \ln w_{CI} + \alpha_{CIy} \ln y \quad (10.35)$$

A priori, le nombre de paramètres à estimer est égal à 15.

Pour que ce système soit dérivé d'une fonction de coût, les paramètres doivent vérifier les contraintes de symétrie et la contrainte d'homogénéité de degré 1 par rapport aux prix.

De plus, les paramètres sont contraints par la définition des variables endogènes (la somme des trois variables endogènes doit être égale à un).

Les conséquences de ces différentes restrictions sur les paramètres sont décrites ci-après.

4.2.1 Contraintes de symétrie

Les contraintes de symétrie sont $\alpha_{KL} = \alpha_{LK}$, $\alpha_{KCI} = \alpha_{CIK}$, $\alpha_{LCI} = \alpha_{CIL}$. Si l'on impose ces contraintes, le nouveau système à estimer s'écrit

$$Sk = \alpha_K + \alpha_{KL} \ln w_L + \alpha_{KK} \ln w_K + \alpha_{KCI} \ln w_{CI} + \alpha_{Ky} \ln y \quad (10.36)$$

$$Sl = \alpha_L + \alpha_{LL} \ln w_L + \alpha_{KL} \ln w_K + \alpha_{LCI} \ln w_{CI} + \alpha_{Ly} \ln y \quad (10.37)$$

$$Sci = \alpha_{CI} + \alpha_{LCI} \ln w_L + \alpha_{KCI} \ln w_K + \alpha_{CICI} \ln w_{CI} + \alpha_{CIy} \ln y \quad (10.38)$$

Le nombre de paramètres à estimer est égal à 12.

4.2.2 Contrainte d'homogénéité de degré 1 par rapport aux prix

Les quatre restrictions à respecter sont

$$\sum_{i=1}^n \alpha_i = 1$$

$$\sum_{i=1}^n \alpha_{iy} = 0$$

$$\sum_{i=1}^n \alpha_{ij} = 0 \quad \forall j = 1, \dots, n$$

$$\sum_{j=1}^n \alpha_{ij} = 0 \quad \forall i = 1, \dots, n$$

Du fait de l'imposition des contraintes de symétrie, la dernière restriction est redondante avec la troisième restriction.

Dans le cas de trois facteurs, les restrictions équivalent à

$$\begin{aligned}
 \alpha_K + \alpha_L + \alpha_{CI} &= 1 \\
 \alpha_{Ky} + \alpha_{Ly} + \alpha_{CIy} &= 0 \\
 \alpha_{KL} + \alpha_{KK} + \alpha_{KCI} &= 0 \\
 \alpha_{LL} + \alpha_{KL} + \alpha_{LCI} &= 0 \\
 \alpha_{LCI} + \alpha_{KCI} + \alpha_{CICI} &= 0
 \end{aligned}
 \tag{10.39}$$

Pour tenir compte de ces contraintes, on exprime les coefficients de l'équation de la part des consommations intermédiaires (équation (10.38)) en fonction des autres coefficients. On obtient

$$\begin{aligned}
 \alpha_{CI} &= 1 - \alpha_K - \alpha_L \\
 \alpha_{CIy} &= -\alpha_{Ky} - \alpha_{Ly} \\
 \alpha_{KCI} &= -\alpha_{KL} - \alpha_{KK} \\
 \alpha_{LCI} &= -\alpha_{LL} - \alpha_{KL} \\
 \alpha_{CICI} &= -\alpha_{LCI} - \alpha_{KCI}
 \end{aligned}
 \tag{10.40}$$

On introduit ces relations dans les équations (10.36), (10.37) et (10.38). On obtient le système suivant

$$\begin{aligned}
 Sk &= \alpha_K + \alpha_{KL} \ln w_L + \alpha_{KK} \ln w_K + (-\alpha_{KL} - \alpha_{KK}) \ln w_{CI} \\
 &\quad + \alpha_{Ky} \ln y
 \end{aligned}
 \tag{10.41}$$

$$\begin{aligned}
 Sl &= \alpha_L + \alpha_{LL} \ln w_L + \alpha_{KL} \ln w_K + (-\alpha_{LL} - \alpha_{KL}) \ln w_{CI} \\
 &\quad + \alpha_{Ly} \ln y
 \end{aligned}
 \tag{10.42}$$

$$\begin{aligned}
 Sci &= (1 - \alpha_K - \alpha_L) + (-\alpha_{LL} - \alpha_{KL}) \ln w_L \\
 &\quad + (-\alpha_{KL} - \alpha_{KK}) \ln w_K + (-\alpha_{LCI} - \alpha_{KCI}) \ln w_{CI} + (-\alpha_{Ky} - \alpha_{Ly}) \ln y
 \end{aligned}
 \tag{10.43}$$

Les paramètres de la dernière équation sont déduits de ceux des premières équations⁷.

⁷ Pour estimer les paramètres, des aléas sont introduits dans chaque équation. La contrainte d'additivité implique que la matrice de variance-covariance des aléas est singulière. Pour résoudre ce problème économétrique, on estime simplement deux équations.

Dans les deux premières équations, certains termes peuvent être regroupés

$$Sk = \alpha_K + \alpha_{KL}(\ln w_L - \ln w_{CI}) + \alpha_{KK}(\ln w_K - \ln w_{CI}) + \alpha_{Ky} \ln y \quad (10.44)$$

$$Sl = \alpha_L + \alpha_{LL}(\ln w_L - \ln w_{CI}) + \alpha_{KL}(\ln w_K - \ln w_{CI}) + \alpha_{Ly} \ln y \quad (10.45)$$

Ces équations peuvent se réécrire

$$Sk = \alpha_K + \alpha_{KL} \ln(w_L/w_{CI}) + \alpha_{KK} \ln(w_K/w_{CI}) + \alpha_{Ky} \ln y \quad (10.46)$$

$$Sl = \alpha_L + \alpha_{LL} \ln(w_L/w_{CI}) + \alpha_{KL} \ln(w_K/w_{CI}) + \alpha_{Ly} \ln y \quad (10.47)$$

Dans l'exemple, l'équation de la part des consommations intermédiaires est omise. On pourrait choisir d'omettre une autre équation (capital ou travail). Dans ce cas, la démarche est similaire, la spécification obtenue dépend de l'équation omise (voir le problème 10.2).

4.2.3 Contraintes d'additivité

Les contraintes d'homogénéité étant imposées, les contraintes d'additivité sont vérifiées (voir tableau 10.1).

4.2.4 Contrainte d'homothétie

Si on impose la contrainte d'homothétie, on a $\alpha_{iy} = 0 \forall i$. Dans ce cas, $\alpha_{Ky} = \alpha_{Ly} = \alpha_{CIy} = 0$. Le nouveau système à estimer s'écrit

$$Sk = \alpha_K + \alpha_{KL} \ln(w_L/w_{CI}) + \alpha_{KK} \ln(w_K/w_{CI}) \quad (10.48)$$

$$Sl = \alpha_L + \alpha_{LL} \ln(w_L/w_{CI}) + \alpha_{KL} \ln(w_K/w_{CI}) \quad (10.49)$$

4.2.5 Interprétation des coefficients : calcul des élasticités

Pour la fonction de coût translog, on peut montrer que l'élasticité de substitution d'Allen s'écrit

$$\sigma_{ij} = \frac{\alpha_{ij} + S_i S_j}{S_i S_j} \quad i \neq j, \quad i, j = K, CI, L \quad (10.50)$$

De plus, l'élasticité de la demande de facteur i par rapport au prix du facteur j s'exprime en fonction de l'élasticité partielle de substitution d'Allen

$$\eta_{ij} = S_j \sigma_{ij} \quad i, j = K, CI, L \quad (10.51)$$

4.2.6 Estimation des paramètres

Pour estimer les coefficients, des aléas ayant les propriétés classiques sont introduits dans chaque équation. Les paramètres sont estimés par la méthode de Zellner itérative.

5. ANALYSE EMPIRIQUE

Description des données

Les données sont des données annuelles françaises sur la période 1978 à 1993 pour une branche agro-alimentaire française : la branche de « la conserve de fruits ». Les données contiennent des informations sur la quantité de travail notée xL (effectifs salariés moyens), la quantité de capital notée xK (millions de francs constants 1980), la quantité de consommations intermédiaires notée xCI (milliers de tonnes), le coût du travail noté wL (francs de l'année courante), le coût du capital wK (francs de l'année courante) et le prix des consommations intermédiaires wCI (indice de prix en base 1980). On dispose également d'une information sur le niveau du prix de l'output noté pv (indice de prix en base 1980) et de la valeur courante des ventes d'output notée $vent$.

Problème 10.1

Construction des variables et analyse descriptive

1. Calculer le niveau de production. Transformer les variables prix des facteurs de production travail et capital de façon à retenir l'année 1980 comme année de référence.
2. Calculer pour chaque facteur de production sa part de dépense par rapport au coût total. Ces parts sont respectivement notées Sk , Sl , et Sci .
3. Analyser graphiquement les différentes variables.

Problème 10.2

Estimation du système de parts de dépense et interprétation des résultats

1. Écrire le système de parts de dépense dérivé d'une fonction de coût translog en imposant les contraintes liées à la fonction de coût et les contraintes d'additivité. Imposer de plus la contrainte d'homothétie. Écrire le système quand l'équation de la part des consommations intermédiaires est omise (système 1), quand l'équation de la part du capital est omise (système 2).
2. Estimer le système 1 par la méthode des MCO, équation par équation. Cette méthode est-elle judicieuse ?
3. Estimer les deux systèmes par la méthode de Zellner. Comparer les coefficients estimés.

4. Vérifier que l'estimation du modèle contraint est indépendante de l'équation omise du système si on utilise la méthode de Zellner itérative. Calculer tous les coefficients.
5. Déterminer les élasticités de substitution partielle d'Allen et les élasticités prix de la demande. Interpréter.

6. ÉLÉMENTS DE CORRIGÉ

Problème 10.1

Construction des variables et analyse descriptive

1. Pour obtenir des variables prix en base 1980, on divise chaque série de prix par sa valeur observée en 1980 et on la multiplie par 100. Les nouvelles séries de prix sont ainsi toutes égales à 100 en 1980.
2. On dispose dans la base de données des informations sur les prix des facteurs et les niveaux des facteurs de production. Pour chaque facteur, on calcule la part du coût du facteur considéré dans le coût total c'est-à-dire

$$S_i = w_i x_i / \sum_i w_i x_i \quad i = L, K, CI$$

Il faut vérifier que, pour chaque année, la somme de ces 3 parts est bien égale à 1.

Le tableau 10.3 est juste une vérification du calcul des parts du coût de chaque facteur considéré dans le coût total, la somme (notée SOMS) est constante (écart type nul) et est égale à 1.

Tableau 10.3 Statistique sur les parts du coût de chaque facteur

Statistics on Series SOMS			
Annual Data From 1978 :01 To 1993 :01			
Observations	16	Variance	6.573841e-033
Sample Mean	1.000000000000	SE of Sample Mean	2.026981e-017
Standard Error	0.000000000000	Signif Level (Mean=0)	0.000000000
t-Statistic	4.93345e+016	Signif Level (Sk=0)	0.00051693
Skewness	-2.34738	Signif Level (Ku=0)	0.03301334
Kurtosis	3.29670		

3. Dans un premier temps, on représente graphiquement l'évolution sur la période d'observation des indices de prix exprimés en base 1980. Le début des années 80 marque une rupture dans l'évolution des prix qui jusque là avaient une évolution à la hausse sensiblement similaire. À cette date, le coût du facteur capital diminue de façon importante. La seconde indication intéressante est la forte augmentation du coût du facteur travail. Pour avoir une information sur l'intensité d'utilisation de chaque facteur de production, on détermine le rapport entre le niveau de chaque facteur et le niveau de production. L'évolution de ce rapport pour chaque facteur (noté $I_i, i = K, L, CI$) est représentée sur la figure 10.2.

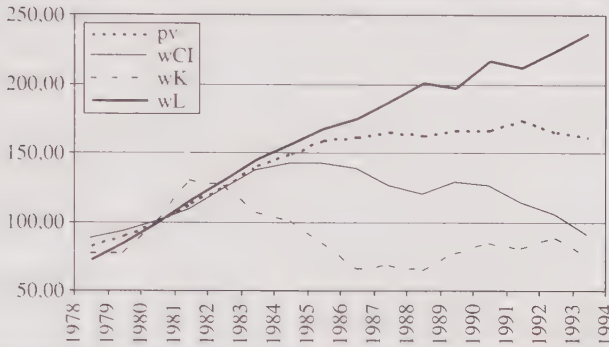


Figure 10.1 Évolution des prix

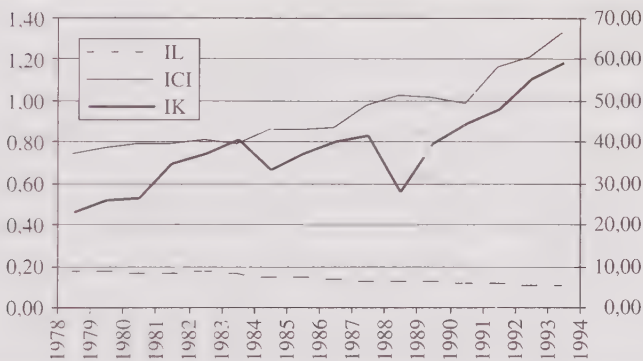


Figure 10.2 Intensité de chaque facteur

Depuis la fin des années 80, la production devient plus intensive en capital et en consommation intermédiaire. L'intensité en travail est décroissante.

Problème 10.2

Estimation du système de parts de dépense et interprétation des résultats

1. Pour que ce système soit dérivé d'une fonction de coût, les paramètres doivent vérifier les contraintes de symétrie et la contrainte d'homogénéité de degré 1 par rapport aux prix. Le système de parts de dépense dérivé d'une fonction de coût translog en imposant les contraintes de symétrie ($\alpha_{ij} = \alpha_{ji}$) et d'homothétie ($\alpha_{iy} = 0$) s'écrit

$$Sk = \alpha_K + \alpha_{KL} \ln w_L + \alpha_{KK} \ln w_K + \alpha_{KCI} \ln w_{CI} \quad (10.52)$$

$$Sl = \alpha_L + \alpha_{LL} \ln w_L + \alpha_{KL} \ln w_K + \alpha_{LCI} \ln w_{CI} \quad (10.53)$$

$$Sci = \alpha_{CI} + \alpha_{LCI} \ln w_L + \alpha_{KCI} \ln w_K + \alpha_{CICI} \ln w_{CI} \quad (10.54)$$

La contrainte d'homogénéité de degré 1 par rapport aux prix impose les relations suivantes

$$\begin{aligned}\alpha_K + \alpha_L + \alpha_{CI} &= 1 \\ \alpha_{KL} + \alpha_{KK} + \alpha_{KCI} &= 0 \\ \alpha_{LL} + \alpha_{KL} + \alpha_{LCI} &= 0 \\ \alpha_{LCI} + \alpha_{KCI} + \alpha_{CICI} &= 0\end{aligned}$$

Comme les conditions d'homogénéité par rapport au prix sont imposées alors les conditions d'additivité sont respectées.

Si l'on exclut l'équation de la part des consommations intermédiaires (système 1), on obtient le système suivant

$$Sk = \alpha_K + \alpha_{KL} \ln(w_L/w_{CI}) + \alpha_{KK} \ln(w_K/w_{CI}) \quad (10.55)$$

$$Sl = \alpha_L + \alpha_{LL} \ln(w_L/w_{CI}) + \alpha_{KL} \ln(w_K/w_{CI}) \quad (10.56)$$

La troisième part se déduisant par solde

$$\begin{aligned}Sci &= (1 - \alpha_K - \alpha_L) + (-\alpha_{LL} - \alpha_{KL}) \ln w_L + \\ &(-\alpha_{KL} - \alpha_{KK}) \ln w_K + (-\alpha_{LCI} - \alpha_{KCI}) \ln w_{CI}\end{aligned} \quad (10.57)$$

avec $\alpha_{LCI} = -\alpha_{LL} - \alpha_{KL}$ et $\alpha_{KCI} = -\alpha_{KL} - \alpha_{KK}$

Si l'on exclut l'équation de la part du capital (système 2), on utilise les contraintes suivantes

$$\begin{aligned}\alpha_K &= 1 - \alpha_{CI} - \alpha_L \\ \alpha_{KK} &= -\alpha_{KL} - \alpha_{KCI} \\ \alpha_{KL} &= -\alpha_{LL} - \alpha_{LCI} \\ \alpha_{KCI} &= -\alpha_{LCI} - \alpha_{CICI}\end{aligned}$$

On introduit ces relations dans les équations (10.52), (10.53) et (10.54). On obtient le système suivant

$$\begin{aligned}Sk &= (1 - \alpha_{CI} - \alpha_L) + (-\alpha_{LL} - \alpha_{LCI}) \ln w_L + \\ &(-\alpha_{KL} - \alpha_{KCI}) \ln w_K + (-\alpha_{LCI} - \alpha_{CICI}) \ln w_{CI}\end{aligned} \quad (10.58)$$

$$Sl = \alpha_L + \alpha_{LL} \ln w_L + (-\alpha_{LL} - \alpha_{LCI}) \ln w_K + \alpha_{LCI} \ln w_{CI} \quad (10.59)$$

$$\begin{aligned}Sci &= \alpha_{CI} + \alpha_{LCI} \ln w_L + (-\alpha_{LCI} - \alpha_{CICI}) \ln w_K \\ &+ \alpha_{CICI} \ln w_{CI}\end{aligned} \quad (10.60)$$

Les deux dernières équations peuvent se réécrire

$$Sl = \alpha_L + \alpha_{LL} \ln(w_L/w_K) + \alpha_{LCI} \ln(w_{CI}/w_K) \quad (10.61)$$

$$Sci = \alpha_{CI} + \alpha_{LCI} \ln(w_L/w_K) + \alpha_{CICI} \ln(w_{CI}/w_K) \quad (10.62)$$

Les deux systèmes présentent des coefficients communs α_L et α_{LL} dans l'équation de la part de dépense du travail.

2. On rajoute des termes d'erreur dans le premier système (formé par les équations (10.55) et (10.56)) et on l'estime par les MCO. La variable $\ln(w_L/w_{CI})$ est notée *LWL*CI dans les tableaux de résultats, la variable $\ln(w_K/w_{CI})$ *LWK*CI.

Tableau 10.4 Résultats des MCO

Dependent Variable SK — Estimation by Least Squares			
Annual Data From 1978 :01 To 1993 :01			
Usable Observations	16	Degrees of Freedom	13
Centered R**2	0.634210	R Bar **2	0.577934
Uncentered R**2	0.999939	T x R**2	15.999
Mean of Dependent Variable	0.9598974928		
Std Error of Dependent Variable	0.0128008313		
Standard Error of Estimate	0.0083162632		
Sum of Squared Residuals	0.0008990830		
Regression F(2,13)	11.2697		
Significance Level of F	0.00144880		
Durbin-Watson Statistic	1.750808		
Q(4-0)	1.393920		
Significance Level of Q	0.84525114		

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	0.9708865153	0.0034865690	278.46474	0.00000000
2.	LWLCI	0.0051670766	0.0068843930	0.75055	0.46629253
3.	LWKCI	0.0407569883	0.0086880941	4.69113	0.00042192

Dependent Variable SL — Estimation by Least Squares			
Annual Data From 1978 :01 To 1993 :01			
Usable Observations	16	Degrees of Freedom	13
Centered R**2	0.571672	R Bar **2	0.505775
Uncentered R**2	0.963209	T x R**2	15.411
Mean of Dependent Variable	0.0065227998		
Std Error of Dependent Variable	0.0020650502		
Standard Error of Estimate	0.0014517536		
Sum of Squared Residuals	0.0000273987		
Regression F(2,13)	8.6753		
Significance Level of F	0.00404155		
Durbin-Watson Statistic	1.923109		
Q(4-0)	1.732027		
Significance Level of Q	0.78489196		

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	0.004921106	0.000608643	8.08537	0.00000199
2.	LWKCI	-0.006293866	0.001516663	-4.14981	0.00114163
3.	LWLCI	-0.001124525	0.001201795	-0.93570	0.36648264

Dependent Variable SCI — Estimation by Least Squares			
Annual Data From 1978 :01 To 1993 :01			
Usable Observations	16	Degrees of Freedom	13
Centered R**2	0.644119	R Bar **2	0.589368
Uncentered R**2	0.968742	T x R**2	15.500
Mean of Dependent Variable	0.0335797075		
Std Error of Dependent Variable	0.0107618114		
Standard Error of Estimate	0.0068962326		
Sum of Squared Residuals	0.0006182543		
Regression F(2,13)	11.7645		
Significance Level of F	0.00121195		
Durbin-Watson Statistic	1.722500		
Q(4-0)	1.354271		
Significance Level of Q	0.85210268		

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	0.024192378	0.002891225	8.36752	0.00000136
2.	LWKCI	-0.034463122	0.007204572	-4.78351	0.00035726
3.	LWLCI	-0.004042552	0.005708859	-0.70812	0.49137358

En estimant équation par équation, on ne prend pas en compte les contraintes de symétrie des coefficients entre les équations. De plus, les corrélations des erreurs entre les équations ne sont pas intégrées. En outre, si on estime le système 2 (système obtenu en omettant l'équation liée au capital) il n'y a aucune garantie d'obtenir les mêmes estimations pour les coefficients communs aux deux systèmes.

3. La méthode de Zellner permet de tenir compte d'une éventuelle corrélation entre les résidus d'estimation des différentes équations. Les contraintes d'additivité des parts des facteurs sont prises en compte en estimant deux équations sur les trois que comporte le système, la dernière équation étant obtenue par solde. On applique tout d'abord la méthode de Zellner au système formé par les équations (10.55) et (10.56). La variable $\ln(w_L/w_{CI})$ est notée *LWLCI* dans les tableaux de résultats, la variable $\ln(w_K/w_{CI})$ *LWKCI*.

Tableau 10.5 Estimation du système en enlevant l'équation sci
Méthode de Zellner

Dependent Variable SK — Estimation by Seemingly Unrelated Regressions			
Annual Data From 1978 :01 To 1993 :01			
Usable Observations	16	Degrees of Freedom	13
Centered R**2	0.576128	R Bar **2	0.510917
Uncentered R**2	0.999929	T x R**2	15.999
Mean of Dependent Variable	0.9598974928		
Std Error of Dependent Variable	0.0128008313		
Standard Error of Estimate	0.0089521899		
Sum of Squared Residuals	0.0010418422		
Durbin-Watson Statistic	1.438723		

	Variable	Coeff	Std Error	T-Stat	Signif
2.	Constant	0.971845934	0.003136618	309.83882	0.00000000
1.	LWLCI	-0.004531865	0.001944334	-2.33081	0.01976364
3.	LWKCI	0.034658016	0.007004273	4.94813	0.00000075

Dependent Variable SL — Estimation by Seemingly Unrelated Regressions

Annual Data From 1978 :01 To 1993 :01

Usable Observations	16	Degrees of Freedom	13
Centered R**2	0.473315	R Bar **2	0.392287
Uncentered R**2	0.954761	T x R**2	15.276
Mean of Dependent Variable	0.0065227998		
Std Error of Dependent Variable	0.0020650502		
Standard Error of Estimate	0.0016098294		
Sum of Squared Residuals	0.0000336902		
Durbin-Watson Statistic	1.505741		

	Variable	Coeff	Std Error	T-Stat	Signif
4.	Constant	0.004909332	0.000894339	5.48934	0.00000004
1.	LWKCI	-0.004531865	0.001944334	-2.33081	0.01976364
5.	LWLCI	0.000767735	0.001322985	0.58031	0.56170870

Pour vérifier si les résultats d'estimation sont dépendants de l'équation omise, on applique la méthode de Zellner en omettant la part associée au capital c'est-à-dire que l'on estime le système formé par les équations (10.61) et (10.62).

On peut comparer directement les coefficients obtenus dans l'équation (10.56) aux valeurs déterminées dans l'estimation de l'équation (10.61). De manière plus précise, on compare les valeurs obtenues pour la constante et les valeurs obtenues pour les coefficients de $\ln(w_L/w_{CI})$ et de $\ln(w_L/w_K)$.

Tableau 10.6 Estimation du système en enlevant l'équation sk
Méthode de Zellner

Dependent Variable SL — Estimation by Seemingly Unrelated Regressions

Annual Data From 1978 :01 To 1993 :01

Usable Observations	16	Degrees of Freedom	13
Centered R**2	0.459236	R Bar **2	0.376042
Uncentered R**2	0.953552	T x R**2	15.257
Mean of Dependent Variable	0.0065227998		
Std Error of Dependent Variable	0.0020650502		
Standard Error of Estimate	0.0016312045		
Sum of Squared Residuals	0.0000345908		
Durbin-Watson Statistic	1.449352		

	Variable	Coeff	Std Error	T-Stat	Signif
2.	Constant	0.0048100034	0.0008170872	5.88677	0.00000000
1.	LWCIK	0.0036614020	0.0017882296	2.04750	0.04060891
3.	LWLK	0.0009863863	0.0009175713	1.07500	0.28237613

On voit qu'il existe des différences dans les paramètres estimés pour la constante c'est-à-dire le paramètre $\widehat{\alpha}_L$ et aussi pour le paramètre $\widehat{\alpha}_{LL}$. Pour obtenir des estimations qui soient indépendantes du choix de l'équation omise, il faut utiliser la méthode de Zellner itérative.

4. Les deux systèmes sont estimés par la méthode de Zellner itérative. Pour chaque système, on ne présente que les résultats liés à l'équation de la part de travail.

La première estimation présentée dans le tableau 10.7 donne les résultats de l'estimation de l'équation (10.56).

Tableau 10.7 Estimation du système en enlevant l'équation sci
Méthode de Zellner itérative

Linear Systems — Estimation by Seemingly Unrelated Regressions	
Iterations Taken	4
Annual Data From 1978 :	01 To 1993 :01
Usable Observations	16

Dependent Variable SL			
Centered R**2	0.496635	R Bar **2	0.602607
Uncentered R**2	0.956764	T x R**2	15.308
Mean of Dependent Variable	0.0065227998		
Std Error of Dependent Variable	0.0020650502		
Standard Error of Estimate	0.0013017899		
Sum of Squared Residuals	0.0000321985		
Durbin-Watson Statistic	1.548974		

	Variable	Coeff	Std Error	T-Stat	Signif
4.	Constant	0.004744401	0.000584198	8.12123	0.00000000
1.	LWKCI	-0.005177729	0.001299673	-3.98387	0.00006780
5.	LWLCI	0.000654288	0.000294519	2.22155	0.02631372

La seconde estimation obtenue à partir du système où l'on omet la part associée au capital correspond à l'équation (10.61). Les résultats sont présentés dans le tableau 10.8.

Tableau 10.8 Estimation du système en enlevant l'équation sk
Méthode de Zellner itérative

Linear Systems — Estimation by Seemingly Unrelated Regressions	
Iterations Taken	5
Annual Data From 1978 :	01 To 1993 :01
Usable Observations	16

Dependent Variable SL			
Centered R**2	0.496702	R Bar **2	0.602660
Uncentered R**2	0.956770	T x R**2	15.308
Mean of Dependent Variable	0.0065227998		
Std Error of Dependent Variable	0.0020650502		
Standard Error of Estimate	0.0013017034		
Sum of Squared Residuals	0.0000321942		
Durbin-Watson Statistic	1.549194		

	Variable	Coeff	Std Error	T-Stat	Signif
2.	Constant	0.0047444189	0.0005839362	8.12489	0.00000000
1.	LWCIK	0.0045248716	0.0011067867	4.08830	0.00004346
3.	LWLK	0.0006535244	0.0002943634	2.22013	0.02641008

On vérifie que l'on retrouve la même valeur pour la constante et la même valeur pour les coefficients de $\ln(w_L/w_{CI})$ et $\ln(w_L/w_K)$.

5. Élasticités de substitution d'Allen.

La positivité de l'élasticité d'Allen entre les consommations intermédiaires et le travail indique que ces facteurs sont substituables sur la période d'observation. Par contre, on remarque que les relations entre le facteur travail et le facteur capital ne sont pas constantes sur la période. En fin de période, ils sont complémentaires.

11

MODÉLISATION D'UN ÉQUILIBRE DE MARCHÉ

1. Introduction
2. Spécification du modèle
3. Analyse empirique
4. Éléments de corrigé

OBJECTIFS DU CHAPITRE

D'un point de vue méthodologique

- Modèle à équations simultanées
- Forme Structurelle du Modèle
- Méthode des Doubles Moindres Carrés et Triples Moindres Carrés
- Simulations statiques et dynamiques

D'un point de vue économique

- Modèle de marché
- Analyse de politique économique

NIVEAU

MASTER, première et deuxième années

CHAMPS ÉCONOMIQUES TRAITÉS

Microéconomie

Analyse d'un marché

1. INTRODUCTION

Pour illustrer le fonctionnement d'un modèle de marché en équilibre, on étudie dans ce chapitre le marché de la viande de bœuf en France. De manière très générale, un modèle de marché comporte deux groupes d'équations : les équations de comportement et les équations traduisant les conditions d'équilibre du marché. Le premier groupe d'équations exprime la réaction des agents économiques participant au marché (en particulier les producteurs et les consommateurs) par rapport à des variations de prix. La spécification de ces équations est issue d'hypothèses particulières sur le comportement des agents participant au marché étudié. Le second groupe d'équations décrit l'équilibre du marché c'est-à-dire impose l'égalité entre ressources et utilisations. Seul le premier groupe comporte des équations avec des paramètres à estimer. On retient comme illustration le marché de la viande de bœuf en France.

La seconde section du chapitre porte sur la description du modèle estimé dans la partie empirique. Dans une première étape, on présente la spécification retenue pour les équations économétriques du modèle et on décrit le comportement des agents participant au marché. Dans un second temps, la forme structurelle du modèle est écrite, puis les conditions d'identification du modèle et les implications sur les méthodes d'estimation sont présentées.

2. SPÉCIFICATION DU MODÈLE

Le modèle étudié dans ce chapitre comprend cinq équations : quatre équations de comportement (équations à estimer) et une équation décrivant l'équilibre du marché. Les spécifications de ces équations sont successivement détaillées.

2.1 Équations de comportement

Les équations à estimer sont l'équation d'offre, l'équation de demande, l'équation de marge et l'équation du solde du commerce extérieur.

2.1.1 Équation d'offre

Pour l'équation d'offre, on se reporte aux développements détaillés dans le chapitre 8 (cf. chapitre 8, La fonction d'offre de viande bovine). On adopte la même spécification à savoir

$$q_t = \alpha_0 + \alpha_1 p_{t-1} + \alpha_2 p_{t-2} + \alpha_3 p_{t-3} + \alpha_4 q_{t-1} + \alpha_5 w_{t-1} + \alpha_6 D74 \quad (11.1)$$

où q_t mesure la quantité offerte l'année t , p_{t-i} , $i = 1, 2, 3$ représente le prix du bien produit en $t-i$, w_{t-1} représente le prix des aliments du bétail l'année $t-1$, $D74$ est

une variable indicatrice qui permet de prendre en compte le boycott du «veau aux hormones» par les consommateurs durant l'année 1973, cet effet ayant eu un impact sur la production l'année suivante.

2.1.2 Équation de demande

L'hypothèse de rationalité du comportement suppose que les fonctions de demande résultent de la maximisation de l'utilité du consommateur sous la contrainte de budget. Cette hypothèse entraîne différentes propriétés sur les fonctions de demande du consommateur (en particulier les propriétés d'additivité, d'homogénéité et de symétrie). Dans les études économétriques appliquées, différentes formes fonctionnelles sont utilisées pour estimer des systèmes complets de fonctions de demande en respectant les contraintes théoriques (par exemple, le système LES, le système AIDS,... pour avoir une synthèse sur les systèmes de demande classiques, se reporter à Deaton et Muellbauer (1980)¹).

Par souci de simplification, on utilise une spécification *ad hoc* pour définir une équation de demande². On retient ainsi une équation linéaire reliant la consommation de viande au prix du bien, aux prix des biens substituables, des biens complémentaires et au revenu. Cette équation s'écrit

$$C_t^* = \beta_0^* + \sum_{j=1}^n \beta_j^* \frac{P_{jt}}{IPC_t} + \theta^* \frac{R_t}{IPC_t} \quad (11.2)$$

où C_t^* représente le niveau de consommation de viande de bœuf souhaité par tête, P_{jt} mesure le prix du bien j , ($j = 1, \dots, n$), IPC_t un indice des prix à la consommation, R_t le revenu disponible des ménages, β_0^* , β_j^* et θ^* les coefficients à estimer.

L'équation de demande est homogène de degré zéro par rapport aux prix et au revenu. Ainsi si tous les prix et le revenu augmentent d'un même pourcentage, la consommation reste inchangée. Cette propriété correspond à l'hypothèse d'absence d'illusion monétaire. Dans la suite du chapitre, les lettres minuscules pour les variables monétaires de l'équation de consommation noteront des variables déflatées par l'indice du prix à la consommation.

Pour améliorer la qualité de l'estimation de l'équation de demande, on suppose qu'il existe une lenteur dans les changements de comportement. Aussi, les

¹ Deaton A.S., Muellbauer J., (1980), *The economics of consumer behaviour*, Cambridge University Press.

² La forme retenue pour l'équation de demande est, ainsi, choisie de façon arbitraire. Cette relation ne permet pas de garantir les propriétés théoriques induites par la théorie de la demande, c'est-à-dire les hypothèses relatives à la rationalité du comportement des consommateurs. Sous des hypothèses particulières associées aux préférences des consommateurs, il serait ainsi plus opportun d'estimer les paramètres de l'équation de viande à partir d'un système étudiant simultanément les dépenses consacrées aux différents biens alimentaires. Rappelons que l'objet du chapitre est d'étudier le fonctionnement d'un modèle de marché. L'amélioration de l'estimation de la fonction de demande est une voie à privilégier pour enrichir le modèle comme on le verra dans la partie empirique.

séries observées de consommation ne correspondent pas au niveau de consommation souhaité par le consommateur.

La réaction des agents économiques aux variations de prix se fait souvent avec lenteur et décalage (Drouet, Mahé (1978))³ d'une part, parce que les agents ont besoin de temps pour réaliser les changements, et d'autre part, parce qu'il existe des habitudes dans le comportement. Pour rendre compte de cette inertie de comportement, l'équation de demande va comporter une variable endogène décalée. De manière plus précise, on va introduire un mécanisme d'ajustement partiel.

De manière analytique, cette hypothèse signifie que l'on peut décrire les variations de la consommation entre des périodes successives de la manière suivante

$$C_t - C_{t-1} = \lambda(C_t^* - C_{t-1}) \quad (11.3)$$

où C_t représente le niveau observé de consommation par tête à la période t , C_{t-1} représente le niveau observé à la période $t-1$, λ représente le coefficient d'ajustement partiel, C_t^* définit le niveau désiré de consommation. Si l'ajustement est instantané ($\lambda = 1$), le niveau de consommation observé est juste égal au niveau de consommation désiré. On fait une hypothèse sur la relation entre C_t^* et les variables exogènes (prix et revenu) (voir équation 13.2).

$$C_t^* = \beta_0^* + \sum_{j=1}^n \beta_j^* p_{jt} + \theta^* r_t$$

Dans cette équation, les coefficients définissent les réponses à long terme du consommateur. Ainsi les coefficients β_j^* mesurent l'impact marginal de long terme des prix sur la consommation et θ^* l'impact marginal de long terme du revenu.

Pour obtenir une relation estimable, on introduit la définition de la consommation désirée dans l'équation d'ajustement partiel. On obtient

$$C_t = \beta_0 + (1 - \lambda)C_{t-1} + \sum_{j=1}^n \beta_j p_{jt} + \theta r_t \quad (11.4)$$

avec $\beta_0 = \lambda\beta_0^*$, $\beta_j = \lambda\beta_j^*$ et $\theta = \lambda\theta^*$.

Le coefficient de la variable endogène décalée permet de calculer le paramètre λ c'est-à-dire la vitesse de réaction des consommateurs, les autres coefficients dans cette équation représentent les impacts marginaux de court terme⁴. Les paramètres estimés permettront d'analyser la rapidité de la réaction des consommateurs aux variations de prix en calculant les élasticités de la demande à court terme et à long

³ Drouet M., Mahé L., (1978), « Le marché du bœuf : analyse et politique de stabilisation », *Statistiques et Études Financières*, Direction de la prévision.

⁴ Si les variables sont exprimées en log, les coefficients s'interprètent comme des élasticités.

terme, par rapport aux prix et au revenu. Dans l'analyse empirique, on dispose du prix d'un bien substituable, le prix de la viande de mouton.

L'équation de consommation s'écrit

$$C_t = \beta_0 + (1 - \lambda)C_{t-1} + \beta_1 p_{bt} + \beta_2 p_{mt} + \theta r_t$$

où p_{bt} définit le prix de la viande bœuf et p_{mt} définit le prix de la viande de mouton. Ces variables sont déflatés par l'indice des prix à la consommation.

2.1.3 Équation de transmission de prix

L'équation de marge définit le prix à la consommation de la viande bovine (noté P_{bt}) comme une fonction linéaire du prix à la production (noté p_t). Cette équation s'écrit

$$P_{bt} = \gamma_0 + \gamma_1 p_t \quad (11.5)$$

Cette équation permet d'étudier la transmission des variations de prix à la production au prix à la consommation.

2.1.4 Équation du commerce extérieur

La dernière équation à estimer du modèle correspond à l'équation du solde du commerce extérieur. Ce solde est, par définition, égal à la différence entre la valeur des exportations et la valeur des importations. Ce solde est noté B_t . On suppose que ce solde dépend du niveau des stocks de l'année précédente (noté s_{t-1}). L'équation s'écrit

$$B_t = \delta_0 + \delta_1 s_{t-1} \quad (11.6)$$

2.2 La forme structurelle du modèle

2.2.1 Équation d'équilibre

La dernière équation du modèle décrit le fonctionnement du marché et correspond à l'équation d'équilibre, c'est-à-dire à l'égalité comptable entre l'offre et la demande. Cette équation permet de déterminer le niveau d'équilibre des prix, les prix vont s'établir de façon à ce que l'offre égalise la demande (égale à la somme de la consommation, du solde du commerce extérieur et à la variation des stocks). Cette équation s'écrit

$$q_t = ctot_t + B_t + (s_t - s_{t-1}) \quad (11.7)$$

où $ctot_t$ représente la consommation totale de bœuf (consommation par tête multipliée par le nombre de consommateurs).

Cette équation est une équation d'identité et ne comporte aucun paramètre à estimer. Les stocks de l'année courante sont considérés de manière exogène.

2.2.2 Écriture du modèle économétrique

Pour spécifier le modèle économétrique, on rajoute des aléas dans chaque équation de comportement.

En résumé, la forme structurelle du modèle étudié s'écrit

$$q_t = \alpha_0 + \alpha_1 p_{t-1} + \alpha_2 p_{t-2} + \alpha_3 p_{t-3} + \quad (11.8)$$

$$\alpha_4 q_{t-1} + \alpha_5 w_{t-1} + \alpha_6 D74 + \varepsilon_{1t} \quad (11.9)$$

$$C_t = \beta_0 + (1 - \lambda)C_{t-1} + \beta_1 \frac{P_{bt}}{IPC_t} + \beta_2 \frac{P_{mt}}{IPC_t} + \theta r_t + \varepsilon_{2t} \quad (11.10)$$

$$P_{bt} = \gamma_0 + \gamma_1 p_t + \varepsilon_{3t} \quad (11.11)$$

$$B_t = \delta_0 + \delta_1 s_{t-1} + \varepsilon_{4t} \quad (11.12)$$

$$q_t = ctot_t + B_t + (s_t - s_{t-1}) \quad (11.13)$$

où ε_{1t} représente l'aléa associé à l'équation d'offre, ε_{2t} définit l'aléa associé à l'équation de consommation, ε_{3t} définit l'aléa de l'équation de marge, et ε_{4t} est l'aléa dans l'équation du solde du commerce extérieur. Les variables endogènes sont C_t , q_t , P_{bt} , p_t et B_t . Les variables prédéterminées correspondent aux variables endogènes décalées C_{t-1} , p_{t-1} , p_{t-2} , p_{t-3} , q_{t-1} et aux variables exogènes P_{mt} , w_{t-1} , r_t , $D74$, IPC_t , s_t et s_{t-1} . Seule la dernière équation qui définit une identité comptable ne comporte pas d'aléa.

2.3 Estimation du modèle

Pour pouvoir estimer le modèle, on doit vérifier que les équations de la forme structurelle du modèle sont identifiées (juste identifiées ou sur-identifiées). Les conditions d'identification s'étudient équation par équation.

On a deux équations, l'équation d'offre (équation 11.8) et l'équation liée au solde du commerce extérieur (équation 11.12), qui ne contiennent aucune variable endogène en variable explicative. La méthode des Moindres Carrés Ordinaires, sous réserve que les aléas aient les bonnes propriétés peut donc être appliquée⁵ sur chaque équation.

Pour les autres équations, on examine successivement les conditions d'ordre et de rang. On peut montrer que ces équations sont sur-identifiées. Aussi, si l'on estime équation par équation on utilisera la méthode des doubles moindres carrés.

Si l'on estime toutes les équations simultanément, on utilisera la méthode des triples moindres carrés (méthode qui reste très sensible aux erreurs de spécification, cf. chapitre 12, Les équations simultanées).

⁵ Toutefois pour l'équation d'offre qui est dynamique, il faut vérifier que les aléas ne sont pas autocorrélés sinon la variable endogène décalée est endogène (voir chapitre 8).

3. ANALYSE EMPIRIQUE

Description des données

L'étude appliquée porte sur le marché du bœuf en France à partir de données annuelles sur la période 1961-1998. Les données sur les quantités mesurées en milliers de tonnes proviennent de l'United States of Department of Agriculture (USDA). Dans le fichier de données on trouve : la production bovine de l'année t notée q_t , la consommation totale $ctot_t$, le solde du commerce extérieur B_t et le niveau des stocks publics en fin d'année s_t .

Par ailleurs, on dispose de différents indices de prix exprimés en base 1970 : prix à la production du bœuf (noté p_t), le prix à la consommation du bœuf (P_{bt}), le prix à la consommation du mouton (P_{mt}), l'indice des prix à la consommation finale (IPC). Enfin, on dispose de l'indice de prix des aliments du bétail (w_t) exprimé en base 1980. Les données sur les prix proviennent du SCEES (Service Central des Enquêtes et Études Statistiques).

Le revenu disponible des ménages français en francs courants est noté (R_t) le niveau de population de la France au 1^{er} janvier de l'année en cours (pop_t).

Problème 11.1

Construction des variables et analyse descriptive

1. Transformer les indices de prix de façon à retenir l'année 1990 comme année de référence.
2. Construire les rapports de prix nécessaires pour estimer les équations d'offre, de demande et de marge. Construire la demande de bœuf par tête (notée C_t) et le revenu réel par tête (noté r_t).
3. Écrire et vérifier l'équation comptable.
4. Analyser graphiquement les différentes variables.

Problème 11.2

Estimation du modèle

1. Écrire le modèle de marché en équilibre. Par rapport au modèle théorique présenté, dans l'application
 - (i) on estimera une équation de consommation par tête en fonction des prix et du revenu par tête afin de prendre en compte un «effet de taille» lié à la population.
 - (ii) dans un souci de simplification, dans l'équation de marge on utilisera les variables de prix à la consommation et à la production déflatées par l'indice des prix à la consommation finale.

Donner la liste des variables endogènes, la liste des variables exogènes.

2. Quelles méthodes d'estimation doit-on appliquer si l'on veut estimer le modèle équation par équation ?
3. Estimer le modèle équation par équation jusqu'en 1995.
4. Estimer toutes les équations en même temps. Commenter.

Problème 11.3

Qualité prédictive du modèle et exemples de simulation

1. Simuler le modèle de manière statique. Comparer les variables d'offre, de consommation, de prix à la consommation et de solde, prédites par le modèle avec les données observées. Comparer avec les données exclues de l'estimation.
2. On veut faire une simulation dynamique. Donner la différence par rapport à la simulation statique.
3. Simuler une multiplication par deux du prix des aliments du bétail.

4. ÉLÉMENTS DE CORRIGÉ

Problème 11.1

Construction des variables et analyse descriptive

1. Pour obtenir les variables prix en base 1990, on divise chaque série par sa valeur observée en 1990 et on la multiplie par 100. Les nouvelles séries de prix sont ainsi toutes égales à 100 en 1990.
2. Pour l'estimation de l'équation de la demande, on a tout d'abord besoin de la variable consommation par tête et du revenu disponible par tête. Ces deux variables sont donc divisées par le niveau de la population. De plus, comme la fonction de demande est homogène de degré zéro par rapport au prix et au revenu, on déflate les différentes variables par l'indice général des prix. Ce calcul permet de garantir la propriété d'homogénéité (absence d'illusion monétaire).
3. L'égalité comptable traduit l'équilibre sur le marché (voir équation (13.13)). L'offre est égale à la somme de la consommation, du solde du commerce extérieur et de la variation de stock. Cette équation est nécessaire quand on utilise le modèle en simulation. Pour vérifier cette égalité, on contrôle que la différence entre l'offre et la consommation, le solde du commerce extérieur et la variation de stock s'annule.
4. On représente les évolutions de la consommation totale et de la production.

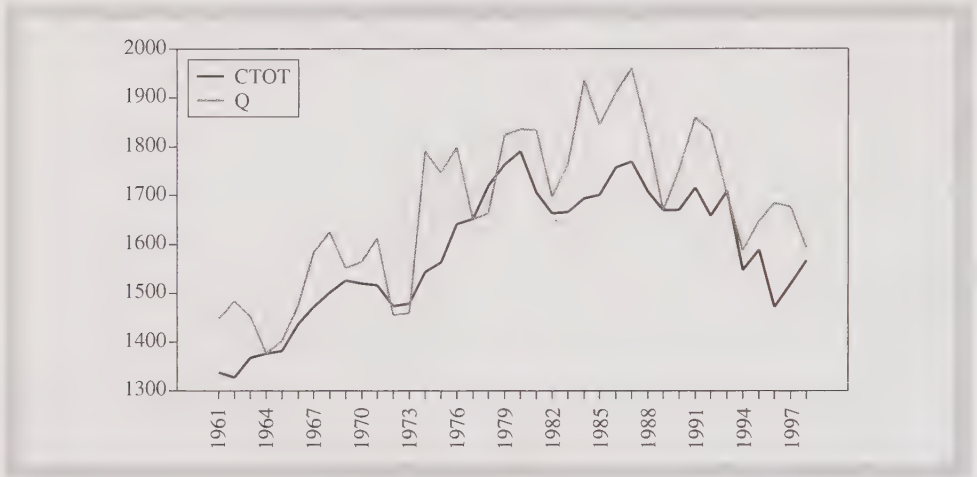


Figure 11.1 Évolution de la consommation totale et de la production

L'évolution de la consommation de la viande bovine est caractérisée par une phase de croissance suivie d'un déclin à partir du début des années quatre-vingt (hormis des chocs conjoncturels comme, par exemple, les boycottages de 1973 et de 1980). Deux facteurs sont généralement avancés pour expliquer cette évolution : le coût élevé de la viande et une certaine désaffection de la part des consommateurs au profit de viandes blanches (porc, poulet).

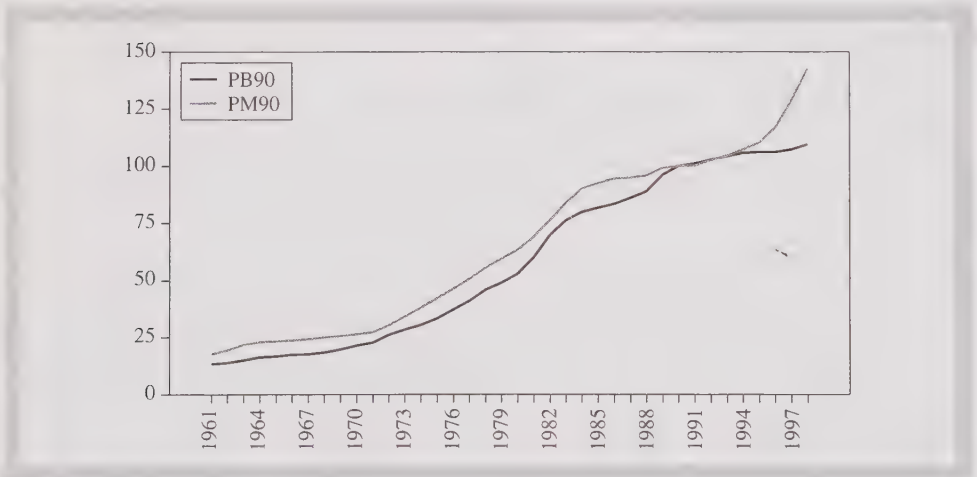


Figure 11.2 Évolution des indices de prix à la consommation du bœuf et du mouton (base 1990)

Les prix à la consommation du mouton et du bœuf ont augmenté de manière assez continue et de façon plutôt linéaire et parallèle jusqu'au début des années quatre-vingt dix. Depuis, on observe une hausse importante du prix du mouton et une relative stagnation du prix du bœuf.

Problème 11.2

Estimation du modèle

1. Le modèle de marché est constitué d'équations de comportement et de l'équation comptable traduisant l'équilibre sur le marché.

$$q_t = \alpha_0 + \alpha_1 p_{t-1} + \alpha_2 p_{t-2} + \alpha_3 p_{t-3} + \quad (11.14)$$

$$\alpha_4 q_{t-1} + \alpha_5 w_{t-1} + \alpha_6 D74 + \varepsilon_{1t}$$

$$C_t = \beta_0 + (1 - \lambda)C_{t-1} + \beta_1 p_{bt} + \beta_2 p_{mt} + \theta r_t + \varepsilon_{2t} \quad (11.15)$$

$$p_{bt} = \gamma_0 + \gamma_1 pr_t + \varepsilon_{3t} \quad (11.16)$$

$$B_t = \delta_0 + \delta_1 s_{t-1} + \varepsilon_{4t} \quad (11.17)$$

$$q_t = ctot_t + B_t + (s_t - s_{t-1}) \quad (11.18)$$

où ε_{1t} représente l'aléa associé à l'équation de consommation, ε_{2t} définit l'aléa associé à l'équation d'offre, ε_{3t} définit l'aléa de l'équation de marge, et ε_{4t} est l'aléa du solde du commerce extérieur. Les variables endogènes sont C_t , q_t , p_{bt} , pr_t et B_t . La variable C_t est égale à $ctot_t/pc_{r_t}$ et la variable $pr_t = p_t/IPC_t$. Les variables prédéterminées correspondent aux variables endogènes décalées C_{t-1} , p_{t-1} , p_{t-2} , p_{t-3} , q_{t-1} et aux variables exogènes p_{mt} , w_{t-1} , r_t , $D74$, s_t et s_{t-1} .

2. Dans le modèle structurel, deux équations ne contiennent aucune variable endogène : l'équation d'offre et l'équation liée au solde du commerce extérieur. La méthode des Moindres Carrés Ordinaires, sous réserve que les aléas aient les bonnes propriétés peut être appliquée.

Pour les autres équations on examine les conditions d'ordre. Ainsi, dans l'équation de demande, on a 8 variables exogènes exclues de l'équation, le nombre de variables endogènes incluses est égal à 2. De manière équivalente, on peut vérifier que le nombre total de variables exclues de cette équation (11) est supérieur au nombre d'équations du modèle moins un (4). La condition d'ordre est donc vérifiée et l'équation est sur-identifiée. Pour l'équation de marge, on a 11 variables exogènes exclues de l'équation pour 2 variables endogènes (ou encore le nombre total de variables exclues de cette équation (14) est supérieur au nombre d'équations du modèle moins un (4)), l'équation est aussi sur-identifiée. Pour la dernière équation du modèle (équation d'équilibre), on n'a pas, par définition, de condition d'identification à établir (il n'y a pas de paramètres à estimer). On peut vérifier par ailleurs que les conditions de rang sont satisfaites.

Le résultat donné par la condition d'ordre et de rang implique que si on veut estimer équation par équation on utilisera la méthode des DMC pour l'équation de demande et l'équation de marge.

3. L'équation d'offre est estimée par la méthode des MCO. Avant de commenter ces résultats, il est nécessaire de vérifier la validité de l'hypothèse de non

corrélation des aléas. Ce modèle est autorégressif c'est-à-dire, la variable endogène décalée figure parmi les variables explicatives. Si les perturbations de l'équation sont corrélées entre elles, la variable q_{t-1} n'est plus exogène. Les estimateurs fournis par les MCO sont alors biaisés et non convergents. Puisqu'une variable dépendante retardée apparaît parmi les variables explicatives, on ne peut pas utiliser le test de Durbin-Watson. On utilise la statistique h de Durbin (cf. chapitre 8, La fonction d'offre de viande bovine).

Tableau 11.1a Estimation de l'équation d'offre par les Moindres Carrés Ordinaires

Linear Regression — Estimation by Least Squares			
Dependent Variable Q			
Annual Data From 1964 :01 To 1995 :01			
Usable Observations	32	Degrees of Freedom	25
Centered R**2	0.764691	R Bar **2	0.708216
Uncentered R**2	0.998040	T x R**2	31.937
Mean of Dependent Variable	1695.1875000		
Std Error of Dependent Variable	157.8532265		
Standard Error of Estimate	85.2675862		
Sum of Squared Residuals	181764.03159		
Regression F(6,25)	13.5405		
Significance Level of F	0.00000084		
Durbin-Watson Statistic	2.091229		

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	385.8623543	243.1179897	1.58714	0.12504966
2.	P{1}	-0.7493947	5.7005248	-0.13146	0.89646316
3.	P{2}	13.4222979	5.8967885	2.27620	0.03166073
4.	P{3}	-8.3773407	3.7824608	-2.21479	0.03611541
5.	Q{1}	0.7348774	0.1672295	4.39443	0.00017904
6.	W{1}	-3.3287454	5.8308081	-0.57089	0.57317121
7.	D74	249.0636669	91.7274757	2.71526	0.01183243

Normal Statistic = -0.795943 with Significance Level 0.42606538

La statistique du h est égale à -0.79, au seuil de 5 %, cette valeur est inférieure à 1.64, on ne rejette donc pas l'hypothèse nulle d'absence d'autocorrélation des erreurs.

On voit qu'à court terme l'élasticité de l'offre est négative par rapport aux prix à court terme. Ainsi, à court terme, une baisse anticipée des prix par les producteurs les conduit à abattre les animaux présents sur l'exploitation. Cette baisse anticipée entraîne une augmentation de la production. Inversement, une hausse des prix les incite à conserver les animaux pour bénéficier

de prix plus élevés plus tard, donc une hausse des prix provoque une diminution des abattages, c'est-à-dire à une diminution de la production de viande. La variable indicatrice marquant le boycott du «veau aux hormones» par les consommateurs a un effet significatif.

L'équation de demande est estimée par la méthode des doubles moindres carrés. La liste des instruments est donnée par l'ensemble des variables exogènes : *constante*, p_{t-1} , p_{t-2} , p_{t-3} , q_{t-1} , w_{t-1} , $D74$, C_{t-1} , p_{mt} , r_t , s_t et s_{t-1} .

Tableau 11.1b Estimation de l'équation de demande par les doubles moindres carrés

Linear Regression — Estimation by Instrumental Variables			
Dependent Variable C			
Annual Data From 1964 :01 To 1995 :01			
Usable Observations	32	Degrees of Freedom	27
Centered R**2	0.688199	R Bar **2	0.642006
Uncentered R**2	0.999290	T x R**2	31.977
Mean of Dependent Variable	30.145531799		
Std Error of Dependent Variable	1.463126167		
Standard Error of Estimate	0.875426048		
Sum of Squared Residuals	20.692010657		
J-Specification(7)	5.046201		
Significance Level of J	0.65432498		
Durbin-Watson Statistic	1.958413		

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	- 1.53665949	10.74950677	- 0.14295	0.88738988
2.	PB	- 10.18765730	9.89300068	- 1.02978	0.31224843
3.	PM	11.36904004	4.67689607	2.43089	0.02198011
4.	R	0.08858392	0.04102152	2.15945	0.03986169
5.	C{1}	0.76025089	0.12489593	6.08708	0.00000168

Les prix présents dans la spécification ont les signes attendus : négatif pour le prix du bœuf, positif pour le prix du bien substituable (le prix du mouton). Le revenu a un effet positif significativement différent de zéro. Pour améliorer la qualité économétrique de cette estimation, il serait judicieux d'introduire le prix d'autres viandes substituables (prix de viandes blanches) ou d'estimer de manière simultanée les dépenses de consommation consacrées à différents types de viandes.

L'équation de transmission des prix est estimée par la méthode des DMC. Elle permet d'expliquer la formation du prix à la consommation.

Tableau 11.1c Estimation de l'équation de marge par les doubles moindres carrés

Linear Regression — Estimation by Instrumental Variables			
Dependent Variable PB			
Annual Data From 1964 :01 To 1995 :01			
Usable Observations	32	Degrees of Freedom	30
Centered R**2	0.358591	R Bar **2	0.337210
Uncentered R**2	0.999518	T x R**2	31.985
Mean of Dependent Variable	0.9864980236		
Std Error of Dependent Variable	0.0274799865		
Standard Error of Estimate	0.0223719754		
Sum of Squared Residuals	0.0150151585		
J-Specification(10)	15.093271		
Significance Level of J	0.12869817		
Durbin-Watson Statistic	0.712964		

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	0.9185382198	0.0219833483	41.78336	0.00000000
2.	PR	0.0537545509	0.0171045961	3.14270	0.00375220

L'équation du solde du commerce extérieur est estimée par la méthode des MCO. On utilise les mêmes instruments que précédemment.

Tableau 11.1d Estimation de l'équation du commerce extérieur par les Moindres Carrés Ordinaires

Linear Regression — Estimation by Least Squares			
Dependent Variable B			
Annual Data From 1964 :01 To 1995 :01			
Usable Observations	32	Degrees of Freedom	30
Centered R**2	0.433292	R Bar **2	0.414402
Uncentered R**2	0.781799	T x R**2	25.018
Mean of Dependent Variable	83.625000000		
Std Error of Dependent Variable	67.228426207		
Standard Error of Estimate	51.446154894		
Sum of Squared Residuals	79401.205600		
Regression F(1,30)	22.9373		
Significance Level of F	0.00004221		
Durbin-Watson Statistic	1.564417		

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	38.102349479	13.155090781	2.89640	0.00698206
2.	S{1}	0.593612395	0.123945814	4.78929	0.00004221

Le test de Durbin-Watson montre que l'on conserve l'hypothèse de non corrélation des perturbations.

La spécification utilisée pour l'équation du commerce extérieur est sommaire (le R^2 n'est pas très élevé). Une amélioration possible de l'ajustement serait de prendre en compte les conditions du marché mondial par l'introduction notamment, du prix de la viande pour des pays ayant un poids important dans les échanges de viande.

4. Les équations économétriques sont estimées de manière simultanée.

Pour estimer le modèle de manière simultanée, la méthode des TMC est utilisée. Cette méthode est sensible aux erreurs de spécification. La liste des instruments est identique à celle utilisée par les DMC.

Tableau 11.2a Estimation de l'équation d'offre par les triples moindres carrés

Linear Systems – Estimation by System Instrumental Variables

Iterations Taken	2
Annual Data From 1964 :	01 To 1995 :01
Usable Observations	32

Dependent Variable Q			
Centered R**2	0.747521	R Bar **2	0.809101
Uncentered R**2	0.997897	T x R**2	31.933
Mean of Dependent Variable	1695.1875000		
Std Error of Dependent Variable	157.8532265		
Standard Error of Estimate	68.9691652		
Sum of Squared Residuals	195026.57566		
Durbin-Watson Statistic	1.674048		

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	592.9210234	182.7115898	3.24512	0.00117401
2.	P{1}	1.3643022	4.3457485	0.31394	0.75356698
3.	P{2}	10.7876931	3.8996457	2.76633	0.00566918
4.	P{3}	-6.3234353	2.5242330	-2.50509	0.01224196
5.	Q{1}	0.5958153	0.1242884	4.79381	0.00000164
6.	W{1}	-4.2913070	4.1896511	-1.02426	0.30571081
7.	D74	177.9629162	60.7161494	2.93106	0.00337803

On retrouve pour l'équation d'offre des résultats similaires à l'estimation précédente, les variables qui n'avaient pas d'effets significativement différents de zéro (prix décalé sur une période, prix des intrants) ne sont toujours pas significatives.

Tableau 11.2b Estimation de l'équation de demande par les triples moindres carrés

Dependent Variable C			
Centered R**2	0.670657	R Bar **2	0.762567
Uncentered R**2	0.999250	T x R**2	31.976
Mean of Dependent Variable	30.145531799		
Std Error of Dependent Variable	1.463126167		
Standard Error of Estimate	0.712939615		
Sum of Squared Residuals	21.856164465		
Durbin-Watson Statistic	1.702694		

	Variable	Coeff	Std Error	T-Stat	Signif
8.	Constant	7.57595403	9.28227618	0.81617	0.41440053
9.	PB	-10.84807097	8.16413960	-1.32875	0.18393166
10.	PM	7.95695451	3.41716362	2.32853	0.01988419
11.	R	0.05877732	0.03013699	1.95034	0.05113582
12.	C{1}	0.67324523	0.10213016	6.59203	0.00000000

On obtient des résultats équivalents à la première estimation. Notons que les valeurs des R^2 n'ont pas de signification dans un modèle à équations simultanées (pour plus d'information sur ce point voir Berndt⁶, p. 468).

Tableau 11.2c Estimation de l'équation de marge par les triples moindres carrés

Dependent Variable PB			
Centered R**2	0.361899	R Bar **2	0.569976
Uncentered R**2	0.999521	T x R**2	31.985
Mean of Dependent Variable	0.9864980236		
Std Error of Dependent Variable	0.0274799865		
Standard Error of Estimate	0.0180203443		
Sum of Squared Residuals	0.0149377091		
Durbin-Watson Statistic	0.711883		

	Variable	Coeff	Std Error	T-Stat	Signif
13.	Constant	0.9167147184	0.0210554226	43.53818	0.00000000
14.	PR	0.0551968961	0.0163765886	3.37048	0.00075038

⁶ Berndt E.R. (1991), *The practice of econometrics, classic and contemporary*, Addison-Wesley.

Le prix à la production a un effet significatif sur le prix au détail.

Tableau 11.2d Estimation de l'équation du commerce extérieur par les triples moindres carrés

Dependent Variable B					
Centered R**2	0.431841	R Bar **2	0.617111		
Uncentered R**2	0.781240	T x R**2	25.000		
Mean of Dependent Variable	83.625000000				
Std Error of Dependent Variable	67.228426207				
Standard Error of Estimate	41.599648372				
Sum of Squared Residuals	79604.414257				
Durbin-Watson Statistic	1.559926				

	Variable	Coeff	Std Error	T-Stat	Signif
15.	Constant	40.736099955	12.628847997	3.22564	0.00125692
16.	S{1}	0.559268460	0.118043870	4.73780	0.00000216

On obtient des résultats équivalents à la première estimation pour l'équation du commerce extérieur.

Problème 11.3

Qualité prédictive du modèle et exemples de simulation

Pour juger de la qualité du modèle, on compare les résultats estimés par le modèle aux valeurs réelles de différentes variables. On effectue tout d'abord des simulations statiques puis des simulations dynamiques.

On a estimé le modèle jusqu'en 1995, on utilise les résultats de ces estimations pour prévoir les valeurs pour les années non utilisées dans les estimations (1996-1998). On utilise les résultats des TMC pour calculer les simulations du modèle. Pour faire exécuter le modèle il faut aussi prendre en compte l'équation décrivant l'équilibre sur le marché, équation qui permet de calculer le niveau d'équilibre des prix.

1. Dans un premier temps, on effectue des simulations statiques i.e pour les équations dynamiques du modèle (équations de demande et d'offre) on utilise les valeurs observées pour les variables décalées dans le calcul des prévisions. Les valeurs observées et simulées sont présentées sur les figures 13.3, 13.4, 13.5 et 13.6.

La courbe Q représente la quantité observée du niveau du production, la seconde courbe Q1 les valeurs prévues par le modèle. L'ajustement obtenu par le modèle est de bonne qualité, les points de retournement de la série observée sont bien reproduits. Les valeurs prédites pour la fin de période sont très proches des valeurs réelles.

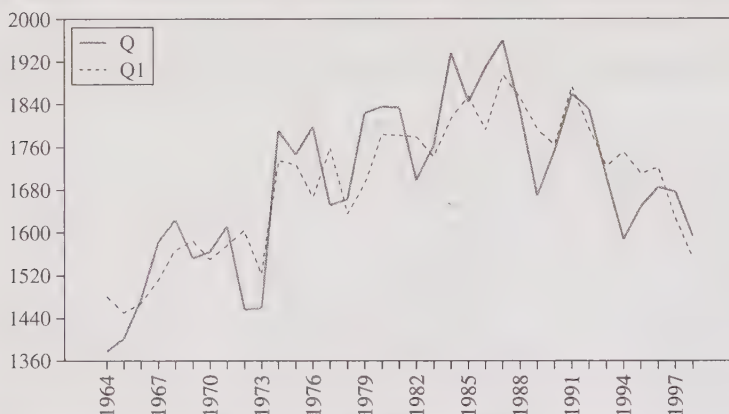


Figure 11.3 Comparaison de la production observée à la production simulée (statique)

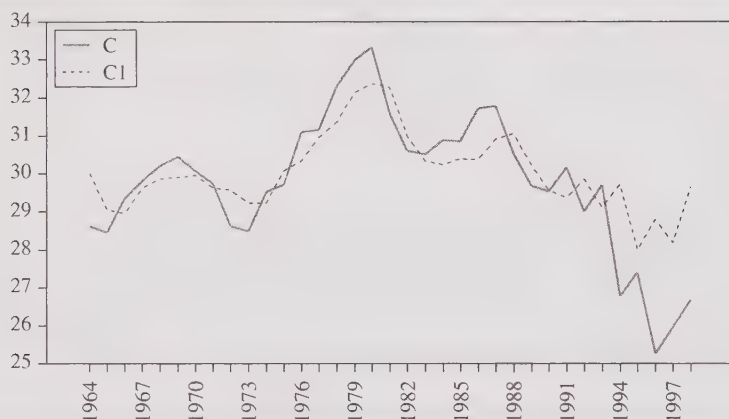


Figure 11.4 Comparaison de la consommation observée à la consommation simulée (statique)

La courbe C représente la quantité observée du niveau de consommation, la seconde courbe C1 les valeurs prévues par le modèle. On remarque que l'ajustement est de très mauvaise qualité sur la fin de période. Comme on l'a indiqué dans la partie descriptive, la fin de période d'observation marque une rupture dans les comportements de consommation de viande bovine, rupture que notre modèle n'arrive pas à reproduire. Pour améliorer la qualité prédictive du modèle on devrait retravailler la spécification de cette équation de demande. Une extension possible serait d'estimer un système de demande de viandes en utilisant par exemple un système AIDS. Cette spécification permettrait d'intégrer les substitutions-complémentarités existant entre les consommations de différents types de

viande, substitution principalement à l'origine de la variation de la consommation de viande bovine.

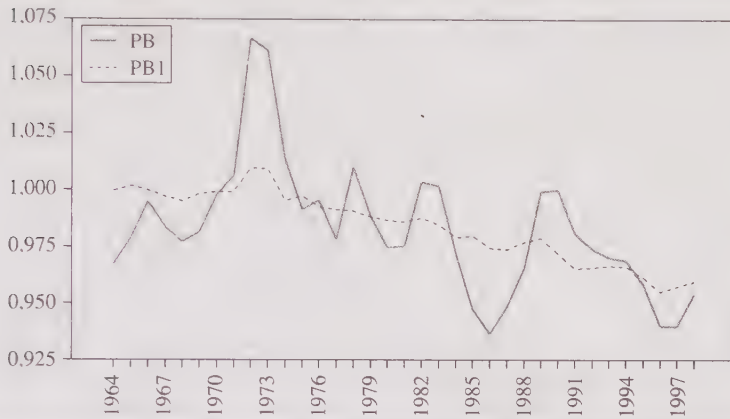


Figure 11.5 Comparaison du prix à la consommation du boeuf observé au prix simulé (statique)

La courbe PB représente le prix à la consommation observé et PB1 les valeurs prévues. La série simulée retrace la tendance à la baisse du prix, en revanche, les points de retournements de la série observée ne sont pas bien pris en compte.

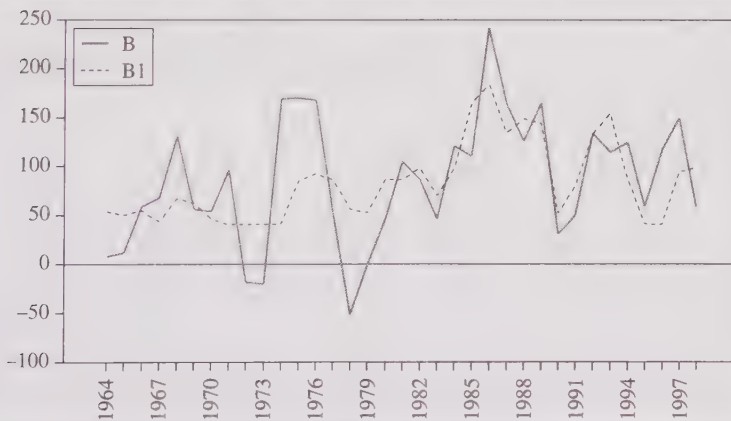


Figure 11.6 Comparaison du solde du commerce extérieur observé au solde simulé (statique)

La courbe B représente le solde du commerce extérieur observé, la seconde courbe B1 les valeurs prévues par le modèle. Même si la spécification de l'équation du commerce extérieur est très simple, on arrive à reproduire

(surtout depuis le début des années quatre-vingt) la série du solde du commerce extérieur.

2. Dans un second temps, on effectue des simulations dynamiques i.e pour les équations dynamiques du modèle (équations de demande et d'offre) on utilise les valeurs prédites par le modèle pour les variables décalées dans le calcul des prévisions.

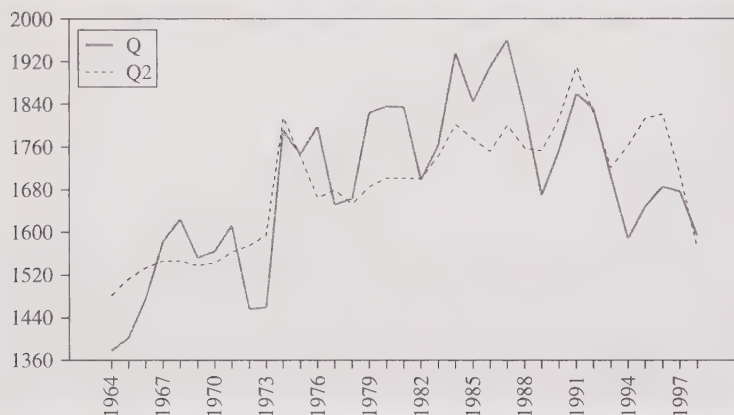


Figure 11.7 Comparaison de la production observée à la production simulée (dynamique)

La courbe Q représente la quantité observée du niveau de production, la seconde courbe Q2 les valeurs prévues par le modèle. On voit que l'ajustement est de moins bonne qualité que dans le premier cas.

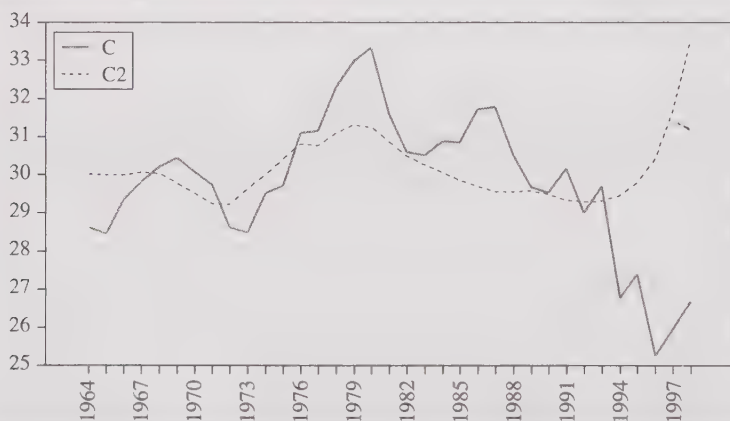


Figure 11.8 Comparaison de la consommation observée à la consommation simulée (dynamique)

La courbe C représente la quantité observée du niveau de consommation, la seconde courbe C2 les valeurs prévues par le modèle. On voit que les problèmes rencontrés lors des précédentes prédictions sont accentués dans le cas de simulations dynamiques. La rupture de la qualité de l'ajustement déjà observée pour la simulation statique est, en effet, encore plus flagrante dans le cas de cette deuxième simulation.

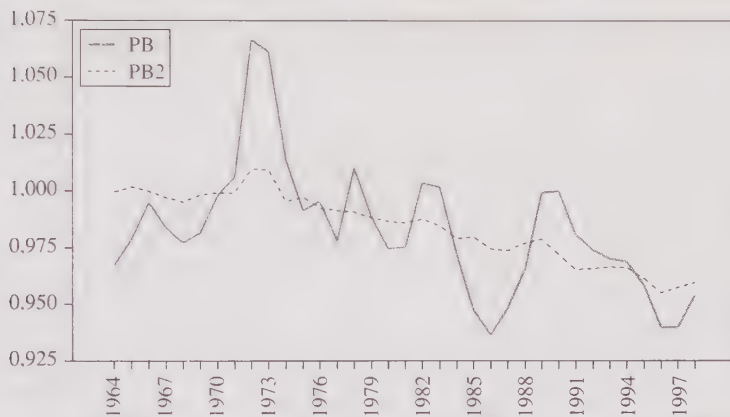


Figure 11.9 Comparaison du prix à la production du boeuf observé au prix simulé (dynamique)

La courbe PB représente la série de prix à la consommation observée et PB2 la série simulée. La simulation est identique à la simulation statique car la série de prix à la production utilisée est celle du prix observé.

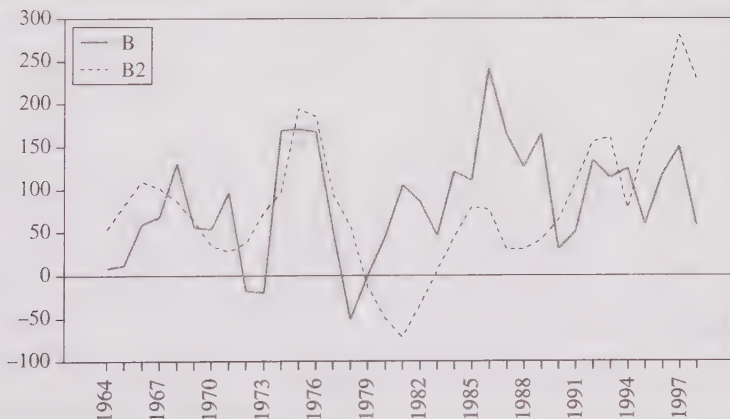


Figure 11.10 Comparaison du solde du commerce extérieur observé au solde simulé (dynamique)

La courbe B représente le solde du commerce extérieur observé, la seconde courbe B2 les valeurs prévues par le modèle. On voit là-aussi que les résultats sont de moins bonne qualité.

Les simulations dynamiques sont moins bonnes car les erreurs de prévision se reportent d'une année à l'autre. Cependant, il est important de réaliser ces simulations pour mieux apprécier la qualité prédictive du modèle. Celle-ci peut également être jugée avec le critère de Theil (tableau 13.3).

Tableau 11.3 Test de Theil

Forecast Statistics for Series Q					
Step	Mean Error	Mean Abs Error	RMS Error	Theil U	N.Obs
1	-19.39963541	59.17763440	77.89886596	0.78019	11
2	-20.29206484	64.04786373	81.63383717	0.52794	10
3	-35.60101688	58.11001486	76.62120772	0.47084	9
4	-48.13996107	57.28494964	77.98223491	0.56813	8

Forecast Statistics for Series C					
Step	Mean Error	Mean Abs Error	RMS Error	Theil U	N.Obs
1	-0.129455686	0.872100058	1.134589378	0.99346	11
2	-0.188914117	0.912797202	1.180842030	0.84843	10
3	-0.358452954	0.865670734	1.162204930	0.70014	9
4	-0.513108899	0.864030250	1.192906259	0.65621	8

Forecast Statistics for Series PB					
Step	Mean Error	Mean Abs Error	RMS Error	Theil U	N.Obs
1	-0.002789418	0.017113354	0.020824492	1.29636	11
2	0.000124796	0.015631534	0.019366543	0.72322	10
3	0.004331261	0.013175772	0.016079045	0.49154	9
4	0.008042207	0.011653206	0.014508097	0.40657	8

Forecast Statistics for Series B					
Step	Mean Error	Mean Abs Error	RMS Error	Theil U	N.Obs
1	-0.54818592	30.31746981	34.25232843	0.47712	11
2	4.94229214	27.80392013	31.35348549	0.37185	10
3	-0.96313920	24.43866968	26.78271832	0.35772	9
4	-4.81679000	23.76024499	26.37196436	0.27806	8

Forecast Statistics for Series S					
Step	Mean Error	Mean Abs Error	RMS Error	Theil U	N.Obs
1	-18.85144949	58.91712364	83.06970317	1.03465	11
2	-25.23435698	60.31107346	85.95542138	0.71823	10
3	-34.63787768	60.41260059	88.41524624	0.69754	9
4	-43.32317107	63.60861698	92.96583032	0.81660	8

Le critère de Theil (colonne Theil U dans le tableau 13.3) est une statistique souvent utilisée pour juger de l'information apportée par le modèle. Quand cette statistique est inférieure à 1 (cas en général vérifié dans l'exemple), le modèle apporte des prévisions meilleures que des prévisions naïves.

3. On présente les résultats d'une simulation d'un doublement du prix des aliments sur toute la période.

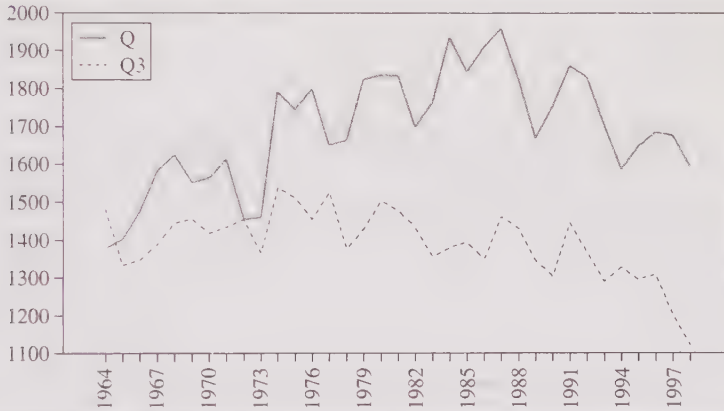


Figure 11.11 Comparaison de la production observée à la production simulée avec un prix des aliments multiplié par 2

On constate que l'augmentation du prix des aliments entraîne une forte chute de la production notée Q3 sur la figure 13.11.

12

INTRODUCTION À L'ÉCONOMÉTRIE DES DONNÉES DE PANEL

1. Présentation du modèle de base
2. Le modèle à effets fixes individuels
3. Le modèle à effets individuels aléatoires

OBJECTIFS DU CHAPITRE

Présenter les bases de l'économétrie des panels

- Modèles statiques et dynamiques à effets fixes
- Modèles statiques et dynamiques à effets aléatoires
- Estimateurs Within, Between, GMM
- Test de Fisher
- Test de Breusch-Pagan
- Test de Sargan-Hansen
- Test d'Hausman
- Test d'autocorrélation dans les modèles dynamiques

NIVEAU

MASTER, première et deuxième années

La disponibilité d'échantillons de données de panel est de plus en plus importante. Ces données ayant une dimension temporelle et individuelle permettent d'analyser la dynamique du comportement d'individus hétérogènes. De manière standard un échantillon de données de panel comporte beaucoup d'individus et peu de données temporelles.

L'objectif de ce chapitre est de présenter quelques notions de bases sur l'économétrie des données de panels : les modèles et les méthodes économétriques (pour une présentation plus complète voir Sevestre (2002)¹).

La prise en compte de l'hétérogénéité est fondamentale dans la modélisation et il existe plusieurs modèles. Dans ce chapitre nous présenterons le modèle avec des effets individuels fixes et le modèle avec des effets individuels aléatoires dans un cadre statique puis dynamique.

1. PRÉSENTATION DU MODÈLE DE BASE

Le modèle de base s'écrit

$$y_{it} = \alpha_i + \beta' x_{it} + \varepsilon_{it} \quad i = 1, \dots, N ; \quad t = 1, \dots, T \quad (12.1)$$

Avec

y_{it} : la variable endogène

β : un vecteur de K paramètres $\beta' = (\beta_1, \dots, \beta_k)$

x_{it} : un vecteur de K variables explicatives $x'_{it} = (x_{1it}, \dots, x_{kit})$ supposées strictement exogènes avec $E(x_{kit} \cdot \varepsilon_{it}) = 0$, $k = 1, \dots, K$ et $t' = 1, \dots, T$

Les aléas ε_{it} sont supposés indépendamment et identiquement distribués avec :

$$E(\varepsilon_{it} | x_{1it}, \dots, x_{Kit}) = 0 \text{ et } \text{var}(\varepsilon_{it} | x_{1it}, \dots, x_{Kit}) = \sigma_\varepsilon^2$$

Dans ce modèle les variables explicatives observables sont supposées avoir le même impact sur la variable endogène.

L'hétérogénéité entre les individus est prise en compte par α_i . Dans le modèle à effets individuels fixes α_i est une constante spécifique à chaque individu qui ne varie pas dans le temps. Elle prend en compte des caractéristiques individuelles non observables qui ont un impact direct sur y_{it} .

Dans le modèle à effets individuels aléatoires α_i est composé de 2 éléments : une constante et un aléa. Ainsi, on a $\alpha_i = \alpha + u_i$, avec α une constante et u_i un aléa spécifique à l'individu i . Le modèle (12.1) est un modèle à erreurs composées, il comporte 2 éléments aléatoires u_i et ε_{it} .

¹ P. Sevestre, 2002, Econométrie des données de panel, Dunod.

Le choix entre les 2 modèles dépend du type d'analyse. Si l'objectif est d'analyser le comportement de chaque individu en tant que tel le modèle à effets fixes sera plus adapté. En revanche, lorsque le nombre d'individus est très élevé et que l'objectif est de prédire le comportement des individus de la population dont est tiré l'échantillon, le modèle à effets aléatoires est plus pertinent.

2. LE MODÈLE À EFFETS FIXES INDIVIDUELS

2.1 Estimation du modèle

Sous forme matricielle le modèle à effets fixes s'écrit de la manière suivante :

$$Y = D\alpha + X\beta + \varepsilon \quad (12.2)$$

Avec

Y et ε des vecteurs de dimension $(NT, 1)$

$$Y = \begin{pmatrix} y_{11} \\ \cdot \\ y_{1T} \\ \cdot \\ y_{N1} \\ \cdot \\ y_{NT} \end{pmatrix} \quad \text{et} \quad \varepsilon = \begin{pmatrix} \varepsilon_{11} \\ \cdot \\ \varepsilon_{1T} \\ \cdot \\ \varepsilon_{N1} \\ \cdot \\ \varepsilon_{NT} \end{pmatrix}$$

On note

$$Y_i = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \cdot \\ y_{iT} \end{pmatrix} \quad \text{et} \quad \varepsilon_i = \begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \cdot \\ \varepsilon_{iT} \end{pmatrix}$$

D est une matrice de dimension (NT, N) composée de N variables indicatrices.

$D = (D_1, \dots, D_N)$ où chaque vecteur D_i de dimension $(NT, 1)$ représente une variable indicatrice qui prend la valeur 1 pour l'individu i et 0 pour les autres.

$$D = \begin{pmatrix} 1 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 1 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 1 \end{pmatrix} = (D_1 \ D_2 \ \dots \ D_N)$$

X est une matrice de dimension (NT, K)

$$X = \begin{pmatrix} x_{111} & x_{211} & \dots & x_{K11} \\ \cdot & \cdot & \cdot & \cdot \\ x_{11T} & x_{21T} & \dots & x_{K1T} \\ x_{121} & x_{221} & \dots & x_{K21} \\ \cdot & \cdot & \cdot & \cdot \\ x_{12T} & x_{22T} & \dots & x_{K2T} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_{1N1} & x_{2N1} & \dots & x_{KN1} \\ \cdot & \cdot & \cdot & \cdot \\ x_{1NT} & x_{2NT} & \dots & x_{KNT} \end{pmatrix}$$

α et β sont 2 vecteurs de paramètres de dimension $(N, 1)$ et $(K, 1)$ respectivement

$$\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdot \\ \cdot \\ \alpha_N \end{pmatrix} \text{ et } \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \beta_K \end{pmatrix}$$

Le modèle peut également s'écrire sous la forme suivante

$$\boxed{Y = Z\delta + \varepsilon} \tag{12.3}$$

Avec $Z = [D \ X]$ et $\delta = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$

Lorsque les aléas sont non autocorrélés et homoscédastiques la matrice de variance-covariance des aléas est donnée par

$$Var(\varepsilon|X) = E(\varepsilon\varepsilon'|X) = \sigma_\varepsilon^2 \mathcal{Z} I = \begin{pmatrix} \sigma_\varepsilon^2 I & 0 & \dots & 0 \\ 0 & \sigma_\varepsilon^2 I & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \sigma_\varepsilon^2 I \end{pmatrix}$$

La matrice est de dimension (NT, NT) .

Ce modèle peut être estimé avec la méthode des MCO (Least Square Dummy Variable).

$$\hat{\delta} = (Z'Z)^{-1} Z'Y \quad (12.4)$$

L'estimateur est sans biais et la matrice de variance-covariance est donnée par

$$\text{Var}(\hat{\delta}) = \sigma_\varepsilon^2 (Z'Z)^{-1} \quad (12.5)$$

Si les aléas sont distribués selon une loi normale l'estimateur suit également une loi normale.

Pour estimer la matrice de variance-covariance il faut $\hat{\sigma}_\varepsilon^2$

$$\hat{\sigma}_\varepsilon^2 = \frac{e'e}{NT - N - K} \text{ avec } e \text{ le vecteur des résidus MCO de dimension } (NT, 1)$$

$$e = Y - Z\hat{\delta}$$

On estime ainsi les $N + K$ paramètres du modèle comme dans un modèle de régression linéaire classique. Cependant, lorsque N est élevé il est usuel d'utiliser le théorème de Frisch-Waugh et donc de décomposer l'estimation du modèle en 2 étapes. Dans un premier temps on estime le vecteur de paramètre β et dans un second temps on déduit le vecteur α .

En effet, l'équation (12.4) peut se décomposer de la manière suivante

$$\begin{aligned} (Z'Z)\hat{\delta} = Z'Y &\Leftrightarrow \begin{pmatrix} D' \\ X' \end{pmatrix} (D \quad X) \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} D' \\ X' \end{pmatrix} Y \\ &\Leftrightarrow \begin{pmatrix} D'D & D'X \\ X'D & X'X \end{pmatrix} \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} D'Y \\ X'Y \end{pmatrix} \Leftrightarrow \begin{cases} D'D\hat{\alpha} + D'X\hat{\beta} = D'Y & \text{(a)} \\ X'D\hat{\alpha} + X'X\hat{\beta} = X'Y & \text{(b)} \end{cases} \end{aligned}$$

L'équation (a) permet d'obtenir le vecteur $\hat{\alpha}$

$$\hat{\alpha} = (D'D)^{-1} D' (Y - X\hat{\beta}) \quad (12.6)$$

Ainsi (b) devient

$$\begin{aligned} X'D (D'D)^{-1} D' (Y - X\hat{\beta}) + X'X\hat{\beta} &= X'Y \\ \Leftrightarrow X' (I - D (D'D)^{-1} D') X\hat{\beta} &= X' (I - D (D'D)^{-1} D') Y \end{aligned} \quad (c)$$

Soit $M_d = I - D(D'D)^{-1}D'$, matrice symétrique et idempotent $M_d = M'_d = M^2_d$ de dimension (NT, NT) .

L'équation (c) s'écrit : $X'M_dX\hat{\beta} = X'M_dY$

On obtient ainsi $\hat{\beta}$

$$\boxed{\hat{\beta} = (X'M_dX)^{-1} X'M_dY} \tag{12.7}$$

Pour estimer le vecteur $\hat{\delta}$ on peut soit (si le nombre d'observation le permet) estimer directement avec les MCO en une étape soit procéder à l'estimation de $\hat{\beta}$ avec (12.7) puis avec (12.6) calculer $\hat{\alpha}$. Le développement de l'équation (12.6) pour une observation i donne

$$\hat{\alpha}_i = \bar{y}_i - \hat{\beta}_1\bar{x}_{1i} - \hat{\beta}_2\bar{x}_{2i} - \dots - \hat{\beta}_K\bar{x}_{Ki}, \quad i = 1, \dots, N$$

Étant donné les caractéristiques de la matrice M_d on note que $\hat{\beta}$ est l'estimateur des MCO du modèle

$$M_dY = M_dX\beta + M_d\varepsilon$$

Il s'agit de l'estimateur Within autrement dit de l'estimateur des MCO sur les données centrées par rapport à leur moyenne individuelle.

$\hat{\beta} = \hat{\beta}_w$ avec $\hat{\beta}_w$ l'estimateur Within.

En effet, on montre que

$$M_d = \begin{pmatrix} M_O & 0 & \dots & 0 \\ 0 & M_O & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \vdots & \dots & M_O \end{pmatrix} = I \otimes M_O$$

La matrice M_O est de dimension (T, T)

$$M_O = I - \frac{1}{T}ii'$$

Avec i un vecteur de dimension $(T, 1)$ de 1 et

$$ii' = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} (1 \quad 1 \quad \dots \quad 1) = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix}$$

Le vecteur $M_O Y_i$ est donné par

$$M_O Y_i = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{pmatrix} - \frac{1}{T} \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{pmatrix}$$

Soit

$$M_O Y_i = \begin{pmatrix} y_{i1} - \bar{y}_i \\ y_{i2} - \bar{y}_i \\ \vdots \\ y_{iT} - \bar{y}_i \end{pmatrix} \text{ avec } \bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}$$

D'où

$$M_d Y = \begin{pmatrix} y_{11} - \bar{y}_1 \\ \vdots \\ y_{1T} - \bar{y}_1 \\ \vdots \\ y_{N1} - \bar{y}_N \\ \vdots \\ y_{NT} - \bar{y}_N \end{pmatrix}$$

De la même manière

$$M_d \varepsilon = \begin{pmatrix} \varepsilon_{11} - \bar{\varepsilon}_1 \\ \vdots \\ \varepsilon_{1T} - \bar{\varepsilon}_1 \\ \vdots \\ \varepsilon_{N1} - \bar{\varepsilon}_N \\ \vdots \\ \varepsilon_{NT} - \bar{\varepsilon}_N \end{pmatrix} \text{ et } M_d X = \begin{pmatrix} x_{111} - \bar{x}_{11} & x_{211} - \bar{x}_{21} & \dots & x_{K11} - \bar{x}_{K1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{11T} - \bar{x}_{11} & x_{21T} - \bar{x}_{21} & \dots & x_{K1T} - \bar{x}_{K1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1N1} - \bar{x}_{1N} & x_{2N1} - \bar{x}_{2N} & \dots & x_{KN1} - \bar{x}_{KN} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1NT} - \bar{x}_{1N} & x_{2NT} - \bar{x}_{2N} & \dots & x_{KNT} - \bar{x}_{KN} \end{pmatrix}$$

avec $\bar{\varepsilon}_i = \frac{1}{T} \sum_{t=1}^T \varepsilon_{it}$ et avec $\bar{x}_{ki} = \frac{1}{T} \sum_{t=1}^T x_{kit}$

L'estimateur Within est donc l'estimateur des MCO sur le modèle centré

$$y_{it} - \bar{y}_i = \beta_1 (x_{1it} - \bar{x}_{1i}) + \beta_2 (x_{2it} - \bar{x}_{2i}) + \dots + \beta_K (x_{Kit} - \bar{x}_{Ki}) + (\varepsilon_{it} - \bar{\varepsilon}_i)$$

$i = 1, \dots, N$ et $t = 1, \dots, T$

Sous l'hypothèse de stricte exogénéité des variables explicatives, lorsque N tend vers l'infini, T restant fini, l'estimateur est convergent si $\lim_{N \rightarrow \infty} \frac{X' M_d X}{N} = M_d^N$, M_d^N étant une matrice définie positive.

2.2 Test de la présence d'effets individuels fixes

Dans le modèle à effets fixes individuels

$$y_{it} = \alpha_i + \beta' x_{it} + \varepsilon_{it} \quad i = 1, \dots, N ; \quad t = 1, \dots, T$$

α_i permet de spécifier l'hétérogénéité des individus.

Si les individus sont totalement homogènes alors α_i est égal à une constante α quelque soit i . Le modèle dans ce cas s'écrit :

$$y_{it} = \alpha + \beta' x_{it} + \varepsilon_{it} \quad i = 1, \dots, N ; \quad t = 1, \dots, T \quad (12.8)$$

Il faut alors estimer $K + 1$ paramètres au lieu de $K + N$.

Pour tester $H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_N = \alpha$ on peut appliquer un test de contraintes de Fisher.

Le modèle à effets fixes individuels correspond au modèle sans contraintes et le modèle (12.8) est un modèle qui comporte $N - 1$ contraintes.

Si on note $e'e$ la somme des carrés des résidus du modèle sans contraintes et $e*'e^*$ la somme des carrés des résidus du modèle avec contraintes la statistique de Fisher est donnée par

$$F = \frac{(e'e - e*'e^*)/N - 1}{e'e/NT - N - K} \quad (12.9)$$

Sous l'hypothèse nulle la statistique F est distribuée selon une loi de Fisher à $(N - 1, NT - N - K)$ degrés de liberté. Si la statistique est inférieure à la valeur tabulée alors l'hypothèse nulle est acceptée. Sinon, elle est refusée et on doit retenir le modèle à effets fixes. Cependant, il faut dans ce cas être vigilant car le rejet de l'hypothèse indique qu'il y a hétérogénéité des individus.

Sous l'hypothèse nulle le modèle (12.8) peut être estimé avec la méthode des MCO. L'estimateur ainsi obtenu peut s'écrire comme une moyenne pondérée des estimateurs Within et Between. Il est alors possible d'analyser les résultats par rapport aux variations intra et inter-individus². Soit $\hat{\beta}_{MCO}$ l'estimateur MCO du modèle centré

$$y_{it} - \bar{y} = \beta_1 (x_{1it} - \bar{x}_1) + \dots + \beta_K (x_{Kit} - \bar{x}_K) + \beta_1 (\varepsilon_{it} - \bar{\varepsilon})$$

² Dans les données de panel on dispose d'observations individuelles et temporelles. La variabilité totale peut se décomposer entre la variabilité inter-individus et la variabilité intra-individus $\sum_i \sum_t (y_{it} - \bar{y})^2 = \sum_i \sum_t (y_{it} - \bar{y}_i)^2 + T \sum_i (\bar{y}_i - \bar{y})^2$

$$\text{Avec } \bar{y} = \frac{1}{NT} \sum_{t=1}^T \sum_{i=1}^N y_{it}, \bar{x}_k = \frac{1}{NT} \sum_{t=1}^T \sum_{i=1}^N x_{kit}, \bar{\varepsilon} = \frac{1}{NT} \sum_{t=1}^T \sum_{i=1}^N \varepsilon_{it}$$

Soient

\bar{X} , X^w et X^b respectivement les matrices des données centrées $(x_{kit} - \bar{x}_k)$, $(x_{kit} - \bar{x}_{ki.})$, $(\bar{x}_{ki.} - \bar{x}_k)$ de dimension $(NT, 1)$, $(NT, 1)$ et $(N, 1)$.

\bar{Y} , Y^w et Y^b respectivement les matrices des données centrées $(y_{it} - \bar{y}_i)$, $(x_{kit} - \bar{x}_{ki.})$, $(\bar{y}_i - \bar{y})$ de dimension $(NT, 1)$, $(NT, 1)$ et $(N, 1)$.

On a les égalités

$$\bar{X}'\bar{X} = X^{w'}X^w + TX^{b'}X^b$$

$$\bar{X}'\bar{Y} = X^{w'}Y^w + TX^{b'}Y^b$$

$$\hat{\beta}_{MCO} = (\bar{X}'\bar{X})^{-1} \bar{X}'\bar{Y} = (X^{w'}X^w + TX^{b'}X^b)^{-1} (X^{w'}Y^w + TX^{b'}Y^b)$$

$$\text{Or } \hat{\beta}_w = (X^{w'}X^w)^{-1} (X^{w'}Y^w) \text{ et } \hat{\beta}_b = (X^{b'}X^b)^{-1} (X^{b'}Y^b)$$

$\hat{\beta}_b$ étant l'estimateur Between.

Par conséquent

$$(X^{w'}X^w) \hat{\beta}_w = (X^{w'}Y^w) \text{ et } (X^{b'}X^b) \hat{\beta}_b = (X^{b'}Y^b)$$

D'où

$$\hat{\beta}_{MCO} = (X^{w'}X^w + TX^{b'}X^b)^{-1} \left((X^{w'}X^w) \hat{\beta}_w + T (X^{b'}X^b) \hat{\beta}_b \right)$$

Soit $F^w = (X^{w'}X^w + TX^{b'}X^b)^{-1} (X^{w'}X^w)$ la matrice de pondération d'ordre (K, K)

$$\boxed{\hat{\beta}_{MCO} = F^w \hat{\beta}_w + (I - F^w) \hat{\beta}_b} \quad (12.10)$$

2.3 Généralisation du modèle à effets fixes

Le modèle à effets fixes individuels peut être généralisé pour inclure des effets fixes temporels et une constante³ :

$$\boxed{y_{it} = \lambda + \alpha_i + \phi_t + \beta' x_{it} + \varepsilon_{it} \quad i = 1, \dots, N; \quad t = 1, \dots, T} \quad (12.11)$$

³ On peut également estimer un modèle à effets fixes temporels $y_{it} = \phi_t + \beta' x_{it} + \varepsilon_{it}$, le traitement dans ce cas est totalement similaire à celui des effets fixes individuels, l'estimateur within est obtenu avec des données centrées par rapport à leur moyenne temporelle :

$$\bar{y}_{.t} = \frac{1}{N} \sum_{i=1}^N y_{it} \text{ et } \bar{x}_{k.t} = \frac{1}{N} \sum_{i=1}^N x_{kit}$$

Pour estimer ce modèle, on doit imposer des contraintes d'identification

$$\sum_i \alpha_i = \sum_t \phi_t = 0$$

On obtient $\widehat{\beta}$ à partir de l'estimation du modèle centré

$$y_{it}^* = \beta' x_{it}^* + \varepsilon_{it}^* \quad i = 1, \dots, N ; \quad t = 1, \dots, T$$

Avec

$$\begin{aligned} y_{it}^* &= y_{it} - \bar{y}_{i.} - \bar{y}_{.t} - \bar{y} \\ x_{kit}^* &= x_{kit} - \bar{x}_{ki.} - \bar{x}_{k.t} - \bar{x}_k \\ \varepsilon_{it}^* &= \varepsilon_{it} - \bar{\varepsilon}_{i.} - \bar{\varepsilon}_{.t} - \bar{\varepsilon} \end{aligned}$$

Les moyennes temporelles étant données par

$$\bar{y}_{.t} = \frac{1}{N} \sum_{i=1}^N y_{it}, \quad \bar{x}_{k.t} = \frac{1}{N} \sum_{i=1}^N x_{kit} \quad \text{et} \quad \bar{\varepsilon}_{.t} = \frac{1}{N} \sum_{i=1}^N \varepsilon_{it}$$

Ayant estimé $\widehat{\beta}$ on peut en déduire le coefficient de la constante et les effets fixes individuels et temporels

$$\begin{aligned} \widehat{\lambda} &= \bar{y} - \widehat{\beta}_1 \bar{x}_1 - \widehat{\beta}_2 \bar{x}_2 - \dots - \widehat{\beta}_K \bar{x}_K \\ \widehat{\alpha}_i &= (\bar{y}_{i.} - \bar{y}) - \widehat{\beta}_1 (\bar{x}_{1i.} - \bar{x}_1) - \widehat{\beta}_2 (\bar{x}_{2i.} - \bar{x}_2) - \dots - \widehat{\beta}_K (\bar{x}_{Ki.} - \bar{x}_k) \\ \widehat{\phi}_t &= (\bar{y}_{.t} - \bar{y}) - \widehat{\beta}_1 (\bar{x}_{1.t} - \bar{x}_1) - \widehat{\beta}_2 (\bar{x}_{2.t} - \bar{x}_2) - \dots - \widehat{\beta}_K (\bar{x}_{K.t} - \bar{x}_k) \end{aligned}$$

2.4 Modèles à effets individuels fixes dynamiques

Le modèle dynamique à effets fixes individuels s'écrit

$$y_{it} = \alpha_i + \beta' x_{it} + \lambda y_{it-1} + \varepsilon_{it} \quad i = 1, \dots, N ; \quad t = 1, \dots, T \quad (12.12)$$

On a vu que l'estimateur MCO des coefficients β et λ est identique à l'estimateur Within c'est-à-dire à l'estimateur MCO des coefficients du modèle centré par rapport aux moyennes individuelles

$$\begin{aligned} y_{it} - \bar{y}_{i.} &= \beta_1 (x_{1it} - \bar{x}_{1i.}) + \beta_2 (x_{2it} - \bar{x}_{2i.}) + \dots + \beta_K (x_{Kit} - \bar{x}_{Ki.}) \\ &\quad + \lambda (y_{it-1} - \bar{y}_{i.-1}) + (\varepsilon_{it} - \bar{\varepsilon}_{i.}) \\ i &= 1, \dots, N \quad \text{et} \quad t = 1, \dots, T \end{aligned} \quad (12.13)$$

$$\text{Avec } \bar{y}_{i,-1} = \frac{1}{T-1} \sum_{t=1}^{T-1} y_{it}$$

Dans ce modèle $\bar{y}_{i,-1}$ par construction est corrélé à $\bar{\varepsilon}_i$, et donc à $\varepsilon_{it} - \bar{\varepsilon}_i$. En effet $\bar{\varepsilon}_i$ contient ε_{it-1} qui est corrélé à y_{it-1} .

Par conséquent l'estimateur Within est biaisé⁴.

Les méthodes d'estimation usuelles se basent sur la méthode des Variables Instrumentales et la méthode des Moments Généralisés.

Anderson et Hsiao suggèrent d'estimer avec la méthode des Variables Instrumentales le modèle en différence première

$$\begin{aligned} y_{it} - y_{it-1} &= \beta_1 (x_{1it} - x_{1it-1}) + \beta_2 (x_{2it} - x_{2it-1}) + \dots + \beta_K (x_{Kit} - x_{Kit-1}) + \\ &\quad \lambda (y_{it-1} - y_{it-2}) + (\varepsilon_{it} - \varepsilon_{it-1}) \\ i &= 1, \dots, N \text{ et } t = 3, \dots, T \end{aligned} \tag{12.14}$$

On note $v_{it} = \varepsilon_{it} - \varepsilon_{it-1}$ les aléas du modèle en différence première. v_{it} suit un processus MA(1) lorsque les aléas ε_{it} ne sont pas autocorrélés.

$$\begin{aligned} E(v_{it} | x_{1it}, \dots, x_{Kit}) &= 0 & E(v_{it}^2 | x_{1it}, \dots, x_{Kit}) &= 2\sigma_\varepsilon^2 \\ E(v_{it}v_{it-1} | x_{1it}, \dots, x_{Kit}) &= -\sigma_\varepsilon^2 & E(v_{it}v_{it} | x_{1it}, \dots, x_{Kit}) &= 0 \quad t \neq t \text{ et } t \neq t-1 \end{aligned}$$

Par conséquent la matrice de variance covariance des aléas du modèle en différence s'écrit

$$\text{Var}(v | X) = \sigma_\varepsilon^2 (I \otimes \Phi) = \sigma_\varepsilon^2 \Pi = \Gamma$$

Avec

$$\Phi = \begin{pmatrix} 2 & -1 & . & . & . & . & 0 \\ -1 & 2 & -1 & . & . & . & 0 \\ 0 & -1 & 2 & -1 & . & . & . \\ . & . & -1 & 2 & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & -1 \\ 0 & . & . & . & . & -1 & 2 \end{pmatrix}$$

Φ est de dimension $(T-2, T-2)$ et I est une matrice identité d'ordre (N, N) .

Les variables y_{it-2} et $y_{it-2} - y_{it-3}$ sont corrélées $y_{it} - y_{it-1}$ mais non corrélées à v_{it} . Anderson et Hsiao proposent d'utiliser soit y_{it-2} soit $y_{it-2} - y_{it-3}$ comme variable instrumentale.

⁴ Il est convergent lorsque N et T tendent vers l'infini.

On note X^D la matrice des variables explicative du modèle en différence première, elle est de dimension $(N(T - 2), K + 1)$

$$X^D = [\Delta X \quad \Delta Y_{-1}]$$

La matrice des instruments Z^D est donnée par

$$Z^D = [\Delta X \quad Y_{-2}] \text{ ou } Z^D = [\Delta X \quad \Delta Y_{-2}]$$

Elle est de la même dimension que X^D .

L'estimateur est donné par

$$\boxed{\begin{bmatrix} \widehat{\beta} \\ \widehat{\lambda} \end{bmatrix}} = (X^{D'} P_{Z^D} X^D)^{-1} (X^{D'} P_{Z^D} \Delta Y) \tag{12.15}$$

Avec

$$P_{Z^D} = Z^D (Z^{D'} Z^D)^{-1} Z^{D'}$$

X^D et Z^D étant de la même dimension on a

$$\boxed{\begin{bmatrix} \widehat{\beta} \\ \widehat{\lambda} \end{bmatrix}} = (Z^{D'} X^D)^{-1} (Z^{D'} \Delta Y) \tag{12.16}$$

Ces estimateurs sont très faciles à mettre en œuvre mais dans la pratique les résultats obtenus ne sont pas satisfaisants. Leur précision est faible et ceci en raison, en outre, de la non prise en compte de l'autocorrélation des aléas v_{it} et du nombre d'instruments utilisés (1 seul)⁵.

Arrelano et Bond proposent un estimateur de la méthode des Moments Généralisés (GMM). L'estimateur prend en compte le problème d'autocorrélation des aléas du modèle en différence et proposent un ensemble d'instruments plus important qu'Anderson et Hsiao.

Dans le modèle en différence première toutes les données sont observées à partir de $t = 3$

$$t = 3$$

$$y_{i3} - y_{i2} = \beta_1 (x_{1i3} - x_{1i2}) + \beta_2 (x_{2i3} - x_{2i2}) + \dots + \beta_K (x_{Ki3} - x_{Ki2}) + \lambda (y_{i2} - y_{i1}) + (\varepsilon_{i3} - \varepsilon_{i2})$$

$$i = 1, \dots, N$$

⁵ Si x_{it} est un vecteur de K variables explicatives $x'_{it} = (x_{1it}, \dots, x_{kit})$ faiblement exogènes avec $E(x_{kit} \cdot \varepsilon_{it}) = 0, k = 1, \dots, K$ et $t \leq t'$ on peut également utiliser comme variables instrumentales X_{-2} ou ΔX_{-2} dans ce cas $Z^D = [\Delta X \quad \Delta Y_{-2} \quad X_{-2}]$ ou $Z^D = [\Delta X \quad \Delta Y_{-2} \quad \Delta X_{-2}]$.

Dans un premier temps ils estiment le modèle en différence

$$\begin{bmatrix} \hat{\beta} \\ \hat{\lambda} \end{bmatrix} = (X^{D'} P_{\Pi} X^D)^{-1} (X^{D'} P_{\Pi} \Delta Y)$$

Avec

$$P_{\Pi} = Z^D (Z^D \Pi Z^D)^{-1} Z^{D'}$$

Dans un second temps, ils utilisent les résidus de la première estimation pour estimer Γ

$$\boxed{\begin{bmatrix} \hat{\beta} \\ \hat{\lambda} \end{bmatrix}} = (X^{D'} P_{\Gamma} X^D)^{-1} (X^{D'} P_{\Gamma} \Delta Y) \quad (12.17)$$

Avec

$$P_{\Gamma} = Z^D (Z^{D'} \hat{\Gamma} Z^D)^{-1} Z^{D'}$$

Cette procédure en 2 étapes permet de ne pas avoir à spécifier de manière précise la forme de l'autocorrélation et de l'hétéroscadasticité des aléas v_{it} .

2.5 Tests sur les modèles à effets individuels fixes dynamiques

1. Test de validité des instruments

La validité des instruments peut être testée avec le test de Sargan-Hansen. Dans la méthode des variables instrumentales et des moments généralisés il est fondamental que les instruments soient non corrélés aux aléas du modèle. Dans le test de Sargan-Hansen on teste l'hypothèse nulle que les instruments ne sont pas asymptotiquement corrélés aux aléas. La statistique du test est distribuée selon une loi du χ^2 sous réserve que les aléas ont été corrigés des problèmes d'autocorrélation et d'hétéroscadasticité.

Soit \hat{v} le vecteur des résidus du modèle estimé de dimension $(N(T-2), 1)$, la statistique est donnée par

$$\boxed{Q_S = \hat{v}' P_{\Gamma} \hat{v}} \quad (12.18)$$

Elle converge en loi vers un χ^2 à $p-k-1$ degrés de liberté p correspondant aux nombres de variables instrumentales autrement dit au nombre de colonnes de la matrice des instruments Z^d .

2. Test d'absence d'effets individuels

Pour tester la présence d'effets individuels dans le modèle dynamique lorsque les variables explicatives x sont supposées strictement exogènes et les aléas homoscédastiques et non autocorrélés on peut appliquer le test d'Hausman.

Sous l'hypothèse nulle pas d'effets individuels l'estimateur MCO est le meilleur sinon on doit utiliser un estimateur des Variables Instrumentales ou GMM. On calcule la statistique

$$W = \left(\widehat{\delta}_{VI/GMM} - \widehat{\delta}_{MCO} \right)' \left(\widehat{Var}(\widehat{\delta}_{VI/GMM}) - \widehat{Var}(\widehat{\delta}_{MCO}) \right)^{-1} \left(\widehat{\delta}_{VI/GMM} - \widehat{\delta}_{MCO} \right) \quad (12.19)$$

Avec $\widehat{\delta} = \begin{bmatrix} \widehat{\beta} \\ \widehat{\lambda} \end{bmatrix}$

W est distribuée selon une loi du χ^2 .

3. Tests d'autocorrélation

Dans les modèles de panel dynamiques les méthodes des Variables Instrumentales et des Moments Généralisés préconisent l'emploi des variables décalées comme variables instrumentales. Les instruments sont valides si les aléas du modèle ne sont pas autocorrélés.

Si les aléas ε_{it} ne sont pas autocorrélés alors les aléas du modèle en différence première v_{it} suivent un processus MA(1) et les variables décalées $y_{it-2}, y_{it-3}, y_{it-4}, \dots$ sont des instruments valides. En revanche si les aléas ε_{it} sont autocorrélés alors les aléas du modèle en différence première v_{it} suivent un processus MA(2) les variables $y_{it-3}, y_{it-4}, \dots$ restent des instruments valides mais pas y_{it-2} .

Pour tester l'autocorrélation d'ordre 1 des aléas ε_{it} on peut utiliser la statistique de Sargan/Hansen. Cette statistique doit être calculée, d'une part avec les résidus d'estimation obtenus lorsque parmi les instruments figure la variable y_{it-2} elle est notée Q_{S0} et, d'autre part, avec les résidus d'estimation obtenus lorsque parmi les instruments ne figure pas la variable y_{it-2} elle est notée Q_{S1} . Si une différence notable entre les 2 statistiques de Sargan apparaît alors on considère que la variable y_{it-2} est corrélée à l'aléa v_{it} en raison d'une autocorrélation des aléas ε_{it} . Dans le cas de la méthode de Arrellano et Bond la statistique est distribuée selon un χ^2 à $p_0 - p_1$ degrés de liberté.

Arrellano et Bond proposent également de calculer la statistique m_2 qui teste l'hypothèse d'absence d'autocorrélation à l'ordre 2 des aléas v_{it} , sous l'hypothèse alternative les aléas v_{it} ne suivent pas un processus MA(1) et par conséquent les aléas ε_{it} sont autocorrélés. Sous l'hypothèse nulle la statistique m_2 est distribuée asymptotiquement selon une loi normale. Lorsqu'elle est inférieure en valeur absolue à 1,64 on l'accepte.

Soit $X^* = [X \ Y_{-1}]$

$$m_2 = \frac{\widehat{v}'_{-2} \widehat{v}^r}{\varphi} \quad (12.20)$$

Avec

$$\begin{aligned} \varphi &= \sum_i \widehat{v}'_{i,-2} \widehat{v}_i^r \widehat{v}_i^{r'} \widehat{v}_{i,-2} \\ &\quad - 2 \widehat{v}'_{-2} X^{*r} (X^{*'} P_{\Gamma} X^*)^{-1} X' Z^D (Z^{D'} \widehat{\Gamma} Z^D)^{-1} \sum_i Z_i^{D'} \widehat{v}_i \widehat{v}_i^{r'} \widehat{v}_{i,-2} \\ &\quad + \widehat{v}'_{-2} X^{*r} Var_{as}(\widehat{\delta}) X^{*r'} \widehat{v}_{-2} \end{aligned}$$

Avec \widehat{v}_{-2} le vecteur des résidus décalé de 2 périodes et \widehat{v}^r le vecteur des résidus dans lequel on omet les 2 premières observations temporelles afin de calculer la somme des carrés des résidus $\widehat{v}'_{-2}\widehat{v}^r$. Dans la formule l'exposant r indique que les 2 premières observations temporelles sont omises pour effectuer le calcul.

3. LE MODÈLE À EFFETS INDIVIDUELS ALÉATOIRES

3.1 Estimation du modèle

Le modèle à effets individuels aléatoires ou modèle à erreurs composées s'écrit

$$y_{it} = \alpha + \beta'x_{it} + \mu_{it} \quad i = 1, \dots, N; \quad t = 1, \dots, T \quad (12.21)$$

Avec : $\mu_{it} = u_i + \varepsilon_{it}$

Les variables explicatives du modèle sont supposées strictement exogènes.

L'aléa est composé de 2 termes d'erreurs non corrélés entre eux :

- un aléa stable dans le temps u_i qui prend en compte l'influence des variables omises dans le modèle, ces variables sont spécifiques à chaque individu mais sont stables dans le temps. On suppose que

$$E(u_i | x_{1it}, \dots, x_{Kit}) = 0, \quad E(u_i^2 | x_{1it}, \dots, x_{Kit}) = \sigma_u^2, \\ E(u_i u_j | x_{1it}, \dots, x_{Kit}) = 0 \quad i \neq j$$

- un aléa ε_{it} qui prend en compte les variables omises propres à chaque individu qui varient dans le temps. On suppose que

$$E(\varepsilon_{it} | x_{1it}, \dots, x_{Kit}) = 0, \quad E(\varepsilon_{it}^2 | x_{1it}, \dots, x_{Kit}) = \sigma_\varepsilon^2, \\ E(\varepsilon_{it}\varepsilon_{jt} | x_{1it}, \dots, x_{Kit}) = 0 \quad i \neq j \text{ ou } t \neq t'$$

Ainsi

$$E(\mu_{it} | x_{1it}, \dots, x_{Kit}) = 0, \quad E(\mu_{it}^2 | x_{1it}, \dots, x_{Kit}) = \sigma_u^2 + \sigma_\varepsilon^2, \\ E(\mu_{it}\mu_{it'} | x_{1it}, \dots, x_{Kit}) = \sigma_u^2 \quad t \neq t', \quad E(\mu_{it}\mu_{jt'} | x_{1it}, \dots, x_{Kit}) = 0 \quad i \neq j \text{ et } t \neq t'$$

Soit μ_i de dimension $(T, 1)$ le vecteur des aléas, $i = 1, \dots, N$

$$\mu_i = \begin{pmatrix} \mu_{i1} \\ \mu_{i2} \\ \vdots \\ \mu_{iT} \end{pmatrix}$$

La matrice de variance-covariance de μ_i est de dimension (T, T) et s'écrit

$$\begin{aligned} \text{Var}(\mu_i | x_{1i}, \dots, x_{Ki}) &= E(\mu_i \mu_i' | x_{1i}, \dots, x_{Ki}) \\ &= \begin{pmatrix} \sigma_u^2 + \sigma_\varepsilon^2 & \sigma_u^2 & \cdot & \cdot & \sigma_u^2 \\ \sigma_u^2 & \sigma_u^2 + \sigma_\varepsilon^2 & \cdot & \cdot & \sigma_u^2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \sigma_u^2 & \cdot & \cdot & \cdot & \sigma_u^2 + \sigma_\varepsilon^2 \end{pmatrix} = \sigma_\varepsilon^2 I_T + \sigma_u^2 i i' = \Omega \end{aligned}$$

La matrice des covariances entre μ_i et μ_j est égale à 0.

Soit μ le vecteur des aléas du modèle (12.21)

$$\mu = \begin{pmatrix} \mu_{11} \\ \cdot \\ \cdot \\ \mu_{1T} \\ \cdot \\ \cdot \\ \mu_{N1} \\ \cdot \\ \cdot \\ \mu_{NT} \end{pmatrix}$$

La matrice de variance-covariance des aléas est de dimension (NT, NT) , elle est donnée par

$$\text{Var}(\mu | X) = E(\mu \mu' | X) = \begin{pmatrix} \Omega & 0 & \cdot & \cdot & 0 \\ 0 & \Omega & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \Omega \end{pmatrix} = I \otimes \Omega = \Lambda$$

Sous forme matricielle le modèle s'écrit

$$\boxed{Y = X^+ \beta^+ + \mu} \tag{12.22}$$

Avec X^+ la matrice des variables explicatives de dimension $(NT, K + 1)$

$$X^+ = \begin{pmatrix} 1 & x_{111} & \cdot & \cdot & x_{K11} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{11T} & \cdot & \cdot & x_{K1T} \\ 1 & x_{121} & \cdot & \cdot & x_{K21} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{12T} & \cdot & \cdot & x_{K2T} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{1N1} & \cdot & \cdot & x_{KN1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{1NT} & \cdot & \cdot & x_{KNT} \end{pmatrix}$$

Et β^+ le vecteur des paramètres à estimer de dimension $(K + 1, 1)$

$$\beta^+ = \begin{pmatrix} \alpha \\ \beta_1 \\ \vdots \\ \beta_K \end{pmatrix}$$

Les aléas du modèle ne sont pas hétéroscédastiques, il n'y a pas de corrélation des aléas entre les individus mais par individu il y a une autocorrélation des aléas.

Parmi les estimateurs non biaisés de β^+ l'estimateur des MCG est le plus précis.

Pour estimer le modèle (12.22) avec la méthode MCG il faut déterminer la matrice de transformation $\Lambda^{-1/2}$ et donc la matrice $\Omega^{-1/2}$.

Cette matrice est donnée par

$$\Omega^{-1/2} = I - \frac{\theta}{T} ii' \text{ avec } \theta = 1 - \frac{\sigma_\varepsilon}{\sqrt{T\sigma_u^2 + \sigma_\varepsilon^2}}$$

La transformation pour Y_i est

$$\begin{aligned} \Omega^{-1/2} Y_i &= \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{pmatrix} - \frac{\theta}{T} \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{pmatrix} \\ &= \begin{pmatrix} y_{i1} - \theta \bar{y}_i \\ y_{i2} - \theta \bar{y}_i \\ \vdots \\ y_{iT} - \theta \bar{y}_i \end{pmatrix} \end{aligned}$$

De la même manière pour X_i^+

$$\Omega^{-1/2} X_i^+ = \begin{pmatrix} 1 - \theta & x_{1i1} - \theta \bar{x}_{1i} & \dots & x_{Ki1} - \theta \bar{x}_{Ki} \\ 1 - \theta & x_{1i2} - \theta \bar{x}_{1i} & \dots & x_{Ki2} - \theta \bar{x}_{Ki} \\ \vdots & \vdots & \ddots & \vdots \\ 1 - \theta & x_{1iT} - \theta \bar{x}_{1i} & \dots & x_{KiT} - \theta \bar{x}_{Ki} \end{pmatrix}$$

L'estimateur des MCG est obtenu en estimant avec la méthode des MCO le modèle transformé suivant⁷

$$y_{it} - \theta \bar{y}_i = \alpha(1 - \theta) + \beta_1(x_{1it} - \theta \bar{x}_{1i}) + \beta_2(x_{2it} - \theta \bar{x}_{2i}) + \dots + \beta_K(x_{Kit} - \theta \bar{x}_{Ki}) + (\mu_{it} - \theta \bar{\mu}_i)$$

$$i = 1, \dots, N \text{ et } t = 1, \dots, T$$

$$\widehat{\beta}^+ = (X^{+\prime} \Lambda^{-1} X^+)^{-1} X^{+\prime} \Lambda^{-1} Y \quad (12.23)$$

La matrice de variance-covariance est donnée par

$$\text{Var}(\widehat{\beta}^+) = \sigma_w^2 (X^{+\prime} \Lambda^{-1} X^+)^{-1} \quad (12.24)$$

L'estimateur des MCG est non biaisé et efficace. Si les aléas sont distribués selon une loi normale l'estimateur suit également une loi normale. De plus, il est convergent lorsque les variables explicatives sont strictement exogènes et lorsque $\lim_{N \rightarrow \infty} \frac{X' \Lambda^{-1} X}{N} = \Lambda^N$, Λ^N étant une matrice définie positive.

Cependant, pour effectuer cette estimation il faut connaître le paramètre θ et donc les variances des aléas u_i et ε_{it} . Celles-ci étant inconnues on doit appliquer l'estimateur des Moindres Carrés Quasi Généralisés (MCQG) dont les propriétés ne sont bonnes qu'asymptotiquement lorsqu'on dispose d'un estimateur convergent de θ et que $p \lim_{N \rightarrow \infty} \frac{X' \widehat{\Lambda}^{-1} X}{N} = \widehat{\Lambda}^N$, $\widehat{\Lambda}^N$ étant une matrice définie positive.

L'estimateur des MCQG est donné par

$$\widehat{\beta}^+ = (X^{+\prime} \widehat{\Lambda}^{-1} X^+)^{-1} X^{+\prime} \widehat{\Lambda}^{-1} Y \quad (12.25)$$

Avec

$$\widehat{\theta} = 1 - \frac{\widehat{\sigma}_\varepsilon}{\sqrt{T \widehat{\sigma}_u^2 + \widehat{\sigma}_\varepsilon^2}}$$

Pour estimer le paramètre θ on utilise les résultats d'estimation des modèles Within et Between.

Dans le modèle Within les données sont exprimées en écarts aux moyennes individuelles

$$y_{it} - \bar{y}_i = \beta_1(x_{1it} - \bar{x}_{1i}) + \beta_2(x_{2it} - \bar{x}_{2i}) + \dots + \beta_K(x_{Kit} - \bar{x}_{Ki}) + (\varepsilon_{it} - \bar{\varepsilon}_i)$$

Il y a NT observations et K paramètres à estimer

⁷ Lorsque que le paramètre θ est égal à 1 on obtient l'estimateur within. On remarque par ailleurs que lorsque T tend vers l'infini θ tend vers 1, ainsi, quand N et T tendent vers l'infini l'estimateur within est convergent et asymptotiquement efficace.

L'estimateur MCO des paramètres du modèle est non biaisé. On note w_{it} , les aléas du modèle Within $w_{it} = \varepsilon_{it} - \bar{\varepsilon}_i$, e_{wit} les résidus et $\hat{\sigma}_\varepsilon^2$ la variance estimée des aléas ε_{it}

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_i \sum_t e_{wit}^2}{NT - N - K}$$

$\hat{\sigma}_\varepsilon^2$ est un estimateur non biaisé de σ_ε^2 .

Dans le modèle Between on utilise les moyennes individuelles, le modèle s'écrit

$$\bar{y}_i = \alpha + \beta_1 \bar{x}_{1i} + \beta_2 \bar{x}_{2i} + \dots + \beta_K \bar{x}_{Ki} + \bar{\varepsilon}_i + u_i$$

Il y a N observations et $K + 1$ paramètres.

On note e_{bi} les résidus MCO et $\hat{\sigma}_b^2$ la variance estimée des aléas du modèle Between

$$\hat{\sigma}_b^2 = \frac{\sum_i e_{bi}^2}{N - (K + 1)}$$

$\hat{\sigma}_b^2$ est un estimateur non biaisé de $\sigma_u^2 + \frac{\sigma_\varepsilon^2}{T}$

On peut ainsi en déduire un estimateur sans biais et convergent de θ

$$\hat{\theta} = 1 - \frac{\hat{\sigma}_\varepsilon}{\sqrt{T \hat{\sigma}_b^2}}$$

L'estimateur des MCQG peut s'écrire comme une moyenne pondérée des estimateurs Within et Between.

$$\boxed{\hat{\beta}_{MCG} = F^w \hat{\beta}_w + (I - F^w) \hat{\beta}_b} \tag{12.26}$$

Avec $F^w = (X^w' X^w + (1 - \theta)^2 T X^b' X^b)^{-1} (X^w' X^w)$

3.2 Test des effets individuels aléatoires : le test de Breush Pagan

Pour tester la présence d'effets individuels aléatoires on teste l'hypothèse de nullité de la variance des aléas u_i

$$H0 : \sigma_u^2 = 0$$

$$H1 : \sigma_u^2 \neq 0$$

On calcule la statistique LM

$$\boxed{LM = \frac{NT}{2(T - 1)} \left(\frac{\sum_i \left(\sum_t \hat{\mu}_{it} \right)^2}{\sum_i \sum_t \hat{\mu}_{it}^2} - 1 \right)^2} \tag{12.27}$$

Avec $\hat{\mu}_{it}$ les résidus MCO du modèle

Sous l'hypothèse nulle la statistique LM est distribuée selon une loi du χ^2 à 1 degré de liberté. Si LM est inférieure à 3,84 (valeur correspondant à un seuil de 5 %) on conclut à l'absence d'effets individuels aléatoires.

Le test de Breush-Pagan est un test bilatéral or une variance ne peut-être que positive. Honda a proposé de tester

$$H0 : \sigma_u^2 = 0$$

$$H1 : \sigma_u^2 > 0$$

Et de calculer la statistique g

$$g = \sqrt{LM} = \sqrt{\frac{NT}{2(T-1)} \left(\frac{\sum_i \left(\sum_t \hat{\mu}_{it} \right)^2}{\sum_i \sum_t \hat{\mu}_{it}^2} - 1 \right)^2} \quad (12.28)$$

On compare g à 1,64 valeur de la table de la loi normale pour un seuil de 5 %.

3.3 Modèles à effets individuels aléatoires corrélés avec les régresseurs

On considère ici que dans le modèle à erreurs composées les effets individuels spécifiques aléatoires non observables sont corrélés aux variables explicatives autrement dit que $E(u_i | x_{1it}, \dots, x_{Kit}) \neq 0$, on suppose que les autres hypothèses sur les aléas sont vérifiées.

Les variables explicatives n'étant plus strictement exogènes, les estimateurs MCO, MCG et MCQG sont biaisés et non convergents.

1. Estimation

L'approche la plus simple consiste à éliminer les effets individuels aléatoires u_i en estimant avec les MCO le modèle Within ou avec les MCG le modèle en différence première.

En effet, on obtient un estimateur sans biais et convergent des paramètres β en estimant avec la méthode des MCO le modèle

$$y_{it} - \bar{y}_i = \beta_1 (x_{1it} - \bar{x}_{1i}) + \beta_2 (x_{2it} - \bar{x}_{2i}) + \dots + \beta_K (x_{Kit} - \bar{x}_{Ki}) + (\varepsilon_{it} - \bar{\varepsilon}_i)$$

les aléas du modèle sont homoscédastiques et non autocorrélés selon nos hypothèses sur les ε_{it} .

De même, on obtient un estimateur sans biais et convergent en appliquant les MCG au modèle en différence dont les aléas sont autocorrélés d'ordre 1 (cf ci-dessus).

$$y_{it} - y_{it-1} = \beta_1 (x_{1it} - x_{1it-1}) + \beta_2 (x_{2it} - x_{2it-1}) + \dots + \beta_K (x_{Kit} - x_{Kit-1}) + (\varepsilon_{it} - \varepsilon_{it-1})$$

Cette solution est possible lorsqu'il n'y a pas de variable explicative x_{kit} stable dans le temps c'est-à-dire telle que $x_{kit} = x_{ki} \forall t$. Dans le modèle Within et dans le modèle en différence ce type de variable disparaît. Dans ce cas il est préférable d'appliquer la méthode des variables instrumentales.

2. Test de corrélation entre les effets individuels et les variables explicatives

Le test d'Hausman peut être utilisé pour tester la corrélation entre les aléas u_i et les variables explicatives. Sous l'hypothèse nulle $E(u_i | x_{1it}, \dots, x_{Kit}) = 0$ et sous l'hypothèse alternative $E(u_i | x_{1it}, \dots, x_{Kit}) \neq 0$. Sous l'hypothèse nulle les estimateurs Within $\hat{\beta}_w$ et MCQG $\hat{\beta}^+$ sont convergents mais l'estimateur des MCQG est asymptotiquement le meilleur. Sous l'hypothèse alternative l'estimateur des MCQG est non convergent contrairement à l'estimateur Within.

Sous l'hypothèse nulle la statistique s'écrit

$$W = \left(\hat{\beta}_w - \hat{\beta}^+ \right)' \left(\widehat{Var}(\hat{\beta}_w) - \widehat{Var}(\hat{\beta}^+) \right)^{-1} \left(\hat{\beta}_w - \hat{\beta}^+ \right) \quad (12.29)$$

La statistique W est distribuée selon une loi du χ^2 à q degrés de liberté avec q le nombre de variables explicatives potentiellement corrélées à l'aléa u_i . Si la valeur calculée de la statistique est plus faible que la valeur tabulée alors on accepte l'hypothèse nulle de non corrélation. On estime dans ce cas le modèle avec l'estimateur des MCQG.

Le test peut-être pratiqué de la même façon en remplaçant $\hat{\beta}_w$ par l'estimateur MCO du modèle en différence première ou l'estimateur des variables instrumentales.

3.4 Modèles à effets individuels aléatoires dynamiques

Le modèle dynamique à effets individuels aléatoires s'écrit

$$y_{it} = \alpha + \beta' x_{it} + \lambda y_{it-1} + \mu_{it} \quad i = 1, \dots, N ; \quad t = 1, \dots, T \quad (12.30)$$

Avec : $\mu_{it} = u_i + \varepsilon_{it}$

On suppose que les variables explicatives x_{kit} sont strictement exogènes et donc non corrélées aux aléas.

L'aléa u_i intervenant pour toutes les observations $t = 1, \dots, T$, la variable endogène décalée y_{it-1} est corrélée à u_i et de fait à μ_{it} . Ainsi, l'estimateur des MCQG du modèle à effets individuels aléatoires n'est pas convergent. L'origine de la corrélation provient de la corrélation entre la valeur initiale de y_{it} et u_i . En effet le modèle dynamique à effets individuels aléatoires peut se réécrire de la manière suivante

$$\begin{aligned} y_{it} &= \alpha + \beta' x_{it} + \lambda (\alpha + \beta' x_{it-1} + \lambda y_{it-2} + u_i + \varepsilon_{it-1}) + u_i + \varepsilon_{it} && \Leftrightarrow \\ y_{it} &= \alpha(1 + \lambda) + \beta'(x_{it} + \lambda x_{it-1}) + \lambda^2 (\alpha + \beta' x_{it-2} + \lambda y_{it-3} + u_i + \varepsilon_{it-2}) \\ &\quad + u_i(1 + \lambda) + \varepsilon_{it} + \lambda \varepsilon_{it-1} && \Leftrightarrow \\ &\cdot \\ &\cdot \\ y_{it} &= \alpha(1 + \lambda + \dots + \lambda^{t-1}) + \beta'(x_{it} + \lambda x_{it-1} + \dots + \lambda^{t-1} x_{i1}) + \lambda^t y_{i0} \\ &\quad + u_i(1 + \lambda + \dots + \lambda^{t-1}) + \varepsilon_{it} + \lambda \varepsilon_{it-1} + \dots + \lambda^{t-1} \varepsilon_{i1} \end{aligned}$$

soit

$$y_{it} = \alpha \frac{1 - \lambda^t}{1 - \lambda} + \sum_{j=0}^{t-1} \sum_{k=1}^K \frac{1 - \lambda^t}{1 - \lambda} \beta_k x_{kit-j} + \lambda^t y_{i0} + \frac{1 - \lambda^t}{1 - \lambda} u_i + \sum_{j=0}^{t-1} \frac{1 - \lambda^t}{1 - \lambda} \varepsilon_{t-j}$$

Il faut donc appliquer des méthodes basées sur l'estimateur des variables instrumentales. Il existe diverses méthodes et les méthodes présentées dans le cadre du modèle dynamique à effets individuels fixes peuvent être appliquées ici.

On peut également signaler la méthode de Blundell et Bond à laquelle on fait souvent référence. Blundell et Bond ont souligné le manque de robustesse des résultats des estimations obtenues avec la méthode GMM de Arellano et Bond en raison de la faible corrélation entre les instruments choisis et les variables explicatives du modèle. Afin d'améliorer la robustesse des résultats ils proposent d'empiler le modèle en niveau et en différence première et de l'estimer avec GMM. Pour l'équation en différence les instruments de Arellano et Bond sont retenus et pour l'équation en niveau on utilise des instruments en différence première. Le modèle estimé est le suivant pour $t = 3, \dots, T$.

$$\boxed{\begin{cases} \Delta Y = \Delta X + \beta^+ + \lambda \Delta Y_{-1} + \Delta \varepsilon \\ Y = X + \beta^+ + \lambda Y_{-1} + u + \varepsilon \end{cases}} \quad (12.31)$$

et la matrice des instruments est donnée par

$$Z = \begin{pmatrix} Z^D & 0 \\ 0 & Z^N \end{pmatrix}$$

avec

$$Z^D = [\Delta X \quad W] \text{ et } Z^N = [\Delta X \quad \widetilde{W}]$$

$$\widetilde{W} = \begin{pmatrix} \widetilde{W}_1 \\ \widetilde{W}_2 \\ \vdots \\ \widetilde{W}_N \end{pmatrix} \text{ et } \widetilde{W}_i = \begin{pmatrix} \Delta y_{i2} & & & \\ & \Delta y_{i3} & & \\ & & \ddots & \\ & & & \Delta y_{iT-1} \end{pmatrix}$$

Dans le modèle en niveau les variables en différence sont des instruments valides s'ils sont non corrélés aux aléas μ . Cette condition est satisfaite si on suppose que

$$y_{i0} = \alpha + \beta' x_{i0} + \frac{u_i}{1 - \lambda} + \varepsilon_{i0}$$

En effet dans ce cas

$$\Delta y_{i1} = y_{i1} - y_{i0} = \alpha + \beta' x_{i1} + (\lambda - 1) y_{i0} + u_i + \varepsilon_{i1}$$

Or

$$(\lambda - 1) y_{i0} = \alpha(\lambda - 1) + (\lambda - 1)\beta' x_{i0} - u_i + (\lambda - 1)\varepsilon_{i0}$$

D'où

$$\begin{aligned} \Delta y_{i1} &= \alpha\lambda + \beta' [x_{i1} + (\lambda - 1)x_{i0}] - u_i + u_i + \varepsilon_{i1} + (\lambda - 1)\varepsilon_{i0} \\ &= \alpha\lambda + \beta' [x'_{i1} + (\lambda - 1)x_{i0}] + \varepsilon_{i1} + (\lambda - 1)\varepsilon_{i0} \end{aligned}$$

Ainsi Δy_{i1} est non corrélé à u_i et donc à $\mu_{i2} = u_i + \varepsilon_{i2}$ puisque les aléas ε_{it} sont supposés non autocorrélés.

Plus généralement Δy_{t-1} ⁸ est non corrélé à μ_{it} et peut-être utilisé comme instrument pour estimer le modèle en niveau.

Tous les tests présentés dans le cadre d'un modèle dynamique à effets individuels fixes s'appliquent également au modèle dynamique à effets individuels aléatoires.

⁸
$$\Delta y_{it} = \sum_{j=0}^{t-2} \sum_{k=1}^K \frac{1 - \lambda^{t-2}}{1 - \lambda} \beta_k \Delta x_{kit-j} + \lambda^{t-1} \Delta y_{i1} + \sum_{j=0}^{t-2} \frac{1 - \lambda^{t-1}}{1 - \lambda} \varepsilon_{t-j}$$

13

ANALYSE DU MARCHÉ DU TRAVAIL NON QUALIFIÉ

1. Introduction
2. Les modèles de négociation
3. Le modèle du droit à gérer
4. Éléments de corrigés

OBJECTIFS DU CHAPITRE

D'un point de vue méthodologique

- Travailler avec des données de panel
- Distinguer des effets fixes et aléatoires individuels
- Appliquer les tests de Breush-Pagan et d'Hausman

D'un point de vue économique

- Effectuer une analyse du marché du travail
- Présenter les modèles de négociation entre les salariés et les entreprises
- Estimer le modèle du droit à gérer

NIVEAU

MASTER, première et deuxième années

CHAMPS ÉCONOMIQUES TRAITÉS

Macroéconomie réelle

Économie du travail

1. INTRODUCTION

L'objectif de ce chapitre est d'examiner le marché du travail des salariés non qualifiés en France à partir d'un panel constitué de 31 secteurs français d'activité pour la période 1988-1992, le marché étant alors en situation de déséquilibre, avec un taux de chômage de la main-d'œuvre non qualifiée d'environ 12 % au cours de cette période.

Dans le modèle néoclassique, le marché du travail est représenté par des équations d'offre et de demande de travail fonctions du salaire réel. L'offre de travail est une fonction croissante du salaire réel. Cette équation est obtenue par le biais d'un programme de maximisation de l'utilité des individus qui doivent effectuer un arbitrage entre travail et loisir. L'équation de demande est une fonction décroissante du salaire réel. Les entreprises déterminent leur demande de travail de manière à minimiser leur coût de production et ainsi à maximiser leur profit. De la confrontation de l'offre et de la demande sur le marché du travail résulte l'emploi et le salaire d'équilibre. Si l'offre est plus élevée que la demande le salaire réel diminue, et inversement si la demande est plus élevée que l'offre le salaire réel augmente. Dans ce schéma, le chômage est lié à un manque de flexibilité des salaires, c'est-à-dire à des rigidités sur le marché du travail. Si le salaire est fixé à un niveau où l'offre est supérieure à la demande, le marché est en situation de déséquilibre et le chômage apparaît.

Depuis le début des années 1980, en lien avec le développement de la théorie des jeux, de nouveaux outils de modélisation ont été mis en œuvre afin d'expliquer le fort degré d'inertie du taux de chômage observé en France et dans la plupart des pays européens. Les modèles de négociation salariale, fondés sur la théorie des jeux, intègrent les interactions stratégiques entre les acteurs qui participent à la formation des niveaux de salaires et de l'emploi : les syndicats et les firmes. Les négociations entre syndicats et firmes peuvent porter à la fois sur les niveaux de salaire et d'emploi ou seulement sur le niveau des salaires. Les comportements individuels sont alors analysés dans un cadre de concurrence imparfaite.

Ce cadre d'analyse est retenu ici car, en France comme dans l'ensemble des pays industrialisés, les conditions de travail sont définies par négociations. Celles-ci mettent en relation des représentants des dirigeants de l'entreprise avec des syndicats de salariés qui peuvent être organisés par métiers ou au niveau de l'entreprise, de la branche, de la région ou de la nation. Le niveau de négociation privilégié en France est traditionnellement celui de la branche où les accords prennent la forme de conventions collectives (ensemble de dispositions négociées entre un ou plusieurs employeurs et les représentants des salariés). Appliquées depuis la fin du XIX^e siècle, elles ont connu un nouvel élan en 1982 avec les lois Auroux qui rendent obligatoires une négociation annuelle des salaires et une négociation quinquennale des classifications. Sur les 1270 branches que compte l'activité française, 733 sont régies par une convention collective, ce qui représente plus de 80 % des effectifs salariés.

Dans une première section, les modèles de négociation sont présentés. Dans une deuxième section, le modèle du droit à gérer est examiné. Dans la troisième et la quatrième section on effectue une analyse empirique du modèle.

2. LES MODÈLES DE NÉGOCIATION

Il existe de nombreux modèles de négociation entre les entreprises et les salariés. On s'intéresse ici à celui de Manning (1987)¹, qui présente l'intérêt d'intégrer plusieurs modèles théoriques de négociation selon que la négociation porte sur les salaires et/ou l'emploi.

De manière générale, les modèles de négociation entre syndicats et firmes s'appuient sur des fondements microéconomiques, plus spécifiquement sur la théorie des jeux. La plupart traitent le problème des négociations comme un problème de marchandage de Nash dans lequel interviennent deux acteurs : le syndicat, supposé être composé de membres identiques, et l'entreprise. Les salaires (w) et l'emploi (l) sont les solutions du problème

$$\text{Max } (U - U^*)^\theta (\pi - \pi^*)^{(1-\theta)} \quad (13.1)$$

ou de manière équivalente

$$\text{Max } [\theta \text{Log}(U - U^*) + (1 - \theta) \text{Log}(\pi - \pi^*)] \quad (13.2)$$

où U est la fonction d'utilité du syndicat et π la fonction de profit de l'entreprise. Les niveaux U^* et π^* représentent respectivement l'utilité du syndicat et le profit de l'entreprise pendant la situation de statu quo, correspondant à la période pendant laquelle il y a négociation et donc pas encore d'accord. Enfin, θ indique le pouvoir de négociation du syndicat. Sous certaines conditions, la solution de Nash du programme précédent est équivalente au résultat d'un jeu non coopératif entre la firme et le syndicat.

La complexité de ce programme réside d'abord dans la formalisation de la fonction d'utilité du syndicat. Le point de vue le plus fréquemment adopté est de considérer le syndicat des salariés comme un consommateur ou une entreprise, son objectif étant alors de maximiser une fonction d'utilité dont les arguments sont les salaires et l'emploi. On a donc $U \equiv U(w, l)$.

Par ailleurs, le programme de négociation doit stipuler l'objet des négociations, les salaires et/ou l'emploi. Manning (1987) propose un modèle de marchandage séquentiel. Il y a d'abord un marchandage sur les salaires (ou l'emploi), puis sur l'emploi (ou les salaires), le pouvoir de négociation du syndicat étant différent à chaque étape. On considère ici le cas où la firme et le syndicat négocient dans une première étape le salaire et dans une seconde étape l'emploi. La solution du problème est alors déterminée en considérant d'abord la seconde étape où le niveau d'emploi est choisi de manière à résoudre le problème

$$\text{Max}_l [\theta_2 \text{Log}(U(w, l) - U^*) + (1 - \theta_2) \text{Log}(\pi(w, l) - \pi^*)] \quad (13.3)$$

¹ Manning A., 1987, « An integration of Trade Union models in a sequential bargaining framework », *The Economic Journal*, 97, pp. 121-139.

où θ_2 représente le pouvoir de négociation du syndicat en matière d'emploi. La solution, supposée unique, du problème définit le niveau d'emploi l comme une fonction des salaires w et du pouvoir de négociation $\theta_2 : l = l(w, \theta_2)$. La première étape du jeu s'écrit alors

$$\underset{w}{Max} [\theta_1 \text{Log} (U(w, l(w, \theta_2)) - U^*) + (1 - \theta_1) \text{Log} (\pi(w, l(w, \theta_2)) - \pi^*)] \quad (13.4)$$

ou θ_1 représente le pouvoir de négociation du syndicat en termes de salaires. Cette présentation en deux étapes a l'inconvénient de ne pas avoir de base théorique précise. Cependant, il fournit un cadre général de programme de négociation. Il intègre, en particulier, le modèle du droit à gérer.

Dans ce modèle, le niveau de l'emploi est déterminé unilatéralement par l'entreprise étant donné le salaire négocié avec le syndicat. L'entreprise, maximisant ses profits, établit sa demande d'emploi de manière à égaliser la productivité marginale du travail au salaire.

Dans ce modèle, $0 < \theta_1 < 1$ et $\theta_2 = 0$ et le programme de négociation s'écrit

$$\underset{w}{Max} [U(w, l)]^{\theta_1} [\pi(w, l)]^{(1-\theta_1)} \quad (13.5)$$

le salaire négocié est une fonction croissante du pouvoir de négociation du syndicat. En revanche, l'emploi diminue avec le pouvoir de négociation car la productivité marginale du travail, égale au salaire, est décroissante. Au final, l'entreprise est conduite à fixer un niveau d'emploi plus faible que celui qui prévaudrait sans syndicat, le salaire négocié étant plus élevé. Le contrat est sous optimal au sens de Pareto. Les couples (w, l) sont sur la courbe de demande de travail.

Ce cadre d'analyse est retenu car il semble correspondre au cas français. En effet, en France, à l'occasion des négociations collectives de multiples thèmes sont abordés : les salaires, les grilles de classification, les primes, les durées du travail, les effectifs etc. Cependant, près de la moitié des accords passés portent sur les rémunérations tandis que seulement 5 % sont relatifs au niveau de l'emploi. Ce constat plaide en faveur du modèle théorique du droit à gérer, selon lequel les salaires font l'objet d'un « marchandage » entre salariés et entreprises, ces dernières gardant le monopole de leur politique d'emploi.

Pour un complément sur les modèles de négociations le lecteur peut se reporter à Cahuc, Zylberberg (1998)².

² Cahuc P., Zylberberg A., 1998, *Économie du travail, la formation des salaires et les déterminants du chômage*, De Boeck Université.

3. LE MODÈLE DU DROIT À GÉRER

3.1 Spécification de l'équation de salaire

Le modèle du droit à gérer de Nickell (1990)³ est repris de manière à spécifier une équation de salaire. Le programme de négociation faisant intervenir l'entreprise et le syndicat, on examine dans un premier temps leur comportement respectif. Dans un second temps, on détermine l'équation de salaire.

3.1.1 Le comportement de l'entreprise

On suppose que les firmes font face à un niveau incertain de demande de biens et que les salaires sont négociés avant que cette incertitude soit levée. Lorsque les salaires sont fixés, l'entreprise choisit le niveau d'emploi, noté l , qui maximise son profit π

$$\underset{l}{\text{Max}} \pi = P_e A F(l) \varepsilon - w l \quad (13.6)$$

où P_e est le prix anticipé du bien, A l'indice de productivité globale des facteurs, F la fonction de production, ε une variable aléatoire qui représente l'incertitude sur la demande de bien et w le salaire nominal des employés. La condition du premier ordre donne

$$\frac{\partial \Pi}{\partial l} = 0 \iff P_e A \varepsilon F_l = w \iff \frac{\partial F}{\partial l} = \frac{w}{P_e A \varepsilon} \quad (13.7)$$

On en déduit la fonction de demande de travail et la fonction de profit

$$l^* = l \left(\frac{w}{A P_e \varepsilon} \right) \text{ et } \pi^* = \pi \left(\frac{w}{A P_e \varepsilon} \right) \quad (13.8)$$

Dans le modèle du droit à gérer, la demande de travail est déterminée de manière unilatérale par l'entreprise. Les niveaux de l'emploi et des profits de la firme sont fonction notamment du salaire réel anticipé, le salaire nominal étant fixé lors des négociations entre les syndicats et l'entrepreneur.

3.1.2 Le comportement du syndicat

Lors des négociations salariales, le syndicat cherche à maximiser sa fonction d'utilité. La difficulté dans ce modèle réside dans la spécification de la fonction d'utilité $U(w, l)$, on retient généralement la fonction d'utilité espérée d'un membre représentatif du syndicat. Cette dernière est définie en référence à deux situations possibles, la première correspond au cas où l'individu est réemployé suite aux négociations et la seconde se réfère au cas où l'individu est licencié. L'utilité espérée de l'individu est alors fonction de sa probabilité d'être licencié.

³ Nickell S., 1990, « Unemployment survey », *The economic Journal*, 100, pp. 391-439.

La probabilité d'être licencié pour les employés en place (les insiders) est nulle lorsque la demande de travail de la firme est supérieure au nombre d'insiders (m). Dans le cas contraire elle est notée pl . Par conséquent, la probabilité qu'un salarié de l'entreprise soit licencié s'écrit

$$\begin{aligned} pr &= \text{Pr } ob(m \leq l).0 + \text{Pr } ob(m > l).pl \\ &= \text{Pr } ob(m > l).pl \end{aligned} \quad (13.9)$$

Cette probabilité s'accroît avec le nombre d'insiders car la probabilité que le nombre d'insiders est plus élevée que la demande de travail augmente. De même, pr est une fonction décroissante de l . Par conséquent, la demande d'emploi étant une fonction décroissante du salaire réel anticipé, la probabilité d'être licencié est une fonction croissante de ce salaire réel

$$pr = pr \left(m, \frac{w}{AP^e} \right) \quad (13.10)$$

avec

$$\frac{\partial pr}{\partial w} > 0 \text{ et } \frac{\partial pr}{\partial m} > 0$$

Ayant défini la probabilité d'un individu d'être licencié on peut déterminer l'utilité espérée du membre représentatif du syndicat. Cette utilité est la somme de l'utilité qu'il obtient avec une probabilité $(1 - pr)$, en cas de réembauche et de l'utilité qu'il obtient avec une probabilité pr en cas de licenciement. On note $u(w)$ l'utilité en cas de réembauche, et $u(w_a)$ l'utilité en cas de licenciement où w_a représente un salaire alternatif versé aux membres du syndicat non réemployés dans l'entreprise à la suite des négociations. L'utilité espérée s'écrit

$$U = (1 - pr).u(w) + (pr).u(w_a) \quad (13.11)$$

Sous l'hypothèse que $u(w) = w$, le terme $(1 - pr).w$ représente l'espérance de gains du salarié s'il est réemployé. Pour une personne licenciée de l'entreprise, deux cas de figures existent : soit elle est réemployée à l'extérieur et perçoit un salaire we , soit elle est au chômage et perçoit un revenu B . Par conséquent si elle reste inoccupée pendant une proportion de temps t , croissante avec le taux de chômage u , ses revenus s'élèvent au montant

$$we.(1 - t(u)) + t(u).B \quad (13.12)$$

En notant b le ratio $\frac{B}{we}$, l'expression précédente s'écrit

$$we.(1 - t(u).(1 - b))$$

Si on suppose que $u(w_a) = we.(1 - t(u).(1 - b))$, alors le terme

$$pr.u(w_a) = pr.[we.(1 - t(u).(1 - b))]$$

indique l'espérance de gains d'un salarié licencié suite aux négociations entre le syndicat et l'entreprise.

L'utilité espérée d'un membre du syndicat est finalement donnée par

$$U = (1 - pr).w + pr. [we.(1 - t(u).(1 - b))] \quad (13.13)$$

Enfin, il reste du point de vue du syndicat à déterminer U^* le niveau d'utilité pendant la période de statu quo. On suppose, pour simplifier, qu'il est égal au montant des revenus perçus en cas de licenciement. Ainsi

$$U^* = we.(1 - t(u).(1 - b)) \quad (13.14)$$

3.1.3 Programme de négociation des salaires

Ayant défini le comportement des deux acteurs de la négociation salariale, on peut établir le programme de négociation afin d'établir l'équation de salaire.

Si on suppose que $\pi^* = 0$, le programme de Nash de négociation des salaires s'écrit

$$\underset{w}{Max} [U(w, l) - U^*]^{\theta_1} [\pi(w, l)]^{(1-\theta_1)} \quad (13.15)$$

En remplaçant $U(w, l)$ et U^* respectivement par les équations (13.13) et (13.14), le programme (13.15) devient

$$\underset{w}{Max} \left[\left(1 - pr. \left(m. \frac{w}{AP_e} \right) \right) (w - we.(1 - t(u).(1 - b))) \right]^{\theta_1} \left[\pi \left(\frac{w}{AP_e} \right) \right]^{1 - \theta_1} \quad (13.16)$$

La solution générale de ce problème implique que le salaire négocié est une fonction de we , AP_e , m , u , b et θ_1

$$w = w(we, AP_e, m, u, b, \theta_1) \quad (13.17)$$

avec $w_{we}, w_{AP_e}, w_b, w_{\theta_1} > 0$ et $w_m, w_u < 0$.

La fonction (13.17) est homogène de degré 1 par rapport à we et AP_e . Elle peut donc s'écrire sous la forme

$$w = w \left(\frac{we}{AP_e}, m, u, b, \theta_1 \right) \quad (13.18)$$

Les salaires, dans ce modèle, sont influencés par des facteurs liés au fonctionnement général du marché du travail (we, u, b) et par des facteurs spécifiques à la

firme et au syndicat (AP^c, m, θ_1). En particulier les salaires sont reliés négativement au nombre d'insiders et au taux de chômage, et positivement aux autres facteurs.

3.2 Évaluation du modèle du droit à gérer

On se propose de tester empiriquement la validité du modèle du droit à gérer avec des données sectorielles françaises sur les emplois non qualifiés. On retient ici la spécification log linéaire proposée par Nickell (1990)

$$\text{Log}w = \alpha + \beta_1 \text{Log}w_e + \beta_2 \text{Log}AP_e + \beta_3 \text{Log}m + \beta_4 u + \beta_5 b + \beta_6 \theta_1 + \varepsilon \quad (13.19)$$

avec ε un aléa.

Pour estimer ce modèle, on dispose d'un panel cylindré⁴ de 31 secteurs de l'activité française de la NAP 40 sur la période 1988-1992. Les séries utilisées sont issues de plusieurs sources statistiques diffusées par l'INSEE : les Déclarations Annuelles de Données Sociales (DADS), les Enquêtes sur la Structure des Emplois (ESE) et les comptes nationaux annuels.

3.2.1 Les données issues des DADS

Ces déclarations annuelles sont obligatoires pour tout établissement employant au moins un salarié. Les informations recueillies sont relatives aux caractéristiques des salariés (salaire, qualification, statut de l'emploi occupé, secteur d'activité, âge, ...) et à celles des établissements (code APE, effectifs suivant les CSP, masse salariale, ...). Des deux exploitations disponibles, on utilise celle relative aux salariés pour construire les séries de salaires et d'effectifs.

Dans l'élaboration de nos séries de salaires et d'effectifs, on se limite aux salariés permanents à temps complet des secteurs privé et semi-public dans la mesure où ces salariés disposent d'un pouvoir de négociation.

Les salaires sont les salaires nets annuels moyens. Ils intègrent donc, en plus du salaire de base les rémunérations pour heures supplémentaires ainsi que les primes qu'elles soient définitives, et intégrées au salaire de base, ou aléatoires, comme celles relatives aux performances de l'entreprise ou des individus. On tient compte ainsi de la partie du salaire influencée par la négociation.

Ces salaires sont ventilés, d'une part selon le secteur d'activité de la NAP 40 (pour des raisons de disponibilité de certaines données, nous avons écarté les secteurs de la pêche, des postes et télécommunications, des services non marchands, et nous avons agrégé les secteurs du commerce), d'autre part selon les PCS (Profession et Catégorie Socio-professionnelle).

Les estimations sont réalisées pour les emplois non qualifiés. La qualification d'un salarié peut être appréciée par son niveau de formation, ainsi que par le niveau de qualification de son emploi décrit par la nomenclature PCS. L'utilisation de ce

⁴ On dispose pour chaque secteur d'un nombre d'observations temporelles identique.

dernier critère permet de distinguer les professions intermédiaires et d'encadrement des professions d'employés et d'ouvriers. Cette partition reflète en outre les différences de formation puisque le pourcentage des salariés du premier groupe dotés au minimum d'un diplôme professionnel atteint 60 %, alors qu'il n'est que de 10 % pour le second (Mihoubi, 1997). Le niveau de formation conduit donc au même clivage que le niveau de qualification de l'emploi occupé. On retient ici une partition standard dans laquelle le groupe des non qualifiés se compose des employés et des ouvriers.

3.2.2 Les données issues de l'ESE

Les effectifs par PCS et par secteur ont été calculés en appliquant aux effectifs (exprimés en années travail) toutes PCS confondues fournies par les DADS la répartition entre les PCS issue de l'Enquête sur la Structure des Emplois.

3.2.3 Les données issues des comptes nationaux annuels

Deux séries sont utilisées

- le taux de chômage des ouvriers et des employés calculés selon la définition du BIT,
- le prix implicite à la production, mesuré par le rapport entre la valeur ajoutée de chaque secteur en prix courants et cette valeur ajoutée en prix constants.

Enfin, pour estimer l'équation de salaire dérivée du programme de négociation il faut au préalable évaluer le pouvoir de négociation du syndicat représenté par la variable θ_1 . Il n'existe pas de statistiques précises sur ce pouvoir de négociation pour chacun des secteurs entre 1987 et 1992. Sur données françaises, certaines études utilisent le taux de syndicalisation pour approximer le pouvoir de négociation du syndicat, cependant ces chiffres ne reflètent pas réellement le pouvoir de négociation, de plus, les données ne sont pas désagrégées selon la NAP40. D'autres études spécifient le pouvoir de négociation comme un paramètre structurel de l'équation de salaire, or dans la spécification proposée ici il s'agit d'une variable explicative et non d'un paramètre. Pour résoudre ce problème un modèle Probit a été construit pour évaluer ce pouvoir de négociation⁵.

Dans le fichier de données chapitre13.xls on dispose, pour chaque secteur et chaque année de 1988 à 1992, des séries suivantes

- *wouv* : salaire annuel net moyen des ouvriers exprimé en francs courants
- *wempl* : salaire annuel net moyen des employés exprimé en francs courants
- *low* : nombre d'ouvriers exprimé en années travail
- *lempl* : nombre d'employés exprimé en années travail
- *l* : effectifs salariés non qualifiés ($l = low + lempl$)

⁵ Les détails de cette évaluation peuvent être obtenus auprès des auteurs. De manière succincte : on a créé une variable indicatrice prenant la valeur 1 lorsque dans un secteur il y a beaucoup d'accords conclus et peu de conflits. Cette variable a été utilisée comme variable expliquée dans un modèle probit dont le but est d'évaluer la probabilité que dans un secteur il y a beaucoup d'accords et peu de conflits. La probabilité ainsi estimée est utilisée comme une variable proxy du pouvoir de négociation du syndicat.

- w : salaire annuel net moyen des salariés non qualifiés calculé comme une moyenne pondérée des salaires des ouvriers et des employés.

$$w = \frac{w_{empl}.l_{empl} + w_{ouv}.l_{ouv}}{l}$$

- w_{ouv} : taux de chômage des ouvriers
- w_{empl} : taux de chômage des employés
- u : moyenne du taux de chômage des ouvriers et des employés
- P : prix implicite à la production
- pns : pouvoir de négociation du syndicat évalué avec un modèle Probit
- m : nombre d'insider supposé identique aux effectifs salariés non qualifiés de l'année précédente ($m = l_{-1}$).
- w_e : salaire alternatif que les salariés non qualifiés peuvent espérer percevoir en cas d'échec des négociations. Il est mesuré par le salaire annuel net moyen des salariés non qualifiés pratiqué dans les autres secteurs.

Problème 13.1

Estimation de l'équation de salaire

On propose d'estimer l'équation de salaire⁶

$$\begin{aligned} \text{Log}w_{s,t} = & \alpha + \beta_1 \text{Log}w_{e,s,t} + \beta_2 \text{Log}P_{s,t}^e + \beta_3 \text{Log}m_{s,t} + \beta_4 u_t \\ & + \beta_5 pns_{s,t} + \varepsilon_{s,t} \end{aligned} \quad (13.20)$$

$$s = 1, \dots, 31, \quad t = 1988, \dots, 1992 \quad (13.21)$$

Les prix anticipés $P_{s,t}^e$ sont supposés identiques aux prix observés.

1. Estimer avec la méthode des MCO le modèle et commenter les résultats.
2. Expliquer pourquoi notamment les variables salaire extérieur et taux de chômage sont probablement endogènes.
3. Estimer avec la méthode des Variables Instrumentales le modèle et commenter vos résultats.

Problème 13.2

Évaluation de l'équation de salaire avec des effets individuels

Dans le problème 13.1, on fait l'hypothèse implicite que l'équation de salaire est identique dans tous les secteurs d'activité de l'économie française car les paramètres estimés sont indépendants du secteur considéré

$$(\alpha_s, \beta_{1s}, \beta_{2s}, \beta_{3s}, \beta_{4s}, \beta_{5s}) = (\alpha, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5) \quad \forall s = 1, \dots, 31$$

⁶ Notons que, par rapport à l'équation (13.19), la variable b est omise. Cette variable est un ratio qui est supposé demeuré constant au cours des 4 années et dans les différents secteurs d'activité. Autrement dit on suppose que $b_{st} = b \forall s, t$

Cette hypothèse ne peut pas être testée car on ne dispose pas d'un nombre d'observations temporelles suffisant. Cependant, il est possible d'introduire un effet spécifique à chaque secteur d'activité qui serait par exemple lié à sa technologie de production. Dans ce cas on suppose, d'une part, que α_s est différent selon le secteur considéré et, d'autre part, que

$$(\beta_{1s}, \beta_{2s}, \beta_{3s}, \beta_{4s}, \beta_{5s}) = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5) \quad \forall s = 1, \dots, 31.$$

le modèle (13.20) devient

$$\begin{aligned} \text{Log}w_{s,t} = & \alpha_s + \beta_1 \text{Log}we_{s,t} + \beta_2 \text{Log}P_{s,t}^e + \beta_3 \text{Log}m_{s,t} \\ & + \beta_4 u_t + \beta_5 pm_{s,t} + \varepsilon_{s,t} \end{aligned} \quad (13.22)$$

$$s = 1, \dots, 31, \quad t = 1988, \dots, 1992$$

1. Estimer le modèle à effets aléatoires individuels (avec la méthode des MCQG). Appliquer le test de Breusch-Pagan pour tester la présence des effets aléatoires et le test d'Hausman pour tester si les effets individuels sont corrélés aux aléas.
2. Estimer avec la méthode des MCO le modèle (13.22). Les effets individuels estimés sont-ils fixes ou aléatoires ?
Estimer avec la méthode Within le modèle et comparer les résultats.
3. Étant donné les problèmes d'endogénéité des variables, posés dans le problème 13.1, l'estimateur des MCO est biaisé, proposer alors un estimateur convergent des paramètres du modèle (13.22).

Problème 13.3

Comparaison des estimateurs MCO, MCG, Within et Between

Dans un échantillon de données de panel, les observations ont une dimension temporelle et individuelle. Les estimateurs MCO et MCG prennent en compte ces deux dimensions. Ils peuvent s'écrire comme une moyenne pondérée des estimateurs Within et Between.

Afin de simplifier les calculs on va s'écarter du modèle théorique et considérer le modèle

$$\begin{aligned} \text{Log}w_{s,t} = & \beta_1 + \beta_2 \text{Log}we_{s,t} + \beta_3 \text{Log}P_{s,t}^e + \beta_4 \text{Log}m_{s,t} + \varepsilon_{s,t} \\ & s = 1, \dots, 31, \quad t = 1988, \dots, 1992 \end{aligned}$$

1. Estimer le modèle avec la méthode des MCO (modèle sans effets fixes individuels).
2. Estimer le modèle avec la méthode des MCQG (modèle avec effets aléatoires individuels).

3. Estimer le modèle avec les estimateurs Within et Between.
4. Vérifier que les estimateurs des MCO et MCQG peuvent s'écrire comme une moyenne pondérée des estimateurs Within et Between.

4. ÉLÉMENTS DE CORRIGÉS

Problème 13.1

Estimation de l'équation de salaire

1. Lorsqu'on estime avec la méthode des MCO le modèle, on suppose que l'équation de salaire est définie de manière identique dans les 31 secteurs d'activités car les coefficients prennent la même valeur quel que soit le secteur considéré. Les résultats de cette estimation sont donnés dans le tableau 13.1.

Tableau 13.1 Estimation de l'équation de salaire par les MCO

Linear Regression — Estimation by Least Squares			
Dependent Variable LOGW			
Panel(5) of Annual Data From 1//1988 :01 To 31//1992 :01			
Usable Observations	155	Degrees of Freedom	149
Centered R**2	0.349071	R Bar **2	0.327228
Uncentered R**2	0.999866	T x R**2	154.979
Mean of Dependent Variable	11.363483654		
Std Error of Dependent Variable	0.163515148		
Standard Error of Estimate	0.134119525		
Sum of Squared Residuals	2.6802190175		
Regression F(5,149)	15.9807		
Significance Level of F	0.00000000		
Durbin-Watson Statistic	0.360814		

	Variable	Coeff	Std Error	T-Stat	Signif
1.	LOGWE	0.879808188	0.218803342	4.02100	0.00009177
2.	LOGP	-0.126734601	0.069221338	-1.83086	0.06911915
3.	LOGM	-0.104883119	0.014366353	-7.30061	0.00000000
4.	U	-0.382217779	2.032303759	-0.18807	0.85107674
5.	PNS	-0.253322744	0.076621245	-3.30617	0.00118521
6.	Constant	2.093160601	2.584792994	0.80980	0.41934756

La qualité de l'ajustement, évaluée par le coefficient de détermination (R^2), indique que seulement 35 % des variations de la variable endogène sont expliquées par le modèle. Par ailleurs les coefficients associés à la constante et la variable de taux de chômage sont non significativement différents de 0. Conformément au modèle théorique le salaire extérieur agit positivement sur

le salaire négocié, un accroissement de 1 % du salaire extérieur induit un accroissement de 0.88 % du salaire négocié. De même le coefficient de la variable de prix a le signe attendu, sa valeur est cependant faible puisque l'élasticité est de -0.12 % ce qui est loin d'une indexation des salaires sur les prix. Enfin le nombre d'insider a un impact négatif sur les salaires, une augmentation de 1 % du nombre d'insider induit une réduction des salaires de 0.1 %. En revanche, le signe du coefficient de la variable *pns* n'est pas conforme à celui attendu puisqu'il est négatif.

2. Dans le modèle, le salaire négocié dans un secteur d'activité est fonction du salaire extérieur et du taux de chômage, ces deux variables sont supposées exogènes et non corrélées aux aléas. Or, le salaire extérieur d'un secteur d'activité est donné par le salaire moyen négocié dans les autres secteurs, les salaires des différents secteurs étant négociés de manière simultanée *we* apparaît endogène. De même, la variable taux de chômage est endogène car elle est déterminée lorsque la demande d'emploi, fonction du salaire négocié, est fixée. Ces deux variables théoriquement sont endogènes et par conséquent corrélées aux aléas. L'estimateur des MCO calculé précédemment est donc biaisé et le biais ne disparaît pas asymptotiquement. Pour résoudre ce problème d'endogénéité on peut utiliser l'estimateur des Variables Instrumentales.
3. Afin d'appliquer la méthode des Variables Instrumentales il faut déterminer deux variables non corrélées aux aléas, une doit être fortement corrélée à *Logwe* et l'autre à *Logu*. Dans cet exercice on propose d'utiliser respectivement les variables *Logwe* et *Logu* décalées d'une période⁷. Les résultats d'estimation sont donnés dans le tableau 13.2.

Tableau 13.2 Estimation de l'équation de salaire par les Variables Instrumentales

Linear Regression — Estimation by Instrumental Variables			
Dependent Variable LOGW			
Panel(5) of Annual Data From 1//1989 :01 To 31//1992 :01			
Usable Observations	124	Degrees of Freedom	118
Total Observations	154	Skipped/Missing	30
Centered R**2	0.257510	R Bar **2	0.226048
Uncentered R**2	0.999852	T x R**2	123.982
Mean of Dependent Variable	11.383384158		
Std Error of Dependent Variable	0.161306280		
Standard Error of Estimate	0.141908429		
Sum of Squared Residuals	2.3762842518		
Durbin-Watson Statistic	0.651596		

⁷ On note que ces instruments ne sont pas valides lorsque les aléas du modèle sont autocorrélés d'ordre 1.

	Variable	Coeff	Std Error	T-Stat	Signif
1.	LOGWE	1.195196239	0.314844880	3.79614	0.00023360
2.	LOGP	-0.166669275	0.078947991	-2.11113	0.03687204
3.	LOGM	-0.112512305	0.017503172	-6.42811	0.00000000
4.	U	-8.362647237	2.920510362	-2.86342	0.00496119
5.	PNS	-0.294701869	0.093393637	-3.15548	0.00203350
6.	Constant	-0.430585891	3.521548539	-0.12227	0.90289180

Tous les coefficients, sauf celui de la constante sont significativement différents de 0, le signe des coefficients correspond à celui attendu théoriquement sauf celui associé à la variable *pns*.

Problème 13.2

Évaluation de l'équation de salaire avec des effets individuels

Dans ce problème on ignore dans un premier temps le problème d'endogénéité soulevé précédemment, on le réintroduira dans la question 5 de ce problème.

1. Dans le modèle à effets aléatoires individuels le paramètre α_s est égal à $\alpha + \nu_s$ avec α une constante, ν_s un bruit blanc de moyenne nulle et de variance σ_ν^2 et non corrélé à l'aléa $\varepsilon_{s,t}$. Le tableau 13.3 donne les résultats d'estimation, avec la méthode des MCQG, du modèle à effets aléatoires individuels. Les variables de prix et de taux de chômage ne sont pas significativement différentes de zéro, en revanche le signe des coefficients est conforme à celui attendu en particulier le coefficient de la variable *pns* est positif et significativement différent de 0.

Tableau 13.3 Estimation de l'équation de salaire avec des effets aléatoires individuels

Panel Regression — Estimation by Random Effects					
Dependent Variable LOGW					
Panel(5) of Annual Data From 1//1988 :01 To 31//1992 :01					
Usable Observations	155	Degrees of Freedom	149		
Centered R**2	0.995706	R Bar **2	0.995562		
Uncentered R**2	0.999999	T x R**2	155.000		
Mean of Dependent Variable	11.363483654				
Std Error of Dependent Variable	0.163515148				
Standard Error of Estimate	0.010893716				
Sum of Squared Residuals	0.0176822846				

	Variable	Coeff	Std Error	T-Stat	Signif
1.	LOGWE	0.963403453	0.051460876	18.72109	0.00000000
2.	LOGP	-0.023112915	0.035776805	-0.64603	0.51825946
3.	LOGM	-0.050986751	0.017925746	-2.84433	0.00445048
4.	U	-0.001216618	0.412620074	-0.00295	0.99764742
5.	PNS	0.095381759	0.043257500	2.20498	0.02745576
6.	Constant	0.624422895	0.613043094	1.01856	0.30841057

Le test de Breush-Pagan teste l'hypothèse $H_0 : \sigma_v^2 = 0$. Sous H_0 il n'y a pas d'effets aléatoires individuels et on retient l'estimateur des MCO. Cette hypothèse n'est pas acceptée car la statistique LM, égale à 278, est plus élevée que la valeur d'un $\chi^2(1)$. Par conséquent, on retient l'estimation avec la méthode des MCQG qui suppose l'existence d'effets aléatoires individuels.

Lorsqu'on applique le test d'Hausman on obtient une valeur pour la statistique de Wald égale à 11.58. Cette valeur est plus élevée que la valeur tabulée du $\chi^2(5)$ et le niveau de significativité du test est de 4 %. On refuse donc l'hypothèse H_0 selon laquelle les effets individuels ne sont pas corrélés aux variables explicatives. Par conséquent, l'estimateur des MCQG est non convergent et on préconise l'utilisation de l'estimateur within (résultats du tableau 13.4).

- Le modèle à effets fixes individuels, représenté par α_s dans (13.22), est estimé avec les MCO. Les résultats d'estimation restent proches de ceux obtenus avec l'estimateur des MCQG (tableau 13.3). L'estimateur des MCO des paramètres du modèle (13.22) est identique à l'estimateur Within qui applique l'estimateur des MCO sur le modèle (13.22) dont les données sont centrées par rapport à la moyenne individuelle. Les résultats d'estimation des paramètres $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ sont présentés dans le tableau 13.4.

Tableau 13.4 Estimation de l'équation de salaire avec des effets fixes individuels

Panel Regression — Estimation by Fixed Effects			
Dependent Variable LOGW			
Panel(5) of Annual Data From 1//1988 :01 To 31//1992 :01			
Usable Observations	155	Degrees of Freedom	119
Centered R**2	0.995868	R Bar **2	0.994652
Uncentered R**2	0.999999	T x R**2	155.000
Mean of Dependent Variable	11.363483654		
Std Error of Dependent Variable	0.163515148		
Standard Error of Estimate	0.011957673		
Sum of Squared Residuals	0.0170153263		
Regression F(35,119)	819.3641		
Significance Level of F	0.00000000		

	Variable	Coeff	Std Error	T-Stat	Signif
1.	LOGWE	0.966770339	0.023413123	41.29181	0.00000000
2.	LOGP	-0.020309437	0.016236830	-1.25083	0.21345196
3.	LOGM	-0.033985363	0.011697494	-2.90535	0.00437517
4.	U	-0.012555942	0.182391575	-0.06884	0.94523206
5.	PNS	0.119406451	0.019698541	6.06169	0.00000002

Remarque : le choix d'un modèle à effets fixes individuels ou à effets aléatoires individuels dépend souvent de l'échantillon de données, cet échantillon concerne-t-il des individus particuliers (pays de l'OCDE, secteur d'activités, firmes automobile,...) ou des individus choisis de manière aléatoire : dans le

premier cas on aura tendance à préférer un modèle à effets fixes, dans le second un modèle à effets aléatoires.

3. Lorsque le modèle à effets fixes individuels est estimé avec l'estimateur Within, on ne tient pas compte du problème d'endogénéité soulevé dans le problème 13.5, les variables *Logwe* et *Logu* sont endogènes. L'estimateur est par conséquent biaisé. Pour obtenir un estimateur non biaisé Hsiao propose d'écrire l'équation en différence première puis d'utiliser l'estimateur des Variables Instrumentales, les variables instrumentales étant soit les variables endogènes décalées de deux périodes soit les variables endogènes en différence première décalées d'une période. L'estimateur ainsi obtenu est convergent. Le tableau suivant présente les résultats lorsque les variables sont décalées de deux périodes.

Tableau 13.5 Estimation de l'équation de salaire en différence avec la méthode des Variables Instrumentales

Linear Regression — Estimation by Instrumental Variables					
Dependent Variable DLOGW					
Panel(5) of Annual Data From 1//1990 :01 To 31//1992 :01					
Usable Observations	93	Degrees of Freedom	88		
Total Observations	153	Skipped/Missing	60		
Centered R**2	0.316004	\bar{r} Bar **2	0.284913		
Uncentered R**2	0.872441	T x R**2	81.137		
Mean of Dependent Variable	0.0359413833				
Std Error of Dependent Variable	0.0173017208				
Standard Error of Estimate	0.0146308214				
Sum of Squared Residuals	0.0188373623				
Durbin-Watson Statistic	1.694448				

	Variable	Coeff	Std Error	T-Stat	Signif
1.	DLOGWE	1.000886351	0.165149011	6.06050	0.00000003
2.	DLOGP	-0.084629342	0.146897018	-0.57611	0.56600907
3.	DLOGM	-0.030835065	0.074024672	-0.41655	0.67802100
4.	DU	-0.146120289	0.401231617	-0.36418	0.71659764
5.	DPNS	0.225014763	0.200052985	1.12478	0.26374263

Les résultats de ces estimations sont bons en termes de signe des coefficients mais les coefficients ne sont pas significativement différents de 0 sauf pour la variable *Logwe*.

Dans ce chapitre on tente de tester la validité empirique du modèle du droit à gérer pour le marché du travail des salariés non qualifiés avec des données sectorielles françaises. Les résultats obtenus sont peu satisfaisants et ne conduisent pas de manière robuste à valider le modèle théorique. Ce modèle ne doit cependant pas être rejeté car les données utilisées ne sont pas totalement satisfaisantes en ce qui concerne notamment le salaire extérieur et le pouvoir de négociation. De manière générale, il est difficile de tester empiriquement la validité des modèles de négociations salariales en raison du manque de données microéconomiques correspondant aux variables du

modèle théorique mais également du problème lié à la spécification de ces modèles car la formalisation de la fonction objectif du syndicat reste un problème majeur.

Problème 13.3

Comparaison des estimateurs MCO, MCG, Within et Between

Cette question s'éloigne du modèle économique, il s'agit uniquement de vérifier les relations entre les différents estimateurs.

L'estimateur des MCO est une moyenne pondérée des estimateurs Within et Between. Dans l'exercice proposé, on obtient la matrice de pondération

$$\begin{bmatrix} 0.994 & 0.599 & -0.460 \\ -0.002 & 0.143 & 0.019 \\ -0.002 & -0.004 & 0.007 \end{bmatrix}$$

Ainsi on constate que

$$\begin{array}{l} \begin{bmatrix} 0.842 \\ -0.101 \\ -0.072 \end{bmatrix} \\ \text{estimateur MCO} \end{array} = \begin{bmatrix} 0.994 & 0.599 & -0.460 \\ -0.002 & 0.143 & 0.019 \\ -0.002 & -0.004 & 0.007 \end{bmatrix} \begin{array}{l} \begin{bmatrix} 0.979 \\ -0.022 \\ -0.048 \end{bmatrix} \\ \text{estimateur Within} \end{array} + \\ \begin{array}{l} \begin{bmatrix} (1 - 0.994) & 0.599 & -0.460 \\ -0.002 & (1 - 0.143) & 0.019 \\ -0.002 & -0.004 & (1 - 0.007) \end{bmatrix} \\ \text{estimateur Between} \end{array} \begin{array}{l} \begin{bmatrix} -31.66 \\ -0.034 \\ 0.006 \end{bmatrix} \end{array}$$

On retrouve également que l'estimateur MCG est une moyenne pondérée des estimateurs Within et Between avec une matrice de pondération égale à

$$\begin{bmatrix} 0.999 & 0.028 & -0.307 \\ 0.00005 & 0.953 & 0.071 \\ -0.001 & -0.010 & 0.432 \end{bmatrix}$$

Par conséquent

$$\begin{array}{l} \begin{bmatrix} 0.973 \\ -0.025 \\ -0.063 \end{bmatrix} \\ \text{estimateur MCG} \end{array} = \begin{bmatrix} 0.999 & 0.028 & -0.307 \\ 0.00005 & 0.953 & 0.071 \\ -0.001 & -0.010 & 0.432 \end{bmatrix} \begin{array}{l} \begin{bmatrix} 0.979 \\ -0.022 \\ -0.048 \end{bmatrix} \\ \text{estimateur Within} \end{array} + \\ \begin{array}{l} \begin{bmatrix} (1 - 0.999) & 0.028 & -0.307 \\ 0.00005 & (1 - 0.953) & 0.071 \\ -0.001 & -0.010 & (1 - 0.432) \end{bmatrix} \\ \text{estimateur Between} \end{array} \begin{array}{l} \begin{bmatrix} -31.66 \\ -0.034 \\ 0.006 \end{bmatrix} \end{array}$$

14

IMPERFECTION DU MARCHÉ DU CAPITAL ET DÉCISIONS D'INVESTISSEMENT

1. Introduction
2. Le modèle microéconomique
3. Spécification économétrique
4. Application empirique : Estimations sur données de panel françaises
5. Éléments de corrigé
6. Annexe

OBJECTIFS DU CHAPITRE

D'un point de vue méthodologique

- Présentation d'un programme microéconomique d'optimisation dynamique
- Dérivation et spécification empirique de l'équation d'Euler
- Méthodes d'estimation d'une équation dynamique à partir de données de panel
- Estimateur de la méthode des moments généralisés
- Mise en œuvre et interprétation du test de Sargan

D'un point de vue économique

- Modélisation des comportements d'investissement
- Rôle des imperfections du marché du capital
- Effet des contraintes d'endettement
- Effet de l'existence de coûts de transaction

NIVEAU

MASTER, deuxième année

CHAMPS ÉCONOMIQUES TRAITÉS

Microéconomie

Programme d'optimisation dynamique

1. INTRODUCTION

Pour financer ses investissements, une entreprise peut recourir à des emprunts (financements externes) ou elle peut faire appel à ses fonds propres (financements internes). Dans un univers financier parfait, les financements externes et internes sont parfaitement substituables et les décisions d'investissement d'une entreprise sont indépendantes de leur mode de financement et de facteurs financiers. Dans un univers financier imparfait, il existe une dépendance entre les décisions d'investissement et de financement. Depuis la fin des années 1980, de nombreux travaux micro-économiques se sont développés pour modéliser l'influence des contraintes de financement sur le comportement des agents. Ils mettent en évidence l'impact des imperfections du marché du capital dues principalement à une asymétrie d'information entre les emprunteurs et les prêteurs. Ces problèmes d'information peuvent prendre diverses formes et sont liés aux risques associés au projet d'investissement. N'étant pas certain de la qualité des projets d'investissement et ne pouvant pas contrôler parfaitement les comportements des emprunteurs, les intermédiaires financiers utilisent différents mécanismes (exigence de garanties contractuelles, évaluation des projets) pour sélectionner les firmes possédant des capacités réelles de remboursement.

Ces critères de sélection ont deux effets principaux. Tout d'abord, les emprunteurs peuvent être restreints sur leur niveau d'endettement. Les restrictions du montant de l'endettement sont généralement basées sur les caractéristiques financières et sur les avoirs des emprunteurs. De plus, ces critères créent un ensemble de coûts de transaction ex-ante qui représentent une barrière supplémentaire pour accéder au marché du capital (Williamson, 1985)¹. L'asymétrie d'information augmente ainsi le coût relatif des fonds externes par rapport au coût des fonds internes et réduit la substitution entre ces deux sources de financement.

Les conséquences des imperfections sur le marché du capital ont fait l'objet de nombreuses études empiriques. La plupart des applications confirment l'existence d'asymétrie d'information et montrent que l'on ne peut pas considérer un investissement indépendamment de son mode de financement (endettement ou fonds internes).

L'objet de cette application est de rechercher les déterminants des décisions d'investissement des exploitants agricoles français en présence d'imperfections du marché du capital.

Nous présentons tout d'abord un modèle structurel pour analyser le comportement d'investissement de l'exploitant en prenant en compte à la fois l'existence d'un plafond d'endettement, et l'existence de coûts de transactions, ces derniers apparaissant quand de nouveaux emprunts sont contractés. Cette approche structurelle permet de tester à partir d'équations d'Euler l'hypothèse selon laquelle les imperfections du marché du capital affectent les décisions d'investissement des exploitations agricoles. Le modèle de comportement est développé dans la seconde section. La section 3 précise les principes d'estimation. L'application empirique dé-

¹ Williamson O.E. (1985) *The Economic Institutions of Capitalism*, New York, The Free Press.

critère dans la section 4 propose sur des spécifications construites à partir d'un modèle accélérateur-profit, d'un modèle d'investissement basé sur une équation d'Euler.²

2. LE MODÈLE MICROÉCONOMIQUE²

Nous supposons que le ménage gérant l'exploitation agricole maximise, jusqu'à un horizon T , sous diverses contraintes, la somme espérée et actualisée des ressources générées par l'exploitation. Le programme de comportement s'écrit de la manière suivante :

$$\max E_t \sum_{\tau=t}^T \theta'_\tau (\pi_\tau(K_\tau, L_\tau, I_\tau, A_\tau) - r_{\tau-1}d_{\tau-1} + b_\tau - g(b_\tau) - a_\tau) \quad (14.1)$$

$$\pi_\tau(K_\tau, L_\tau, I_\tau, A_\tau) - r_{\tau-1}d_{\tau-1} + b_\tau - g(b_\tau) - a_\tau \geq \bar{C}_\tau \quad (a)$$

$$K_\tau = (1 - \delta)K_{\tau-1} + I_\tau \quad (b)$$

$$d_\tau = d_{\tau-1} + b_\tau - a_\tau \quad (c)$$

$$d_\tau \leq \bar{d}_\tau \quad (d)$$

$$a_\tau \geq \bar{a}_\tau \quad (e)$$

$$b_\tau \geq 0 \quad (f)$$

$$d_T \leq 0 \quad (g)$$

$$d_{t-1} = \bar{d}_{t-1} \quad (h)$$

$$K_{t-1} = \bar{K}_{t-1} \quad (i)$$

où E_t - représente l'espérance à l'instant t , θ'_t - le taux d'escompte exogène, $\pi_\tau(\cdot)$ - la fonction de profit définissant pour chaque période le revenu net de l'exploitation, K_τ - le stock de capital au début de la période, L_τ - le vecteur des facteurs de production

² La partie structurelle reprend un modèle explicité dans l'article suivant Benjamin C., Phimister E., 1997, Transaction costs, farm finance and investment, European Review of Agricultural Economics, vol.24 n° 3/4, 453-466.

variables, I_τ - le niveau d'investissement, A_τ - le vecteur des facteurs fixes, r_τ - le taux d'intérêt, $d_{\tau-1}$ - le montant des dettes au début de la période, b_τ - le montant des nouveaux emprunts contractés, $g(\cdot)$ - la fonction définissant les coûts de transaction, a_τ - le montant des remboursements, \bar{C}_τ - le niveau minimal de ressources (exogène), et δ le taux de dépréciation du capital constant compris entre 0 et 1.

Nous avons supposé que le ménage agricole souhaite maximiser le revenu généré par l'exploitation. Cependant, pour chaque période, l'exploitation doit générer un revenu minimum pour faire face aux besoins de consommation du ménage. Ainsi les besoins en consommation influencent les décisions de production par la contrainte (a). Cette inégalité précise que pour chaque période, le revenu c'est-à-dire le profit de l'exploitation moins les intérêts liés aux dettes précédemment contractées plus, éventuellement un nouvel emprunt, et moins les coûts de transaction, doit au moins être égal à un niveau minimal de ressources.

La contrainte (b) décrit l'équation d'accumulation du capital. Le capital se déprécie à chaque période t du montant δK_{t-1} et s'accroît du niveau de l'investissement brut. La dépréciation est supposée constante dans le temps (Bond, Meghir, 1995)³.

La contrainte (c) définit l'endettement d_τ au début de la période comme la somme du niveau d'endettement de la période précédente et des nouveaux emprunts contractés moins le montant des remboursements.

Les contraintes d'accès au marché du crédit sont prises en compte en intégrant des coûts de transaction explicites associés aux emprunts et en définissant un plafond d'endettement.

Les coûts de transaction sont définis comme une fonction croissante et linéaire avec le niveau de l'emprunt c'est-à-dire $g(b_\tau) = \alpha_\tau b_\tau$.

Le niveau maximal d'endettement possible pour l'exploitation est défini par la contrainte (d). Le niveau de dette maximale \bar{d}_τ est fixé de manière exogène par la banque. L'introduction de cette contrainte constitue l'approche utilisée généralement pour prendre en compte les restrictions de crédit créées par l'asymétrie d'information sur le marché du capital.

La contrainte (e) permet de définir un montant minimum de remboursement au cours de la période τ noté \bar{a}_τ exigé par l'intermédiaire financier.

A l'instant t , le programme de l'exploitant agricole peut se réécrire sous la forme du programme dynamique suivant où les variables d'état sont K_{t-1} et d_{t-1} .

$$V_t(K_{t-1}, d_{t-1}) = \max \left\{ \begin{array}{l} \pi_t((1-\delta)K_{t-1} + I - t, L_t, I_t, A_t) - rt - 1d_{t-1} \\ +b_t - \alpha_t b_t - a_t + \theta_t E_t[V_{t+1}((1-\delta)K_{t-1} \\ + I_t, d_{t-1} + b_t - a_t)] \end{array} \right\} \quad (14.2)$$

³ Bond S. et Meghir C. (1995), Dynamic investment policy and the firm's financial policy, *Review of Economic Studies* 61, 197-222.

sous les contraintes

$$\begin{aligned} \pi_t((1-\delta)K_{t-1} + I_t, L_t, I_t, A_t) - r_{t-1}d_{t-1} + b_t - \alpha_t b_t - a_t &\geq \bar{C} \\ d_t &\geq \bar{d}_t \\ a_t &\geq \bar{a}_t \\ b_t &\geq 0 \end{aligned}$$

Pour résoudre ce programme, on note μ_t le multiplicateur de Lagrange associé à la contrainte de ressources minimales (contrainte (a)), η_t le multiplicateur associé à la contrainte de dettes (contrainte (d)) et φ_t et λ_t les multiplicateurs associés respectivement aux contraintes de remboursement et de nouvel emprunt.

Des conditions du premier ordre et en appliquant le théorème de l'enveloppe, l'équation d'Euler suivante est dérivée⁴ :

$$-\theta_t(1-\delta)E_t \left[(1+\mu_{t+1}) \left(\frac{\partial \pi_{t+1}}{\partial I_{t+1}} \right) \right] = -(1+\mu_t) \frac{\partial \pi_t}{\partial I_t} - (1+\mu_t) \frac{\partial \pi_t}{\partial K_t} \quad (14.3)$$

avec les conditions suivantes sur les valeurs des multiplicateurs ;

$$(1+\mu_t)(1-\alpha_t) + \lambda_t - \eta_t = \theta_t E_t [(1+\mu_{t+1})(1+r_t - \alpha_{t+1}) + \lambda_{t+1}] \quad (14.4)$$

Cette équation d'optimalité décrit l'égalité en espérance de gain et de coût marginaux d'un investissement nouveau. De plus, pour estimer les paramètres du modèle, les variables anticipées en début de période, non observées, sont remplacées par leurs valeurs réalisées ex-post (rationalité des anticipations Hayashi, 1985).

Ce modèle général comprend quatre cas qui décrivent des structures différentes du marché du capital : i) l'existence d'un marché du capital parfait c'est à dire sans coûts de transaction et sans contrainte d'endettement (dans le modèle, ce cas particulier correspond au cas où $\bar{d}_\tau \rightarrow \infty$, $\alpha = 0$), ii) l'existence d'une contrainte d'endettement uniquement (ce cas particulier apparaît quand \bar{d}_τ est fini et $\alpha = 0$), iii) l'existence uniquement de coûts de transaction (ce cas correspond au cas où $\bar{d}_\tau \rightarrow \infty$ et le coefficient α_t est différent de zéro), et iv) l'existence de contrainte d'endettement et de coûts de transaction (ce cas particulier correspond au cas où \bar{d}_τ est fini et $\alpha > 0$).

D'un point de vue empirique, ces quatre cas vont conduire à des spécifications différentes pour expliquer les variations du taux d'investissement.

Dans le premier cas (absence de coûts de transaction et de contrainte d'endettement), on retrouve l'équation d'Euler obtenue sous l'hypothèse de maximisation

⁴ Voir l'annexe pour le détail des calculs.

du profit dans un contexte de perfection du marché du capital (Bond et Meghir, 1995) c'est à dire on obtient la relation suivante⁵ :

$$-(1 - \delta) \frac{\partial \pi_{t+1}}{\partial I_{t+1}} = -(1 + r_t) \frac{\partial \pi_t}{\partial I_t} - (1 + r_t) \frac{\partial \pi_t}{\partial K_t} + v_{t+1} \quad (14.5)$$

où v_{t+1} représente l'erreur d'anticipation qui est supposée non corrélée avec le passé de la variable anticipée. Dans ce cas, l'investissement des exploitations est indépendant des décisions financières de l'exploitation.

Dans le second cas, (existence d'une limite dans le niveau d'endettement) la présence du multiplicateur η_t inobservable dans l'équation (14.4), pour les exploitations pour lesquelles la contrainte de dette est active, implique que le comportement d'investissement de ces exploitations est influencé par les variables financières.

Le troisième cas correspond à la situation où il existe uniquement des coûts de transaction. L'estimation directe du modèle n'est pas possible en raison de la présence de multiplicateurs de Lagrange non observables dans l'équation (14.3). Toutefois si l'exploitation emprunte sur deux périodes successives c'est à dire quand $b_t > 0$ et $b_{t+1} > 0$, on peut dériver, des conditions du premier ordre, en supposant des anticipations rationnelles l'équation suivante (voir l'annexe pour la démonstration de ce résultat).

$$-(1 - \delta) \frac{\partial \pi_{t+1}}{\partial I_{t+1}} = - \left(\frac{1 + r_t - \alpha_{t+1}}{1 - \alpha_t} \right) \frac{\partial \pi_t}{\partial I_t} - \left(\frac{1 + r_t - \alpha_{t+1}}{1 - \alpha_t} \right) \frac{\partial \pi_t}{\partial K_t} + v_{0t+1} \quad (14.6)$$

La spécification est donc similaire à la relation obtenue sous l'hypothèse de perfection du marché du capital, avec la différence que les coefficients des variables dépendent des coûts de transaction.

Si l'exploitation n'emprunte pas dans deux périodes successives l'équation d'Euler dépend des multiplicateurs de Lagrange. Ces derniers sont fonctions des variables d'état et des autres variables exogènes du modèle.

Le quatrième cas autorise la double existence de coûts de transaction et d'un rationnement sur le niveau de dette.

Un résultat issu du modèle théorique est le fait que pour étudier les variations du taux d'investissement, les exploitations ayant contracté des emprunts sur deux années successives doivent être distinguées des autres exploitations.

⁵ Les équations (14.3) et (14.4) sont écrites avec le coefficient des coûts de transaction égal à zéro ainsi que le multiplicateur associé à la contrainte du montant d'endettement.

3. SPÉCIFICATION ÉCONOMÉTRIQUE

Pour dériver la spécification économétrique, les fonctions de production et de coût d'ajustement sont paramétrés. Le profit d'une exploitation à l'instant t s'écrit

$$\pi_t = p_t F(K_t, L_t, A_t) - p_t G(I_t, K_t) - w_t L_t - p_t^I I_t \quad (14.7)$$

où p_t définit le prix à la production, $F(K_t, L_t, A_t)$ la fonction de production, la fonction G les coûts d'ajustement, w_t le vecteur prix pour les inputs variables et p_t^I le prix des biens d'investissement.

Pour la fonction de production, une forme Cobb-Douglas⁶ est retenue ($F(K_t, L_t, A_t) = dK_t^{\gamma_1} L_t^{\gamma_2} A_t^{1-\gamma_1-\gamma_2}$). Le coût d'ajustement est supposé quadratique et homogène de degré zéro, $G(I_t, K_t) = (1/2)bk_t(I_t/K_t - c)^2$. La valeur nette (observable) de l'output est notée Y ($Y = F - G$). Vu les hypothèses sur la fonction de production et la fonction définissant les coûts d'ajustement, la fonction $Y(K_t, L_t, A_t)$ est aussi linéaire homogène. Aussi,

$$\begin{aligned} \frac{\partial \pi_t}{\partial I_t} &= -bp_t \left(\frac{I}{K} \right)_t + bcp_t - p_t^I \\ \frac{\partial \pi_t}{\partial K_t} &= \gamma_1 p_t \left(\frac{Y}{K} \right)_t + (1/2)bp_t(\gamma_1 + 1) \left(\frac{I}{K} \right)_t^2 - bc\gamma_1 p_t \left(\frac{I}{K} \right)_t + (1/2)bp_t c^2 (\gamma_1 - 1) \end{aligned}$$

En utilisant ces formes paramétriques les variations du taux d'investissement sous l'hypothèse de perfection du marché du capital s'écrivent

$$\begin{aligned} \left(\frac{I}{K} \right)_{it+1} &= c(1 - (1/2)c(\lambda_1 - 1)\phi_{t+1} - \phi_{t+1}) + \phi_{t+1}(1 + c\lambda_1) \left(\frac{I}{K} \right)_t \\ &\quad - \phi_{t+1}(1/2)(\lambda_1 + 1) \left(\frac{I}{K} \right)_t^2 - \frac{\gamma_1}{b} \phi_{t+1} \left(\frac{Y}{K} \right)_t + \frac{\phi_{t+1}}{b} Q_t + u_{t+1} \end{aligned} \quad (14.8)$$

où $\phi_{t+1} = (p_t/p_{t+1})(1 + r_t/1 - \delta)$, $Q_t = \frac{1}{p_t} \left(\frac{rp_t^I + (p_t^I - p_{t+1}^I) + \delta p_{t+1}^I}{1 + r_t} \right)$ et u_{t+1} est le terme d'erreur.

Le terme Q_t est équivalent au coût d'usage du capital avec le numérateur égal aux intérêts plus le coût de dépréciation du capital plus le coût d'investissement.

⁶ Pour obtenir une spécification empirique de l'équation d'Euler il est nécessaire de spécifier la productivité marginale du capital. Nous avons choisi de spécifier la fonction de production. Cette forme paramétrique est certes restrictive. Toutefois l'utilisation de données de panel permet à la technologie de production via l'effet fixe de différer entre les exploitations. De plus, si la fonction de production n'est pas paramétrée une spécification similaire pour le taux d'investissement serait obtenue en supposant les marchés des facteurs terre et travail concurrentiels et l'homogénéité de degré 1 de la fonction de production. (voir Bond et Meghir).

Pour l'application empirique des données de panel sont utilisées, aussi la spécification de l'équation à estimer s'écrit :

$$\left(\frac{I}{K}\right)_{i,t+1} = \beta_{1t} \left(\frac{I}{K}\right)_{i,t} + \beta_{2t} \left(\frac{I}{K}\right)_{i,t}^2 + \beta_{3t} \left(\frac{Y}{K}\right)_{i,t} + \rho_i + \sigma_{t+1} + v_{it+1} \quad (14.9)$$

avec $\beta_{1it} = \phi_{it+1}(1 + c\lambda_1)$, $\beta_{2it} = -\phi_{it+1}(1/2)(\lambda_1 + 1)$ et $\beta_{3it} = -\frac{\gamma_1}{\delta} \phi_{it+1}$.

où ρ_i représente l'effet individuel et σ_{t+1} est l'effet temporel⁷. Ces deux effets intègrent l'effet du coût d'usage du capital (Bond et Meghir). Du modèle structurel, on déduit que le coefficient sur I/K doit être positif et ceux de $(I/K)^2$ et Y/K doivent être négatifs.

La spécification précédente correspond au cas où le marché du capital est parfait. Dans le cas où il existe uniquement des coûts de transaction, une spécification symétrique à l'équation (14.9) peut être dérivée pour les exploitations ayant emprunté sur deux périodes successives. La seule différence réside dans la définition des coefficients qui dans ce cas dépendent des coûts de transaction. Ainsi, dans ce cas, il suffit de substituer à $1 + r_t$ l'expression $(1 + r_t - \alpha_{t+1})/(1 - \alpha_t)$. Pour les exploitations n'ayant pas emprunté sur deux périodes successives, les variations du taux d'investissement doivent dépendre de variables supplémentaires.

De manière générale pour les deux autres régimes, le fait d'introduire les contraintes liées aux imperfections de marché (existence de coûts de transaction, plafond d'endettement) introduit dans l'équation d'Euler des termes difficilement estimables liés aux multiplicateurs de Lagrange des contraintes. L'optique retenue dans la partie empirique est d'introduire l'effet de ces multiplicateurs comme une fonction ad hoc de variables de l'exploitation.

4. APPLICATION EMPIRIQUE : ESTIMATIONS SUR DONNÉES DE PANEL FRANÇAISES

L'estimation s'effectue à partir d'un panel cylindré de 1471 entreprises observées sur la période 1987-1992. Pour chaque année et pour chaque entreprise nous avons des données sur les valeurs de l'investissement en capital physique observées l'année t , le stock de capital de l'entreprise en début d'année⁸, le niveau de production et le niveau d'endettement. L'objectif de l'application empirique est l'estimation d'équations de taux d'accumulation à partir de données de panel individuelles. Des spécifications construites à partir d'un modèle accélérateur-profit, d'un modèle d'investissement basé sur une équation d'Euler sont estimées.

⁷ Dans la partie empirique, on suppose que les prix sont identiques pour toutes les exploitations. Leur impact est capté par l'effet spécifique temporel qui correspond à l'omission de variables dont la valeur est identique pour tous les individus. Dans chaque spécification, des variables muettes sont introduites

⁸ Dans nos notations, K_t représente le stock de capital au début de la période t .

Problème 14.1

Construction des variables et analyse descriptive

1. Construire les différentes variables nécessaires à l'estimation : le taux d'accumulation, le ratio entre le niveau de production et le stock de capital, le ratio entre profit et stock de capital, le rapport entre niveau d'endettement et le stock de capital.
2. Présenter et commenter les statistiques descriptives des variables nécessaires à l'estimation et des principales variables en niveau.

Problème 14.2

Estimation d'un modèle accélérateur profit

1. Proposer de manière ad hoc une spécification dynamique basée sur un modèle accélérateur-profit expliquant les variations du taux d'accumulation à partir de données de panel. Interpréter chaque terme de l'équation.
2. Comment prendre en compte les effets temporels ?
3. Estimer l'équation en utilisant un modèle à effets fixes (modèle Within). Est-ce la méthode appropriée ? Proposer une méthode alternative.
4. L'équation est estimée en utilisant la méthode des moments généralisés. Lister les instruments à utiliser. Estimer l'équation. Quels sont les tests à appliquer pour s'assurer de la validité de l'équation ?
5. Quels critiques peut-on faire sur la spécification utilisée ?

Problème 14.3

Estimation d'un modèle basé sur l'équation d'Euler

Les estimations empiriques sont basées sur l'équation (14.9).

1. Ecrire l'équation d'Euler permettant d'expliquer les variations du taux d'accumulation en l'absence de contraintes de liquidité.
2. Estimer l'équation en utilisant la méthode des moments. Après avoir listé les instruments, commenter les résultats. Peut-on conserver l'hypothèse de perfection des marchés financiers ?
3. Proposer une spécification permettant de prendre en compte les contraintes financières et l'existence d'un plafond d'endettement.
4. Quelles améliorations pourrait-on proposer pour estimer cette spécification ?

5. ÉLÉMENTS DE CORRIGÉ

Problème 14.1

Construction des variables et analyse descriptive

1. La variable expliquée au cours de l'analyse empirique est le ratio du niveau d'investissement sur le stock de capital ou taux d'accumulation. Parmi les

variables explicatives qui pourront être utilisées, il faut construire le rapport entre production et stock de capital, le rapport entre profit et stock de capital et le rapport entre le niveau d'endettement et le stock de capital.

- Présentation et commentaires sur les statistiques descriptives des principales variables.

Le tableau suivant présente la moyenne pour chaque variable, le chiffre entre parenthèses donne la valeur de l'écart-type. Les statistiques descriptives (tableau 14.1) montrent un ralentissement de l'investissement et une baisse du niveau de profit depuis le début des années 1990. L'évolution du taux d'accumulation est cohérente avec celles de ces deux variables. Après une augmentation, au début de période le taux de profit décroît à partir de 1990. Par ailleurs, le taux d'endettement est en diminution sur l'ensemble de la période.

Tableau 14.1 Définitions des variables utilisées et statistiques descriptives sur la période 1987-1992

	1987	1988	1989	1990	1991	1992
Investissement	477,83 (825,86)	513,40 (845,59)	529,81 (886,01)	505,95 (925,40)	444,13 (763,35)	408,27 (827,70)
Capital	2385,28 (2698,27)	2253,34 (2094,50)	2231,10 (2081,41)	2168,78 (2171,18)	2181,20 (2190,80)	2284,36 (2400,47)
Profit	1510,23 (1209,67)	1699,09 (1519,62)	1936,99 (1920,56)	1796,32 (1957,94)	1624,07 (1637,85)	1676,49 (2239,06)
I/K taux d'accumulation	0,192 (0,22)	0,204 (0,23)	0,203 (0,23)	0,215 (0,52)	0,186 (0,23)	0,167 (0,22)
Y/K rapport production/ capital	4,394 (8,88)	4,739 (9,81)	5,261 (12,54)	6,004 (16,24)	7,977 (47,17)	7,614 (37,70)
π /K Profit/capital	1,507 (0,53)	1,603 (4,18)	1,977 (6,06)	2,381 (11,86)	2,787 (20,66)	2,524 (18,61)
empr/K rapport endettement/ capital	2,361 (6,932)	2,303 (6,85)	2,401 (7,35)	2,406 (6,17)	2,838 (17,62)	2,114 (5,06)

Problème 14.2

Estimation d'un modèle accélérateur profit

- Écriture de la spécification d'un modèle accélérateur profit

Dans le modèle accélérateur, l'investissement dépend de la demande et augmente avec l'accélération de celle-ci. Le rapport entre le niveau de production et le stock de capital permet de mesurer l'intensité du phénomène d'accélération. Le modèle d'accélérateur simple peut être enrichi par la prise

en compte de variables de profit. L'investissement est fonction non seulement de la croissance des débouchés mais aussi d'une variable de profit. Cette modélisation traduit l'idée qu'une partie des entreprises est contrainte sur la demande et une autre partie sur les conditions de financement.

Dans ce cadre d'analyse, l'équation à estimer s'écrit

$$\left(\frac{I}{K}\right)_{i,t+1} = \alpha_1 \left(\frac{I}{K}\right)_{i,t} + \alpha_2 \left(\frac{Y}{K}\right)_{i,t} + \alpha_3 \left(\frac{\pi}{K}\right)_{i,t} + \rho_i + \sigma_{t+1} + v_{it+1} \quad (14.10)$$

où $\left(\frac{I}{K}\right)_{i,t+1}$ représente le taux d'accumulation, $\left(\frac{Y}{K}\right)_{i,t}$ le terme "accélérateur" et le $\left(\frac{\pi}{K}\right)_{i,t}$ terme financier (rapport entre le profit de l'entreprise et le niveau de capital).

Le terme ρ_i représente l'effet individuel fixe, spécifique à chaque entreprise. Cet effet peut être analysé comme une constante inobservable qui résume l'ensemble des caractéristiques propres à chaque firme par exemple la "qualité du management", le "savoir-faire" du chef d'entreprise.

Le terme σ_{t+1} représente l'effet temporel et le terme v_{it+1} le terme d'erreur classique.

2. Prise en compte des effets temporels

Comme la période d'estimation est courte ($T=6$) et la taille de l'échantillon importante on estime les effets temporels σ_{t+1} en les traitant comme un ensemble de variables indicatrices temporelles dans toutes les équations.

3. Estimation de l'équation en utilisant un modèle Within. Discussion de la méthode d'estimation.

Un traitement fixe de l'équation précédente conduit à l'estimateur Within (moindre carrés après avoir pris les écarts aux moyennes individuelles). La transformation intra-individuelle revient à éliminer les effets fixes individuels en soustrayant de l'équation précédente la moyenne individuelle de l'équation. L'équation suivante est obtenue

$$\frac{I}{K}_{i,t+1} - \overline{\frac{I}{K}}_{i,+1} = \alpha_1 \left(\frac{I}{K}_{i,t} - \overline{\frac{I}{K}}_{i,\cdot}\right) + \alpha_2 \left(\frac{Y}{K}_{i,t} - \overline{\frac{Y}{K}}_{i,\cdot}\right) + \alpha_3 \left(\frac{\pi}{K}_{i,t} - \overline{\frac{\pi}{K}}_{i,\cdot}\right) + (\sigma_{t+1} - \bar{\sigma}) + (v_{it+1} - \bar{v}_i) \quad (14.11)$$

$$\text{où } \overline{\frac{I}{K}}_{i,+1} = \frac{1}{T+1} \sum_{t=1}^{T+1} \frac{I}{K}_{i,t}$$

$$\overline{\frac{I}{K}}_{i,\cdot} = \frac{1}{T} \sum_{t=1}^T \frac{I}{K}_{i,t}$$

$$\overline{\frac{Y}{K}}_{i,\cdot} = \frac{1}{T} \sum_{t=1}^T \frac{Y}{K}_{i,t}$$

$$\overline{\frac{\pi}{K}}_{i,\cdot} = \frac{1}{T} \sum_{t=1}^T \frac{\pi}{K}_{i,t}$$

Les résultats de l'estimation de cette équation sont présentés dans le tableau suivant.

Tableau 14.2 Résultats d'estimation du modèle accélérateur profit en utilisant un modèle à effets fixes (modèle Within)

Variabiles	Paramètres estimés	Écart type
Taux d'accumulation décalé	-0,188	0,012
Ratio production sur capital	$-0,722 \cdot 10^{-2}$	$0,450 \cdot 10^{-3}$
Taux de profit	0,015	$0,913 \cdot 10^{-3}$

Le paramètre de la variable production sur capital est significativement négatif dans l'équation.

Discussion de la méthode d'estimation.

L'existence de la variable retardée soulève la question de la simultanéité de cette variable avec l'erreur résiduelle.

Dans l'équation (14.11), la variable expliquée décalée ($\frac{I}{K}_{i,t} - \bar{\frac{I}{K}_i}$) est corrélée avec le terme ($v_{it+1} - \bar{v}_i$). Aussi, l'estimateur Within n'est pas convergent puisque cette transformation induit une corrélation entre la variable expliquée décalée et le terme d'erreur.

L'estimateur intra-individuel est donc biaisé. La solution consiste à utiliser une méthode d'estimation par variables instrumentales sur l'équation (14.10) écrite en différences premières. On raisonne en différences premières pour supprimer l'effet fixe, on applique l'estimateur de la méthode des moments généralisés en utilisant comme instruments les variables décalées (question suivante).

4. Estimation de l'équation en utilisant la méthode des moments.

L'équation (14.10) est estimée en utilisant l'estimateur la méthode des moments à partir de l'année 1989, les années 1987,1988 permettent d'avoir de l'information sur les variables décalées. Les estimations sont basées sur la méthode de déviations orthogonales proposée par Arellano et Bond. En supposant que les termes d'erreurs de l'équation (14.10) en niveau v_{it+1} ne soient pas autocorrélés, leur différence première doit être orthogonale aux variables explicatives de l'équation (14.10) décalées d'au moins deux périodes. Ainsi, les variables instrumentales utilisées sont le taux d'accumulation, le ratio production sur capital et le taux de profit retardé de 2 à 4 années : $(I/K)_{t-2}$, $(I/K)_{t-3}$, $(I/K)_{t-4}$, $(Y/K)_{t-2}$, $(Y/K)_{t-3}$, $(Y/K)_{t-4}$ et $(\pi/K)_{t-2}$, $(\pi/K)_{t-3}$, $(\pi/K)_{t-4}$. On obtient ainsi 30 instruments (hors constante). Le nombre de conditions d'orthogonalité est égal à 34 et il y a 3 paramètres à estimer. Si les termes d'erreurs en niveau suivent un processus de moyenne mobile d'ordre 1, les premiers instruments valides doivent être pris au 3^{ème} décalage plutôt qu'au second retard.

La validité des instruments est testée par le test général des restrictions sur-identifiantes dit test de Sargan et les tests m1, m2 pour l'autocorrélation des

erreurs à un an, deux ans. La validité des instruments choisis est vérifiée en testant que les variables instrumentales ne soient pas corrélées avec le terme résiduel de l'équation. Le test de Sargan teste ainsi si les moments de l'échantillon correspondant aux conditions d'orthogonalité sont suffisamment proches de zéro. Sous cette hypothèse, la statistique de Sargan suit une loi du khi-deux (avec un degré de liberté égal au nombre de restrictions). Le rejet de l'hypothèse signifie que le modèle de comportement n'est pas valide ou que les instruments ne sont pas correctement choisis. Dans l'application, le nombre de degrés de libertés est égal à 31.

Les statistiques m1 et m2 sont calculées. La statistique m1 permet de tester l'autocorrélation à l'ordre 1 des résidus. La statistique m2 teste si la différence première des résidus n'est pas corrélée. Si cette hypothèse est rejetée, les instruments datés en t-2 ne sont pas valides (Arrelano et Bond, 1991). Les résultats de l'estimation de la spécification " accélérateur profit " sont présentés dans le tableau suivant.

Tableau 14.3 Estimation du modèle d'accélérateur profit en utilisant la méthode des moments

Variables	Paramètres estimés	Écart type
Taux d'accumulation décalé	-0,0495	0,020
Ratio production sur capital	$0,361 \cdot 10^{-2}$	$0,857 \cdot 10^{-3}$
Taux de profit	$0,240 \cdot 10^{-2}$	$0,101 \cdot 10^{-2}$
Test de Sargan	75,82 (31)	
Valeur critique à 5 %	44,99	
Valeur critique à 10 %	52,19	
m1	-1,64	
m2	-1,068	

L'hypothèse d'existence de contrainte de liquidité semble ne pas être rejetée puisque le coefficient de la variable de profit est positif et statistiquement significatif. Mais cette spécification est rejetée au titre du test de Sargan.

Le modèle est rejeté si la valeur de la statistique de Sargan est supérieure à la valeur donnée par la table du khi deux. A $5 \chi^2(31) = 44,99$. La statistique calculée est supérieure à cette valeur, ce qui conduit au rejet de l'hypothèse nulle (validité des instruments et de la spécification). Le modèle accélérateur profit n'est donc pas validé par les données.

5. Critiques sur la spécification utilisée

L'équation économétrique basée sur un modèle accélérateur profit est écrite de manière ad hoc, c'est à dire elle n'est pas déduite d'un modèle structural. Cet empirisme pose alors des problèmes d'interprétation. Ainsi on peut difficilement tirer des conclusions de la significativité de la variable financière dans l'équation précédente : cette variable peut représenter les possibilités financières de l'entreprise ou elle peut représenter un indicateur de profitabilité future de l'entreprise.

Problème 14.3

Estimation d'un modèle basé sur l'équation d'Euler

1. Écriture de l'équation à estimer

L'équation d'Euler est dérivée d'un modèle structurel expliquant le choix d'investissement en présence de contraintes financières. Les estimations empiriques sont basées sur l'équation (14.9)

$$\left(\frac{I}{K}\right)_{i,t+1} = \beta_1 \left(\frac{I}{K}\right)_{i,t} + \beta_2 \left(\frac{I}{K}\right)_{i,t}^2 + \beta_3 \left(\frac{Y}{K}\right)_{i,t} + \rho_i + \sigma_{t+1} + v_{it+1} \quad (14.12)$$

L'existence de la variable retardée soulève la question de la simultanéité de cette variable avec l'erreur résiduelle. La solution est de raisonner en différence première pour supprimer l'effet fixe et d'appliquer l'estimateur de la méthode des moments généralisés en utilisant comme instruments les variables décalées.

2. Méthode d'estimation retenue, tests appliqués pour s'assurer de la validité de chaque modèle

Les résultats de l'estimation de la spécification basée sur une équation d'Euler sont présentés dans le tableau suivant. L'équation est estimée sur la période 1989-1992.

Tableau 14.4 Estimation de l'équation d'Euler sous l'hypothèse du marché du capital parfait

Variabiles	Paramètres estimés	Écart type
Taux d'accumulation décalé	0,38089	0,020
Taux d'accumulation décalé (au carré)	-0,036093	0,692*10 ⁻²
Ratio production sur capital	0,459*10 ⁻²	0,841*10 ⁻³
Test de Sargan	67,31 (31)	
Valeur critique à 5 %	44,99	
Valeur critique à 10 %	52,19	
m1	-1,64	
m2	-1,068	

Deux critères sont utilisés pour s'assurer de la validité du modèle estimé : le test de Sargan et les signes des coefficients. Ces deux critères ne sont pas vérifiés.

Pour les signes des coefficients, le modèle théorique montre que le coefficient de la variable endogène décalée doit être positif et ceux de $(I/K)^2$ et (Y/K) doivent être négatifs. La variable (ratio de la production sur capital) n'a pas le bon signe.

Par ailleurs, le test de suridentification conduit au rejet de la spécification.

Les instruments utilisés pour cette spécification sont $(I/K)_{t-2}$, $(I/K)_{t-3}$, $(I/K)_{t-4}$, $(I/K)_{t-2}^2$, $(I/K)_{t-3}^2$, $(I/K)_{t-4}^2$, et $(Y/K)_{t-2}$, $(Y/K)_{t-3}$.

$(Y/K)_{t-4}$. Il y a 34 conditions liées au moments, il y a 3 paramètres à estimer, la statistique de Sargan a 31 degrés de liberté. Le modèle est rejeté si la valeur de la statistique de Sargan est supérieure à la valeur donnée par la table du khi deux. A 5 %, la valeur lue dans la table est $\chi^2(31) = 44,99$. La statistique calculée est supérieure à cette valeur, ce qui conduit au rejet de l'hypothèse nulle (validité des instruments et de la spécification). L'hypothèse de perfection du marché du capital ne peut donc pas être acceptée.

3. Spécification en intégrant les contraintes financières.

La seconde hypothèse testée est l'existence d'une imperfection du marché du capital due à une contrainte sur le montant d'endettement. Lorsqu'on ne fait plus l'hypothèse de marchés financiers parfaits, l'expression de l'équation d'Euler est modifiée (voir équation 14.4). La difficulté empirique est qu'apparaît alors dans l'équation d'Euler le multiplicateur de Lagrange de la contrainte de financement qui est une variable inobservable fonction des caractéristiques de l'entreprise. Une possibilité est de rajouter dans l'équation précédente des termes pouvant mesurer l'importance de ces contraintes financières. Dans ce cadre l'équation à estimer s'écrit

$$\begin{aligned} \left(\frac{I}{K}\right)_{i,t+1} &= \beta_1 \left(\frac{I}{K}\right)_{i,t} + \beta_2 \left(\frac{I}{K}\right)_{i,t}^2 + \beta_3 \left(\frac{Y}{K}\right)_{i,t} + \beta_4 \left(\frac{\pi}{K}\right)_{i,t} \\ &\quad + \beta_5 \left(\frac{empr}{K}\right)_{i,t} + \rho_{i,t} + \sigma_{t+1} + v_{it+1} \end{aligned}$$

La spécification estimée est l'équation (14.9) où on rajoute deux variables explicatives : le rapport des profits sur le stock de capital (π/K) , et le rapport des emprunts sur le stock de capital, $(empr/K)$. Les instruments sont ceux de la première spécification ainsi que les variables $(\pi/K)_{t-2}$, $(\pi/K)_{t-3}$, $(\pi/K)_{t-4}$, et, $(empr/K)_{t-2}$, $(empr/K)_{t-3}$, $(empr/K)_{t-4}$. Cette spécification est dénommée " l'équation d'Euler augmentée ", équation augmentée des variables financières. Les résultats sont présentés dans le tableau 14.5.

Pour rappel, d'après le modèle théorique l'impact du taux d'accumulation de la variable endogène décalée doit être positif et les impacts des variables de $(I/K)^2$ et (Y/K) doivent être négatifs sur les variations du taux d'accumulation. Ces enseignements sont respectés dans les résultats présentés. Par ailleurs, la spécification n'est pas rejetée suivant le critère de Sargan suivant une probabilité critique de 10 %. Le taux d'accumulation est positivement influencé par la variable de profit. Ce résultat semble indiquer l'existence de contraintes de liquidité. L'estimation est améliorée par rapport au cadre sans contrainte financière (tableau 14.4). L'équation d'Euler augmentée d'une contrainte financières n'est pas rejetée par les données. De plus, les coefficients associés aux deux ratios de variables financières sont significatifs.

Tableau 14.5 Estimation de l'équation d'Euler augmentée des variables financières

Variabiles	Paramètres estimés	Écart type
Taux d'accumulation décalé	+0,0413	0,015
Taux d'accumulation décalé au carré	-0,0309	$0,435 \cdot 10^{-2}$
Ratio production sur capital	$-0,422 \cdot 10^{-2}$	$0,914 \cdot 10^{-3}$
Taux de profit	$0,639 \cdot 10^{-2}$	$0,682 \cdot 10^{-2}$
Niveau d'endettement sur stock de capital	$0,885 \cdot 10^{-2}$	$0,168 \cdot 10^{-2}$
Test de Sargan	69,07 (49)	
Valeur critique à 5 %	66,34	
Valeur critique à 10 %	74,92	
m1	-1,87	
m2	-1,01	

4. Améliorations pour estimer cette spécification

L'estimation pourrait être réalisée sur des sous échantillons d'entreprises classées a priori suivant leur facilité d'accès au marché du crédit (taille, âge, formation du chef d'entreprise). La difficulté rencontrée dans la partie empirique est l'interprétation du rejet de l'équation (14.12). L'approche traditionnelle est de scinder l'échantillon en groupes d'exploitations pour lesquelles on pense a priori qu'elles ne sont pas affectées par des restrictions quantitatives d'endettement. Les critères utilisés correspondent à la taille des exploitations, à la formation du chef de l'exploitation. Si le rejet de l'équation d'Euler est dû uniquement à l'existence de contrainte d'endettement, le rôle des variables financières doit être observé uniquement pour le groupe a priori contraint.

Un autre prolongement serait aussi d'appliquer un résultat issu du modèle structurel : pour étudier les variations du taux d'accumulation, les exploitations ayant contracté des emprunts sur deux années successives doivent être distinguées des autres exploitations. D'après les enseignements du modèle si les résultats sont cohérents avec la double existence de coûts de transaction et de contrainte d'endettement, l'équation (14.9) ne sera pas rejetée par le sous-échantillon des exploitations ayant contracté des emprunts sur des périodes successives.⁹

⁹ Voir Benjamin C., Phimister E., 2001, Imperfection du marché du capital et investissement des exploitations agricoles, vol.77, n° 3, *L'Actualité Économique*, 357-383 pour une illustration de ce point.

6. ANNEXE

6.1 Dérivation des équations (14.3), (14.4) et (14.6)

Le Lagrangien associé au programme d'optimisation (14.2) pour la période t s'écrit :

$$\begin{aligned}
 L = & \pi_t ((1 - \delta)K_{t-1} + I_t, L_t, I_t, A_t) - r_{t-1}d_{t-1} + b_t - \alpha b_t - a_t \\
 & + \theta_t E_t [V_{t+1} ((1 - \delta)K_{t-1} + I_t, d_{t-1} + b_t - a_t)] \\
 & + \mu_t (\pi_t ((1 - \delta)K_{t-1} + I_t, L_t, I_t, A_t) - r_{t-1}d_{t-1} + b_t - \alpha b_t - a_t - \bar{C}_t) \\
 & + \varphi_t (a_t - \bar{a}_t) \\
 & + \eta_t (\bar{d} - d_{t-1} - b_t + a_t) \\
 & + \lambda_t b_t
 \end{aligned}$$

Les conditions du premier ordre sont :

$$L_t : (1 + \mu_t) \frac{\partial \pi_t}{\partial L_t} = 0 \quad \text{([I])}$$

$$I_t : (1 + \mu_t) \left(\frac{\partial \pi_t}{\partial I_t} + \frac{\partial \pi_t}{\partial K_t} \right) + \theta_t E_t \left[\frac{\partial V_{t+1}}{\partial K_t} \right] = 0 \quad \text{([II])}$$

$$b_t : (1 + \mu_t) (1 - \alpha) + \theta_t E_t \left[\frac{\partial V_{t+1}}{\partial d_t} \right] + \lambda_t - \eta_t = 0 \quad \text{([III])}$$

$$a_t : -(1 + \mu_t) - \theta_t E_t \left[\frac{\partial V_{t+1}}{\partial d_t} \right] + \varphi_t = 0 \quad \text{([IV])}$$

$$\begin{aligned}
 \mu_t : & \pi_t ((1 - \delta)K_{t-1} + I_t, L_t, I_t, A_t) - r_{t-1}d_{t-1} + b_t - \alpha b_t - a_t - \bar{C}_t \geq 0, \mu_t \geq 0 \\
 & \mu_t (\pi_t ((1 - \delta)K_{t-1} + I_t, L_t, I_t, A_t) - r_{t-1}d_{t-1} + b_t - \alpha b_t - a_t - \bar{C}_t) = 0
 \end{aligned} \quad \text{([V])}$$

$$\varphi_t : a_t - \bar{a}_t \geq 0, \varphi_t \geq 0, \varphi_t (a_t - \bar{a}_t) = 0 \quad \text{([VI])}$$

$$\eta_t : \bar{d} - d_{t-1} - b_t + a_t \geq 0, \eta_t \geq 0, \eta_t (\bar{d} - d_{t-1} - b_t + a_t) = 0. \quad \text{([VII])}$$

$$\lambda_t : b_t \geq 0, \lambda_t \geq 0, \lambda_t b_t = 0 \quad \text{([VIII])}$$

6.2 Équation (14.3)

Pour déterminer l'équation (14.3) on obtient à partir du Lagrangien (en appliquant le théorème de l'enveloppe) que

$$\begin{aligned}\frac{\partial V_t}{\partial K_t} &= (1 - \delta) \left((1 + \mu_t) \frac{\partial \pi_t}{\partial K_t} + \theta_t E_t \left[\frac{\partial V_{t+1}}{\partial K_{t+1}} \right] \right) \\ &= -(1 - \delta)(1 + \mu_t) \frac{\partial \pi_t}{\partial I_t} \quad (\text{partir de l'Équation [II]})\end{aligned}$$

$$\text{aussi } E_t \left[\frac{\partial V_{t+1}}{\partial K_t} \right] = -(1 - \delta) E_t \left[(1 + \mu_{t+1}) \frac{\partial \pi_{t+1}}{\partial I_{t+1}} \right] \quad \text{[VIII]}$$

En substituant [VIII] dans l'équation [II] on dérive l'équation (14.3)

6.3 Équation (14.4)

L'équation (14.4) est obtenue de façon similaire. Dans un premier temps on applique le théorème de l'enveloppe pour obtenir :

$$\begin{aligned}\frac{\partial V_t}{\partial d_{t-1}} &= -r_{t-1}(1 + \mu_t) + \theta_t E_t \left[\frac{\partial V_{t+1}}{\partial d_t} \right] - \eta_t \\ &= -(1 + \mu_t)(1 + r_{t-1} - \alpha) - \lambda_t \quad (\text{partir de l'Équation [III]})\end{aligned}$$

$$E_t \left[\frac{\partial V_{t+1}}{\partial d_t} \right] = -E_t \left[(1 + \mu_{t+1})(1 + r_t - \alpha) - \lambda_{t+1} \right] \quad \text{([IX])}$$

En intégrant la relation [IX] dans [III] on obtient l'équation (14.4).

6.4 Équation (14.6)

Tout d'abord considérons le cas où la solution optimale pour l'exploitation est de contracter des emprunts à la période t . Aussi, le multiplicateur associé à la contrainte est nul [VII] i $\lambda_t = 0$. L'équation d'Euler s'écrit en substituant pour $(1 + \mu_t)$ dans l'équation (14.3) en utilisant (14.4), :

$$\begin{aligned}E \left[(1 + \mu_{t+1}) \left((1 - \delta) \frac{\partial \pi_{t+1}}{\partial I_{t+1}} \left(\frac{1 + r_T - \alpha}{1 - \alpha} - \frac{\lambda_{t+1}}{(1 - \alpha)(1 + \mu_{t+1})} \right) \right. \right. \\ \left. \left. \left(\frac{\partial \pi_t}{\partial I_t} + \frac{\partial \pi_t}{\partial K_t} \right) \right) \right] = 0\end{aligned}$$

En supposant des anticipations rationnelles on obtient

$$\begin{aligned}- \left((1 - \delta) \frac{\partial \pi_{t+1}}{\partial I_{t+1}} - \left(\frac{1 + r_t - \alpha}{1 - \alpha} - \frac{\lambda_{t+1}}{(1 - \alpha)(1 + \mu_{t+1})} \right) \left(\frac{\partial \pi_t}{\partial I_t} + \frac{\partial \pi_t}{\partial K_t} \right) \right) \\ = \frac{e_{t+1}}{(1 + \mu_{t+1})}\end{aligned}$$

où e_{t+1} est le terme d'erreur avec une moyenne égale à zéro. Aussi pour une exploitation où $b_{t+1} > 0$ ($\Rightarrow \lambda_{t+1} = 0$) l'équation (14.6) est déterminée.

15

ÉCONOMÉTRIE DES SÉRIES NON STATIONNAIRES

1. Les modèles de série temporelle univariés
2. Processus non stationnaire et racine unitaire
3. Cointégration et MCE : l'approche de Engel et Granger
4. Cointégration et MCE : l'approche de Johansen

OBJECTIFS DU CHAPITRE

Effectuer une introduction à l'économétrie des séries temporelles

- Modèles AR, MA, ARMA
- Modélisation d'une série stationnaire univariée
- Processus non stationnaire (séries stationnaires en différence, en écart au trend)
- Tests de racine unitaire (Dickey-Fuller, KPSS)
- Cointégration : approche de Engel et Granger, approche de Johansen
- Modèle à correction d'erreur
- Test d'exogénéité faible

NIVEAU

MASTER, deuxième année

La faible performance en termes de prévision des modèles macroéconométriques à équations simultanées des années 60 a conduit à de nouvelles recherches sur l'économétrie des séries temporelles, surtout depuis les années 80. Le thème en lui-même est vaste et il s'agit ici de présenter quelques éléments importants afin de comprendre les développements récents de la macroéconométrie vers l'analyse des séries temporelles non stationnaires. Beaucoup de variables macroéconomiques caractérisées par un trend important telles que le PIB, la consommation, ... ne sont pas stationnaires et ce chapitre a pour objectif d'introduire les problèmes induits par la non stationnarité des séries, le concept de cointégration et les tests de cointégration ainsi que les Modèles à Correction d'Erreur (MCE). Dans un premier temps, on présente de manière succincte des modèles de série temporelle univariés. Dans un second temps, sont présentés les tests de stationarité. Pour terminer, le concept de cointégration, les tests de cointégration et les MCE sont examinés.

1. LES MODÈLES DE SÉRIE TEMPORELLE UNIVARIÉS

1.1 Exemples

Un modèle de série temporelle univarié décrit le comportement d'une variable en termes de ses valeurs passées. Voici trois exemples :

$$y_t = \mu + \gamma_1 y_{t-1} + \dots + \gamma_p y_{t-p} + \varepsilon_t \tag{15.1}$$

$$y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \tag{15.2}$$

$$y_t = \mu + \gamma_1 y_{t-1} + \dots + \gamma_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \tag{15.3}$$

Le premier exemple correspond à un modèle autorégressif d'ordre p noté $AR(p)$. Le second est un modèle de moyenne mobile d'ordre q noté $MA(q)$ et le dernier est un modèle de moyenne mobile autorégressif noté $ARMA(p, q)$.

Ces différents modèles peuvent s'écrire soit sous une forme MA soit sous une forme AR . Dans le cas d'un $ARMA$ d'ordre (p, q) quelconque

$$y_t (1 - \gamma_1 L - \dots - \gamma_p L^p) = \varepsilon_t (1 - \theta_1 L - \dots - \theta_q L^q) \tag{15.4}$$

avec L l'opérateur de décalage ($L^i y_t = y_{t-i}$, $L^i \varepsilon_t = \varepsilon_{t-i}$).

Soit $\Gamma(L)$ un polynôme en L de degré p et $\Upsilon(L)$ un polynôme en L de degré q

$$y_t \Gamma(L) = \varepsilon_t \Upsilon(L)$$

si les racines de $\Gamma(L) = 0$ ont un module inférieur à 1 alors la forme $MA(\infty)$ du processus $ARMA(p, q)$ s'écrit :

$$y_t = \Psi(L) \varepsilon_t \quad \text{avec} \quad \Psi(L) = \frac{\Upsilon(L)}{\Gamma(L)} \tag{15.5}$$

de même si les racines de $\Upsilon(L) = 0$ ont un module inférieur à 1, la forme $AR(\infty)$ est donnée par

$$\varepsilon_t = \Pi(L)y_t \quad \text{avec} \quad \Pi(L) = \frac{\Gamma(L)}{\Upsilon(L)} = \Psi^{-1}(L) \quad (15.6)$$

La condition que les racines de $\Gamma(L) = 0$ restent dans le cercle unitaire est la condition de stationnarité de la série, et la condition que les racines de $\Upsilon(L) = 0$ restent dans le cercle unitaire est la condition d'invertibilité. Dans un $ARMA(1,1)$ la condition de stationnarité est $|\gamma_1| < 1$ et la condition d'invertibilité est $|\theta_1| < 1$.

1.2 Définition de la stationnarité

Un processus est faiblement stationnaire si les conditions suivantes sont satisfaites

1. $E(y_t)$ est indépendante de t
2. $var(y_t)$ est une constante finie indépendante de t
3. $cov(y_t, y_{t-k})$ est une fonction finie de k et ne dépend pas de t .

Par exemple le processus $AR(1) : y_t = \gamma_1 y_{t-1} + \varepsilon_t$ est stationnaire si $|\gamma_1| < 1$. Dans ce cas

$$\begin{aligned} E(y_t) &= E(\varepsilon_t) = 0 \\ Var(y_t) &= \frac{\sigma_\varepsilon^2}{1 - \gamma_1^2} \\ cov(y_t, y_{t-k}) &= \gamma_1^k \frac{\sigma_\varepsilon^2}{1 - \gamma_1^2} \end{aligned}$$

De manière générale la condition de stationnarité ne dépend que de la partie AR de la série.

1.3 Modélisation d'une série temporelle

La modélisation d'une série temporelle consiste à déterminer les retards p et q donnant la meilleure représentation du mouvement de la série à partir de sa fonction d'autocorrélation d'une part, et sa fonction d'autocorrélation partielle d'autre part¹. Cette modélisation ne s'appuie sur aucune base économique. On parle souvent de «mesure sans théorie».

La fonction d'autocorrélation est obtenue en calculant le rapport entre la covariance des séries y_t et y_{t-k} et la variance de la série y_t

$$r_k = \frac{\frac{1}{T} \sum_{t=1}^T (y_t - \bar{y}) \cdot (y_{t-k} - \bar{y})}{\frac{1}{T} \sum_{t=1}^T (y_t - \bar{y})^2} \quad k = 1, 2, \dots \quad (15.7)$$

¹ Pour un complément sur ces fonctions cf. : Terence C. Mills, 1990, *Time Series Techniques for Economists*, Cambridge University Press.

La fonction d'autocorrélation partielle correspond à la corrélation entre y_t et y_{t-k} lorsque la partie expliquée par les variables y_{t-k-i} ($i < k$) est omise, elle est calculée de la manière suivante

$$r_{11} = r_1$$

$$r_{kk} = \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} \quad k = 2, 3, \dots \quad (15.8)$$

La représentation de la fonction d'autocorrélation c'est-à-dire le corrélogramme de la série et de sa fonction d'autocorrélation partielle permet d'identifier les caractéristiques de la série. Un processus $AR(p)$ a une fonction d'autocorrélation qui décroît de manière exponentielle et/ou sinusoidale et une fonction d'autocorrélation partielle avec des pics pour les p premiers retards. Un processus $MA(q)$ a une fonction d'autocorrélation avec des pics pour les q premiers retards et une fonction d'autocorrélation partielle qui décroît de manière exponentielle et/ou sinusoidale.

Les figures 15.1 et 15.2 présentent respectivement le corrélogramme d'une série caractérisée par un processus $AR(2)$ et une série caractérisée par un processus $MA(3)$.

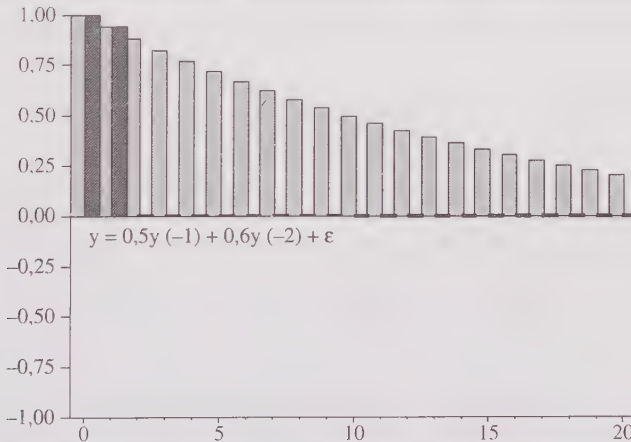


Figure 15.1 Corrélogramme d'une série $AR(2)$

Sur les figures 15.1 et 15.2 sont représentées les fonctions d'autocorrélation et d'autocorrélation partielles des deux séries. De manière alternée apparaissent les valeurs r_k et r_{kk} .

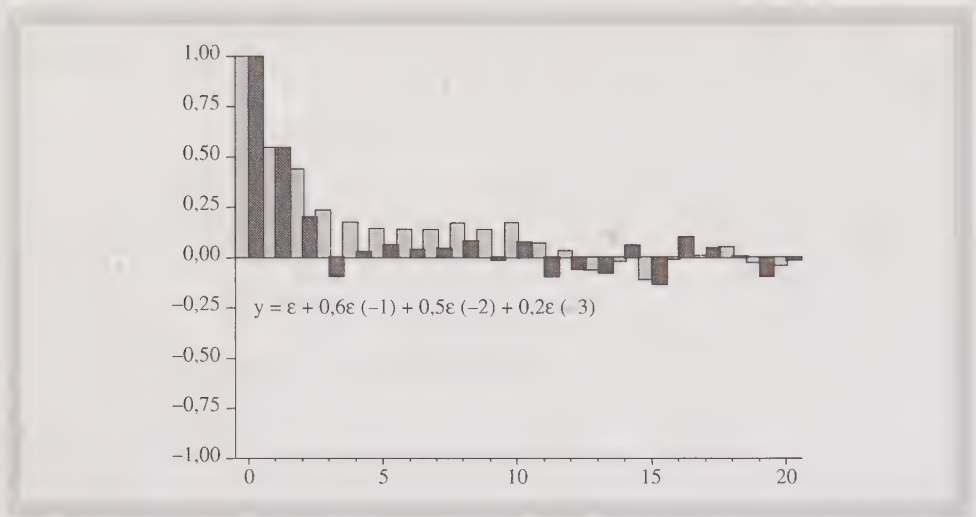


Figure 15.2 Corrélogramme d'une série MA(3)

La modélisation des séries temporelles s'appuie également sur des tests statistiques. Les tests de Box-Pierce et Ljung-Box permettent de tester l'ordre d'autocorrélation de la série à l'aide des statistiques suivantes

$$\begin{aligned} \text{Box-Pierce} &= T \sum_{k=1}^m r_k^2 \\ \text{Ljung-Box} &= T(T+2) \sum_{k=1}^m (T-k)^{-1} r_k^2 \end{aligned} \quad (15.9)$$

avec T le nombre d'observations.

Dans le cas d'un $ARMA(p, q)$ la statistique de Box-Pierce est distribuée asymptotiquement selon un loi du $\chi^2(m-p-q)$. Si la valeur calculée excède la valeur critique du $\chi^2(m-p-q)$, pour un seuil de significativité donné, on rejette l'hypothèse nulle de non-autocorrélation, rejeter l'hypothèse nulle conduisant à accepter l'hypothèse alternative selon laquelle au moins une autocorrélation est non nulle. Sur de petits échantillons la statistique de Ljung-Box donne une meilleure approximation d'une distribution du χ^2 .

On peut également utiliser des critères d'information tels que le critère d'information de Akaike (AIC) et le critère de Schwarz (BIC). Ils permettent d'identifier les valeurs de p et de q dans un modèle $ARMA$. Les décalages p et q sont choisis de manière à minimiser les statistiques suivantes

$$\begin{aligned} AIC(p, q) &= \ln \hat{\sigma}^2 + 2(p+q)T^{-1} \\ BIC(p, q) &= \ln \hat{\sigma}^2 + (p+q)T^{-1} \ln T \end{aligned} \quad (15.10)$$

avec $\hat{\sigma}^2$ l'estimation de la variance de l'erreur d'un $ARMA(p, q)$.

Cette forme de modélisation d'une série temporelle est beaucoup utilisée pour réaliser des prévisions à court terme mais, également dans l'analyse de la stationnarité des séries des modèles économétriques.

2. PROCESSUS NON STATIONNAIRE ET RACINE UNITAIRE

2.1 Les séries non stationnaires

Les séries temporelles non stationnaires peuvent être stationnarisées en calculant leur différence (séries DS) ou en calculant leur écart par rapport au trend (séries TS).

2.1.1 Séries DS

Une série caractérisée par un processus $AR(1)$

$$y_t = \mu + \gamma_1 y_{t-1} + \varepsilon_t \quad (15.11)$$

est stationnaire si $|\gamma_1| < 1$. Par contre, si $\gamma_1 = 1$ la série suit une marche aléatoire avec une dérive μ

$$y_t = \mu + y_{t-1} + \varepsilon_t \quad (15.12)$$

et elle est non stationnaire. En effet, pour une valeur initiale y_0 donnée, l'équation (15.12) peut s'écrire

$$y_t = \mu t + y_0 + \sum_{i=1}^t \varepsilon_i \quad (15.13)$$

or, si $E(\varepsilon_t) = 0$, $var(\varepsilon_t) = \sigma_\varepsilon^2$ et $cov(\varepsilon_t, \varepsilon_s) = 0$ quelque soit t et $s \neq t$ alors

$$E(y_t) = \mu t + y_0 \quad (15.14)$$

et

$$var(y_t) = E(y_t - E(y_t))^2 = E\left(\sum_{i=1}^t \varepsilon_i\right)^2$$

$$var(y_t) = t\sigma_\varepsilon^2 \quad (15.15)$$

L'espérance et la variance de y_t étant des fonctions de t , la série y_t est non stationnaire². Elle a une racine unitaire ($\gamma_1 = 1$) c'est-à-dire qu'elle est intégrée

² Si $\mu = 0$ la série est également non stationnaire car sa variance croît avec t .

d'ordre 1. Par contre la différence première $z_t = y_t - y_{t-1}$ est stationnaire car ε_t est stationnaire

$$z_t = (1 - L)y_t = \mu + \varepsilon_t \quad (15.16)$$

De manière plus générale on définit une série intégrée d'ordre d lorsque la série z_t définie par

$$z_t = (1 - L)^d y_t = \Delta^d y_t \quad (15.17)$$

est stationnaire c'est-à-dire lorsque la série y_t est devenue stationnaire après avoir été différenciée d fois, ce sont des séries DS³. Beaucoup de séries économiques temporelles sont intégrées d'ordre 1 et rarement d'ordre 2 ou 3.

2.1.2 Séries TS

On rencontre également beaucoup de séries qui deviennent stationnaires lorsqu'on calcule leur écart par rapport au trend. Ce sont des séries TS qui caractérisent une non-stationnarité de type déterministe

$$y_t = \mu + \beta t + \varepsilon_t \quad (15.18)$$

avec ε_t un bruit blanc. La série y_t est non stationnaire car $E(y_t)$ dépend du temps

$$E(y_t) = \mu + \beta t \quad (15.19)$$

On note que la variance de y_t ne dépend pas de t , elle est égale à la variance de l'aléa ε_t supposée constante.

Pour stationnariser le processus on peut estimer avec la méthode des MCO les paramètres μ et β . Le processus est alors stationnarisé en calculant

$$y_t - \hat{\mu} - \hat{\beta}t$$

L'écart d'une série par rapport à sa tendance est stationnaire lorsque la série est TS par contre il est non stationnaire si la série est DS. Une série DS doit être différenciée. Autrement dit une série TS est caractérisée par un trend déterministe tandis qu'une série DS est caractérisée par un trend stochastique. Les tests de racine unitaire permettent d'identifier les caractéristiques d'une série.

³ On note qu'un modèle $ARIMA(p, d, q)$ est un modèle $ARMA(p, q)$ dans lequel la série a été différenciée d fois afin d'être stationnarisée

$$\Delta^d y_t = \mu + \gamma_1 \Delta^d y_{t-1} + \dots + \gamma_p \Delta^d y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

2.2 Tests de racine unitaire

Granger et Newbold (1974) ont souligné que si dans un modèle économétrique on utilise des séries non stationnaires les résultats des tests statistiques sont faux, on obtient ce qu'on appelle communément une Spurious Régression. Les tests de Student et de Fisher tendent à ne pas rejeter l'hypothèse d'une corrélation entre la variable expliquée et les variables explicatives alors que la relation est inexistante. Les tests de racine unitaire cherchent à déterminer l'ordre d'intégration des séries. On présente ici les tests de Dickey Fuller et le test de Kwiatkowski, Phillips, Schmidt et Shin (KPSS).

2.2.1 Test de Dickey Fuller (test DF)

Le test de Dickey Fuller teste l'hypothèse H_0 qu'une série contient une racine unitaire, c'est-à-dire qu'elle est non stationnaire, contre l'hypothèse H_1 que la série est stationnaire. Pour réaliser ce test on estime, dans un premier temps, avec la méthode des MCO le modèle suivant

$$(1 - L)y_t = \Delta y_t = (\rho - 1)y_{t-1} + \varepsilon_t \quad (15.20)$$

avec ε_t un bruit blanc distribué selon une loi normale. Dans un second temps, on teste

$$\begin{aligned} H_0 : (\rho - 1) &= 0 && \text{la série est non stationnaire} \\ H_1 : (\rho - 1) &< 0 && \text{la série est stationnaire} \end{aligned}$$

Sous l'hypothèse H_0 la statistique $t = (\hat{\rho} - 1) / \hat{\sigma}_{\hat{\rho}}$ n'est pas distribuée selon une loi de Student comme dans le modèle de régression linéaire multiple. Dickey Fuller (1979) ont alors, à partir de simulations de Monte Carlo, présenté les valeurs critiques du test. Ils ont également fourni les tables statistiques permettant de tester la présence de racine unitaire dans les deux modèles suivants

$$\begin{aligned} \Delta y_t &= \mu + (\rho - 1)y_{t-1} + \varepsilon_t \\ \Delta y_t &= \mu + \beta t + (\rho - 1)y_{t-1} + \varepsilon_t \end{aligned} \quad (15.21)$$

2.2.2 Test de Dickey Fuller Augmenté (test ADF)

Le test précédent s'applique lorsque la série y_t est caractérisée par un processus $AR(1)$. Dickey Fuller ont généralisé ce test à une série caractérisée par un $AR(p)$ ⁴ afin d'obtenir un test plus précis : le test de Dickey Fuller Augmenté.

⁴ Pour effectuer le test ADF il faut au préalable déterminer le nombre de retards p , la méthode consiste à estimer par les MCO un modèle dans lequel sont rajoutés des termes retardés de la série de manière à blanchir les résidus. Le nombre de retards peut être déterminé avec les critères AIC ou BIC.

On estime le modèle

$$y_t = \mu + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \dots + \gamma_p y_{t-p} + \varepsilon_t$$

ou de manière équivalente

$$\Delta y_t = \mu + \gamma^* y_{t-1} + \gamma_1^* \Delta y_{t-1} + \gamma_2^* \Delta y_{t-2} + \dots + \gamma_{p-1}^* \Delta y_{t-(p-1)} + \varepsilon_t \quad (15.22)$$

avec $\gamma^* = (\gamma_1 + \gamma_2 + \dots + \gamma_p) - 1$ et $\gamma_i^* = \sum_{j=i+1}^p -\gamma_j$, $i = 1, \dots, p-1$.

Si l'hypothèse nulle $\gamma^* = 0$, avec comme hypothèse alternative $\gamma^* < 0$, n'est pas refusée alors y_t contient une racine unitaire. Pour tester cette hypothèse on calcule la statistique de Dickey fuller $(\widehat{\gamma^*} / \widehat{\sigma}_{\gamma^*})$.

Le test peut également être effectué lorsque le modèle contient un trend

$$\Delta y_t = \mu + \beta t + \gamma^* y_{t-1} + \gamma_1^* \Delta y_{t-1} + \gamma_2^* \Delta y_{t-2} + \dots + \gamma_{p-1}^* \Delta y_{t-(p-1)} + \varepsilon_t \quad (15.23)$$

Dans les deux tests présentés (DF et ADF), sous l'hypothèse nulle, la série y_t n'est pas stationnaire. Elle contient donc une racine unitaire et la série doit être différenciée afin d'être stationnarisée.

2.2.3 Stratégie du test ADF

On reprend ici la stratégie de test proposée par Harvey.

- (i) La première étape, consiste à estimer le modèle dans sa forme la plus générale

$$\Delta y_t = \mu + \beta t + \gamma^* y_{t-1} + \sum_{i=1}^{p-1} \gamma_i^* \Delta y_{t-i} + \varepsilon_t \quad (15.24)$$

et à utiliser la statistique t_τ pour tester l'hypothèse nulle $\gamma^* = 0$. Si l'hypothèse nulle d'une racine unitaire est rejetée la série est stationnaire.

- (ii) La seconde étape s'applique lorsque l'hypothèse nulle n'est pas refusée. Il est dans ce cas nécessaire de tester, avec la statistique t_β , si $\beta = 0$ et pour confirmer les résultats de tester avec la statistique Φ_3 les hypothèses $\gamma^* = \beta = 0$. Si le trend est significatif on refait un test de racine unitaire, c'est-à-dire on teste à nouveau l'hypothèse nulle $\gamma^* = 0$ en utilisant une loi normale. Si l'hypothèse nulle de racine unitaire est rejetée on conclut que la série est stationnaire. À l'inverse si l'hypothèse est acceptée on conclut que la série contient une racine unitaire.
- (iii) La troisième étape intervient lorsque, dans l'étape 2, le trend n'est pas significatif. On estime dans ce cas le modèle

$$\Delta y_t = \mu + \gamma^* y_{t-1} + \sum_{i=1}^{p-1} \gamma_i^* \Delta y_{t-i} + \varepsilon_t \quad (15.25)$$

et on teste avec la statistique t_μ la présence d'une racine unitaire ($H_0 : \gamma^* = 0$).

Si l'hypothèse nulle est rejetée on conclut que la série ne contient pas de racine unitaire.

Sinon, si on accepte l'hypothèse $\gamma^* = 0$, on teste l'hypothèse nulle $\mu = 0$ avec la statistique $t_{\alpha\mu}$ puis, pour confirmer les résultats, les hypothèses $\mu = \gamma^* = 0$ avec la statistique Φ_1 . Si la dérive est significative on procède à un test de racine unitaire avec une distribution normale. Dans ce dernier cas, si l'hypothèse d'une racine unitaire est rejetée on conclut que la série ne contient pas une racine unitaire et inversement si l'hypothèse nulle n'est pas rejetée on conclut que la série contient une racine unitaire.

- (iv) Si la dérive n'est pas significative on procède à la 4^{ème} étape qui consiste à estimer le modèle sans trend et sans dérive

$$\Delta y_t = \gamma^* y_{t-1} + \sum_{i=1}^{p-1} \gamma_i^* \Delta y_{t-i} + \varepsilon_t \quad (15.26)$$

On teste avec la statistique τ_ρ l'hypothèse nulle de racine unitaire ($H_0 : \gamma^* = 0$). Si l'hypothèse est rejetée la série contient une racine unitaire sinon elle est stationnaire.

2.2.4 Test de Kwiatkowski, Phillips, Schmidt et Shin (KPSS)

La faible puissance⁵ des tests ADF incite à la pratique d'autres tests, par exemple le test KPSS qui prend pour hypothèse nulle la stationnarité de la série.

Kwiatkowski, Phillips, Schmidt et Shin (1992) proposent d'estimer le modèle suivant

$$y_t = \mu_t + \beta t + \varepsilon_t \quad (15.27)$$

où ε_t est un processus stationnaire et μ_t suit une marche aléatoire

$$\mu_t = \mu_{t-1} + u_t, \quad u_t \sim \text{idd}(0, \sigma_u^2)$$

puis de tester

$$H_0 : \sigma_u^2 = 0 \text{ ou } \mu \text{ égale à une constante}$$

$$H_1 : \sigma_u^2 > 0$$

Sous H_0 la série y_t est trend stationnaire tandis que sous H_1 elle est non stationnaire.

Pour effectuer ce test on régresse la série y_t sur une constante et un trend afin de calculer la série des résidus e_t . On construit ensuite la statistique

$$S_t = \sum_{\tau=1}^t e_\tau \quad t = 1, \dots, T$$

⁵ La puissance d'un test est donnée par la probabilité qu'il conduit à rejeter correctement une hypothèse nulle fautive.

et on calcule la statistique LM

$$LM = \frac{1}{\widehat{S}^2} \sum_{t=1}^T S_t^2 \quad (15.28)$$

avec S^2 la variance de long terme⁶ des résidus e_t . Cette variance de long terme est définie par

$$S^2 = \lim_{T \rightarrow \infty} \frac{E(S_T^2)}{T}$$

elle prend en compte toutes les autocorrélations des résidus. L'estimation de la variance de long terme d'une série stationnaire de moyenne nulle peut être réalisée à partir du calcul suivant

$$\widehat{S}^2 = \frac{1}{T} \sum_{t=1}^T e_t^2 + \frac{2}{T} \sum_{i=1}^l w_{il} \sum_{t=i+1}^T e_t e_{t-i}$$

où w_{il} est une fonction de pondération optimale, le test $KPSS$ utilisant la fenêtre de Barlett

$$w_{il} = 1 - \frac{i}{l+1}$$

l correspondant à un nombre de retards qui doit être bien choisi : quand $T \rightarrow \infty$ on doit avoir $l \rightarrow \infty$.

On rejette l'hypothèse de stationnarité lorsque la statistique LM du test $KPSS$ est supérieure à sa valeur critique. Les valeurs critiques de ce test ont été obtenues par simulation comme dans les tests de Dickey-Fuller. Le test peut être également mis en œuvre pour tester l'hypothèse nulle de stationnarité au lieu de l'hypothèse nulle de trend stationnaire, dans ce cas, on omet la variable de trend du modèle (15.27).

Valeurs critiques du test $KPSS$			
Tests	0.10	0.05	0.01
avec trend	0.347	0.463	0.739
sans trend	0.119	0.146	0.216

⁶ La variance de court terme est donnée par

$$\frac{1}{T} \sum_{t=1}^T e_t^2$$

3. COINTÉGRATION ET MCE : L'APPROCHE DE ENGEL ET GRANGER

3.1 Tests de Cointégration

3.1.1 Cointégration : présentation

Une série qui contient d racines unitaires doit être différenciée d fois pour devenir stationnaire, elle est intégrée d'ordre d ($I(d)$). Soient deux séries temporelles y_t et x_t intégrées d'ordre d , toute combinaison linéaire des deux séries sera aussi $I(d)$, par exemple les aléas d'un modèle de régression $\varepsilon_t = y_t - \beta x_t$ sont $I(d)$.

Cependant s'il existe un vecteur β tel que l'ordre d'intégration des aléas est inférieur à d et égal à $d - b$ avec $b > 0$ alors, selon la définition de Engle et Granger, les séries y_t et x_t sont cointégrées d'ordre (d, b) . Par exemple, si la série ε_t est $I(0)$ et les séries y_t et x_t sont $I(1)$ alors les deux séries (y_t et x_t) sont cointégrées d'ordre $(1, 1)$ ($CI(1, 1)$), et le vecteur $(1, -\beta)$ est le vecteur de cointégration.

La notion de cointégration permet de mettre en évidence des relations de long terme stables entre des séries stationnaires. Ce concept reproduit l'existence d'un équilibre de long terme et l'aléa ε_t peut s'interpréter comme une distance à la période t par rapport à cet équilibre. Dans ce cas différencier les séries est contre productif car on cache la relation de long terme entre elles. Ainsi si les séries y_t et x_t sont cointégrées la relation de long terme $y_t = \beta x_t + \varepsilon_t$ peut être estimée avec la méthode des MCO, la régression obtenue ne sera pas Spurious. La relation de cointégration est une relation d'équilibre entre des séries en régime de croissance équilibrée mais des chocs peuvent affecter cette relation à court terme c'est-à-dire avoir des effets temporaires. Le problème consiste donc à déterminer si les séries d'un modèle sont cointégrées puis à estimer la relation de long terme et de court terme entre les variables.

3.1.2 La procédure de Engel et Granger

Soient deux séries $y(t)$ et $x(t)$ $I(1)$, le test de Engel et Granger suggère de procéder aux deux étapes suivantes

1. estimer avec la méthode des MCO la relation de long terme⁷

$$y_t = \beta x_t + \varepsilon_t \quad (15.29)$$

On note e_t les résidus de l'estimation.

⁷ La relation de long terme peut également contenir des termes dynamiques

$$A(L)y_t = B(L)x_t + \varepsilon_t$$

avec

$$\begin{aligned} A(L) &= 1 - \alpha_1 L - \dots - \alpha_p L^p \\ B(L) &= \gamma_0 + \gamma_1 L + \dots + \gamma_p L^p \end{aligned}$$

2. tester à l'aide des tests de Dickey Fuller

$$H_0 : e_t \text{ est } (1)$$

$$H_1 : e_t \text{ est } (0)$$

sur le modèle

$$\Delta e_t = \Psi^* e_{t-1} + \sum_{i=1}^{p-1} \Psi_i^* \Delta e_{t-i-1} + \mu + \omega_t \quad (15.30)$$

où ω_t est un bruit blanc. Sous H_0 les séries ne sont pas cointégrées, la série e_t n'est pas stationnaire.

On note que les valeurs critiques du test ne sont pas identiques aux tables de Dickey-Fuller de la section précédente car la série des résidus est obtenue par une estimation MCO. La distribution de la statistique $\hat{\Psi}^*/\hat{\sigma}_{\hat{\Psi}^*}$ dépend du nombre de variables explicatives dans la relation de long terme. Les tables adaptées sont celles de Engle et Yoo (1997)⁸, Phillips et Ouliaris (1990)⁹.

Si on conclut à la stationnarité de la série des résidus alors les séries sont cointégrées et la relation de long terme entre les deux variables peut-être estimée avec la méthode des MCO. On obtient un estimateur super convergent. Selon cette propriété, si les séries y_t , x_t et ε_t sont respectivement $I(1)$, $I(1)$ et $I(0)$ alors, à mesure que la taille de l'échantillon s'accroît l'estimateur des MCO de β converge vers sa vraie valeur à un taux plus rapide¹⁰ que l'estimateur des MCO calculé avec les variables stationnaires Δy_t et Δx_t . Asymptotiquement les variables $I(1)$ dominent les variables $I(0)$. Tout biais dû à l'endogénéité des variables est capturé par le résidu et conduit à un problème d'efficacité de l'estimateur des MCO, en particulier les aléas de la relation (15.29) sont autocorrélés. Cependant, dans des échantillons de taille finie l'estimateur des MCO de β dans l'équation (15.29) est biaisé. De plus, Phillips et Durlauf (1986) ont montré que la statistique de student correspondant au test standard de significativité de β ($H_0 : \beta = 0$) n'est pas valide, la distribution asymptotique de l'estimateur des MCO de β étant très complexe et non-normale. On ne peut pas, par conséquent, tester la significativité du paramètre β dans la relation de long terme (15.29). Ainsi, avec la méthode de Engle et Granger, on part de l'hypothèse que la relation de long terme entre les séries est décrite par (15.29) mais cette hypothèse généralement suggérée par la théorie économique ne peut pas être testée avec cette procédure. Par contre, les résidus donnent une estimation du déséquilibre et pour obtenir des informations sur la vitesse d'ajustement à l'équilibre on peut estimer un Modèle à Correction d'Erreur.

⁸ Engle, R.F., Yoo B.S., 1987, « Forecasting and Testing in cointegrated systems », *Journal of Econometrics*, 35, 143-159.

⁹ Phillips PCB, Ouliaris S., 1990, « Asymptotic Properties of residual based tests for cointegration », *Econometrica*, 58, 165-193.

¹⁰ Le vecteur de cointégration estimé tend vers sa vraie valeur à un taux de convergence égal à T le nombre d'observations, au lieu de la vitesse habituelle de $T^{1/2}$.

3.2 Modèles à Correction d'Erreur

Afin de reproduire la dynamique d'ajustement vers l'équilibre de long terme on peut utiliser un MCE. Supposons que les séries intégrées d'ordre 1 y_t et x_t sont cointégrées, le vecteur de cointégration étant donné par $(1, -\beta)$ le MCE décrivant la relation entre les deux séries s'écrit

$$\Delta y_t = \gamma \Delta x_t - \lambda (y_{t-1} - \hat{\beta} x_{t-1}) + u_t \quad (15.31)$$

Ce modèle décrit la variation de y_t autour de sa tendance de long terme. $\hat{\beta}$ est l'estimateur des MCO de la relation de long terme et le modèle est cohérent seulement si les variables sont cointégrées car dans ce cas les variables Δy_t , Δx_t et $(y_{t-1} - \hat{\beta} x_{t-1})$ sont $I(0)$ c'est-à-dire stationnaires. À l'équilibre, le terme $y - \hat{\beta} x$ est nul. Pendant les périodes de déséquilibre, il mesure la distance à laquelle le système s'est éloigné de l'équilibre à la période t . L'estimation avec la méthode des MCO du paramètre λ ($\lambda > 0$) donne alors une indication sur la vitesse d'ajustement de y_t vers son niveau d'équilibre, c'est-à-dire comment la variable y_t varie lorsqu'il y a un déséquilibre.

Ce modèle peut être généralisé pour capturer un processus dynamique plus complexe. En général on peut reformuler un MCE de la manière suivante

$$A(L)\Delta y_t = B(L)\Delta x_t - \lambda (y_{t-1} - \hat{\beta} x_{t-1}) + u_t \quad (15.32)$$

avec

$$\begin{aligned} A(L) &= 1 - \alpha_1 L - \dots - \alpha_p L^p \\ B(L) &= \gamma_0 + \gamma_1 L + \dots + \gamma_p L^p \\ \lambda &= 1 - \alpha_1 - \dots - \alpha_p \end{aligned}$$

Enfin, seul le cas de deux variables a été présenté et on a supposé la variable x_t exogène. La procédure de Engel et Granger s'applique également lorsque x_t est endogène. Dans ce cas on estime avec la méthode des MCO la relation de long terme ($y_t = \beta x_t + \varepsilon_t$) afin de tester si les séries sont cointégrées. Si les variables sont cointégrées on estime le VECM (modèle vectoriel à correction d'erreurs) suivant en appliquant les MCO équation par équation

$$\begin{aligned} A_1(L)\Delta y_t &= B_1(L)\Delta x_t - \lambda_1 (y_{t-1} - \hat{\beta} x_{t-1}) + u_{1t} \\ A_2(L)\Delta x_t &= B_2(L)\Delta y_t - \lambda_2 (y_{t-1} - \hat{\beta} x_{t-1}) + u_{2t} \end{aligned}$$

Pour terminer, lorsque le nombre de séries est supérieur à 2 il est possible d'obtenir des séries, intégrées d'ordre différents, cointégrées. Par exemple, soient la série y_t $I(1)$ et les séries x_t et z_t $I(2)$. Si les séries x_t et z_t sont cointégrées avec un vecteur de cointégration $(1, -\theta)$, c'est-à-dire si $x_t - \theta z_t$ est $I(1)$ alors, il peut exister un vecteur de cointégration tel que $y_t - \delta (x_t - \theta z_t)$ est $I(0)$. De manière générale,

soit N le nombre de séries, il peut exister $N - 1$ vecteurs de cointégration linéairement indépendants et donc $N - 1$ relations qui gouvernent l'évolution jointe des variables. Lorsque le nombre de variables devient supérieur à 2 l'approche de Engel et Granger devient insuffisante car elle ne considère qu'une relation de cointégration. On utilise dans ce cas l'approche VAR développée par Johansen pour réaliser les tests de cointégration et construire un VECM à plusieurs équations.

4. COINTÉGRATION ET MCE : L'APPROCHE DE JOHANSEN

4.1 Modèles à Correction d'Erreur à plusieurs équations

On considère un vecteur Z_t composé de N variables potentiellement endogènes. Le modèle VAR (Vecteur Autorégressif) non contraint, incluant k décalages correspondant au vecteur Z_t , s'écrit de la manière suivante

$$\begin{array}{ccccccc} Z_t & = & A_1 & Z_{t-1} & + \dots + & A_k & Z_{t-k} & + & \varepsilon_t & & \varepsilon_t \sim N(0, \Sigma) \\ (N,1) & & (N,N) & (N,1) & & (N,N) & (N,1) & & (N,1) & & \end{array} \quad (15.33)$$

Les modèles VAR permettent d'estimer des relations dynamiques entre des variables endogènes sans imposer a priori des restrictions sur les paramètres. Le système est sous une forme réduite : chaque variable contenue dans le vecteur Z_t est régressé en fonction des variables décalées composant le vecteur Z_t . Il peut être estimé avec la méthode des MCO car les variables explicatives sont prédéterminées.

Exemple d'un modèle VAR composé de 3 variables : z_{1t} , z_{2t} et z_{3t} avec $k = 2$

$$\begin{aligned} z_{1t} &= a_{11}z_{1t-1} + a_{12}z_{2t-1} + a_{13}z_{3t-1} + a_{14}z_{1t-2} + a_{15}z_{2t-2} + a_{16}z_{3t-2} + \varepsilon_{1t} \\ z_{2t} &= a_{21}z_{1t-1} + a_{22}z_{2t-1} + a_{23}z_{3t-1} + a_{24}z_{1t-2} + a_{25}z_{2t-2} + a_{26}z_{3t-2} + \varepsilon_{2t} \\ z_{3t} &= a_{31}z_{1t-1} + a_{32}z_{2t-1} + a_{33}z_{3t-1} + a_{34}z_{1t-2} + a_{35}z_{2t-2} + a_{36}z_{3t-2} + \varepsilon_{3t} \end{aligned}$$

$$Z_t = \begin{pmatrix} z_{1t} \\ z_{2t} \\ z_{3t} \end{pmatrix}, A_1 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, A_2 = \begin{pmatrix} a_{14} & a_{15} & a_{16} \\ a_{24} & a_{25} & a_{26} \\ a_{34} & a_{35} & a_{36} \end{pmatrix},$$

$$\varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{pmatrix}$$

L'équation (15.33) peut être reformulée en un VECM ¹¹ :

$$\Delta Z_t = \Gamma_1 \Delta Z_{t-1} + \dots + \Gamma_{k-1} \Delta Z_{t-(k-1)} + \Pi Z_{t-1} + \varepsilon_t \quad (15.34)$$

¹¹ On peut étendre le VECM de manière à inclure une constante et un trend dans le modèle.

avec $\Gamma_i = \sum_{j=i+1}^k -A_j$, $i = 1, \dots, k-1$ et $\Pi = A_1 + \dots + A_k - I$.

Ce système donne des informations sur les ajustements de court terme et de long terme des variables du modèle, contenues respectivement dans les matrices Γ_i et Π .

En ce qui concerne la relation de long terme, on factorise la matrice Π en $\Pi = \alpha\beta'$ avec α la vitesse d'ajustement et β une matrice de coefficients de long terme telle que $\beta'Z_{t-1}$ donne au plus $N - 1$ relations de cointégration assurant la convergence à long terme de Z_t vers son équilibre. Si Z_t est un vecteur composé de variables $I(1)$ alors les variables ΔZ_{t-i} sont $I(0)$ et ε_t est un bruit blanc et donc $I(0)$ si ΠZ_{t-1} est également $I(0)$. Le problème consiste donc à tester si Π est de rang r avec $0 < r \leq N - 1$, c'est-à-dire s'il existe r vecteurs de cointégration¹².

Pour simplifier on suppose que Z_t est un vecteur composé de variables $I(1)$. La méthode de Johansen consiste à estimer, dans un premier temps, avec la méthode des MCO les deux relations suivantes

$$\begin{aligned} \Delta Z_t &= P_1 \Delta Z_{t-1} + \dots + P_{k-1} \Delta Z_{t-(k-1)} + U_{0t} \\ Z_{t-k} &= T_1 \Delta Z_{t-1} + \dots + T_{k-1} \Delta Z_{t-(k-1)} + U_{kt} \end{aligned}$$

avec P_i et T_i les matrices de coefficients, U_0 et U_k les aléas. On note R_0 et R_k les vecteurs de résidus MCO

$$\begin{aligned} \Delta Z_t - \widehat{P}_1 \Delta Z_{t-1} + \dots + \widehat{P}_{k-1} \Delta Z_{t-(k-1)} &= R_{0t} \\ Z_{t-k} - \widehat{T}_1 \Delta Z_{t-1} + \dots + \widehat{T}_{k-1} \Delta Z_{t-(k-1)} &= R_{kt} \end{aligned}$$

Dans un second temps, l'estimateur du maximum de vraisemblance de β est donné par les vecteurs propres associés aux r plus grandes valeurs propres de

$$|\lambda S_{kk} - S_{k0} S_{00}^{-1} S_{0k}| = 0 \tag{15.35}$$

avec

$$S_{kk} = T^{-1} \sum_{t=1}^T R_{kt} R'_{kt} \text{ et } S_{0k} = T^{-1} \sum_{t=1}^T R_{0t} R'_{kt}$$

Soient $\widehat{\lambda}_1 > \widehat{\lambda}_2 > \dots > \widehat{\lambda}_N$ N valeurs propres et $\widehat{V} = (\widehat{v}_1, \widehat{v}_2, \dots, \widehat{v}_N)$ les vecteurs propres correspondants, les r premiers éléments de \widehat{V} déterminent des combinaisons linéaires de relations stationnaires. La matrice des vecteurs de cointégration $\widehat{\beta}$ est donnée par

$$\widehat{\beta} = (\widehat{v}_1, \widehat{v}_2, \dots, \widehat{v}_r)$$

¹² Si Π est de plein rang alors les variables incluent dans le vecteur Z_t sont $I(0)$ dans ce cas il n'y a pas de problème de Spurious Régression et on peut estimer le système VAR en niveau. A l'opposé si le rang de Π est égal à 0 il n'existe pas de relation de cointégration autrement dit il n'existe pas de combinaison linéaire des variables Z_{it} $I(0)$. On estime dans ce cas le modèle VAR en différence première éliminant ainsi la relation de long terme.

On obtient ainsi r relations de cointégration données par $\hat{v}'_i Z$ ($i = 1, \dots, r$). Ces relations sont des combinaisons linéaires des variables $I(1)$ composant le vecteur Z_t et doivent être $I(0)$ pour être fortement corrélées avec les éléments stationnaires de ΔZ_t présents dans l'équation (15.34).

L'amplitude des valeurs propres donne une indication sur le degré de cette corrélation et sert de base pour les tests de cointégration.

4.2 Tests de cointégration : test de la trace et test de la valeur propre maximale

Dans la relation (15.34) le vecteur ΠZ_{t-1} contient les relations de cointégration, il doit être stationnaire afin que ε_t soit $I(0)$. $\Pi (= \alpha\beta')$ a dans ce cas un rang réduit, il existe $r (\leq N - 1)$ colonnes linéairement indépendantes dans Π , c'est-à-dire que r vecteurs de cointégration sont présents dans β . L'estimation de β selon la procédure de Johansen permet de calculer les différents vecteurs de cointégration $\hat{v}'_i Z$ ($i = 1, \dots, r$). Chaque \hat{v}_i est associé à une valeur propre $\hat{\lambda}_i$ qui indique le degré de corrélation entre $\hat{v}'_i Z$ et la partie stationnaire du modèle (15.34). Les $N - r$ vecteurs $\hat{v}'_i Z$ ($i = r + 1, \dots, N$) restant donnent les combinaisons non stationnaires qui théoriquement sont non corrélées aux variables stationnaires ΔZ_t de l'équation (15.34). Par conséquent, pour les vecteurs propres correspondant à la partie non stationnaire du modèle on a $\hat{\lambda}_i = 0$, $i = r + 1, \dots, N$.

Pour tester l'hypothèse nulle qu'il y a au plus r vecteurs de cointégration et donc $N - r$ racines unitaires on teste l'hypothèse suivante

$$H_0 : \lambda_i = 0, \quad i = r + 1, \dots, N.$$

sous l'hypothèse H_0 seules les r premières valeurs propres sont non nulles.

4.2.1 Test de la Trace

On calcule la statistique de la Trace comme la statistique du ratio de vraisemblance. Soit $\text{Log}L^*$ la valeur du logarithme de la fonction de vraisemblance sous H_0 et $\text{Log}L$ la valeur du logarithme de la fonction de vraisemblance du modèle sans les contraintes

$$\lambda_{\text{trace}} = -2(\text{Log}L^* - \text{Log}L) = -T \sum_{i=r+1}^N \log(1 - \hat{\lambda}_i)$$

$$r = 0, 1, \dots, N - 2, N - 1$$

$$(15.36)$$

Le test de la trace teste l'hypothèse nulle que $r = q$ ($q = 1, \dots, N - 1$) contre l'alternative $r = N$.

4.2.2 Test de la valeur propre maximale

Un autre test permettant de tester les r plus grandes valeurs propres est donné par la statistique λ_{\max}

$$\lambda_{\max} = -T \text{Log} \left(1 - \hat{\lambda}_{r+1} \right)$$

$$r = 0, 1, \dots, N - 2, N - 1$$
(15.37)

avec cette statistique on teste l'hypothèse nulle $r = 0$ contre l'hypothèse $r = 1$, puis $r = 1$ contre $r = 2, \dots$

Les lois des deux tests ont été tabulées dans Johansen(1988)¹³. Lorsque les deux tests donnent des conclusions différentes, on retient en général les résultats du test de la Trace dont la puissance est plus élevée que celle du test de la valeur propre maximale.

4.3 Test d'exogénéité faible

Toutes les variables composant le vecteur Z_t ont été considérées endogènes. Afin de vérifier cette hypothèse on peut appliquer un test d'exogénéité faible.

On considère le VECM suivant dans lequel il existe deux relations de cointégration et $k = 2$

$$\begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \Delta y_{3t} \end{pmatrix} = \Gamma_1 \begin{pmatrix} \Delta y_{1t-1} \\ \Delta y_{2t-1} \\ \Delta y_{3t-1} \end{pmatrix} +$$

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{pmatrix} \begin{pmatrix} \beta_{11} & \beta_{21} & \beta_{31} \\ \beta_{12} & \beta_{22} & \beta_{32} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \end{pmatrix} + \varepsilon_t$$
(15.38)

$\alpha_{11}, \alpha_{21}, \alpha_{31}$ donnent respectivement la vitesse à laquelle $\Delta y_{1t}, \Delta y_{2t}, \Delta y_{3t}$ s'ajustent lorsque la relation de cointégration est donnée par $\beta_{11}y_1 + \beta_{21}y_2 + \beta_{31}y_3$. De même $\alpha_{12}, \alpha_{22}, \alpha_{32}$ donnent respectivement la vitesse à laquelle $\Delta y_{1t}, \Delta y_{2t}, \Delta y_{3t}$ s'ajustent lorsque la relation de cointégration est donnée par $\beta_{12}y_1 + \beta_{22}y_2 + \beta_{32}y_3$.

Si $\alpha_{31} = \alpha_{32} = 0$ alors la relation de cointégration n'intervient pas dans la troisième équation du VECM et on peut, dans ce cas, considérer que la variable y_{3t} est faiblement exogène et estimer une version partielle du VECM initial

¹³ Johansen S., 1988, « Statistical Analysis of Cointegration Vectors », *Journal of Economic Dynamics and Control*, 12, 231-254.

$$\begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} = \Gamma_0 \Delta y_{3t} + \tilde{\Gamma}_1 \begin{pmatrix} \Delta y_{1t-1} \\ \Delta y_{2t-1} \\ \Delta y_{3t-1} \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \beta_{11} & \beta_{21} & \beta_{31} \\ \beta_{12} & \beta_{22} & \beta_{32} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \end{pmatrix} + \varepsilon_t \quad (15.39)$$

Pour tester l'exogénéité faible il faut tester l'hypothèse $H_0 : \alpha_{ij} = 0$ pour $j = 1, \dots, r$, c'est-à-dire il faut tester si la ligne i contient des éléments nuls. Ce test est réalisé en imposant des contraintes sur les lignes de α afin d'appliquer le test du ratio de vraisemblance. La statistique LR est dans ce cas donnée par

$$LR = -2(\text{Log}L^* - \text{Log}L) = T \sum_{i=1}^r \log \frac{1 - \hat{\lambda}_i^*}{1 - \hat{\lambda}_i} \quad (15.40)$$

avec $\hat{\lambda}_i^*$ les $r - 1$ valeurs propres obtenues lorsque le modèle est contraint. Cette statistique est distribuée selon un $\chi^2(r(N - m))$ avec $(N - m)$ le nombre de lignes avec des contraintes.

4.4 Test d'hypothèse sur les relations de cointégration

Les tests de cointégration de Johansen indiquent le nombre de vecteurs de cointégration. Les estimations données par une colonne particulière de β ne sont pas unique nécessairement. Il est nécessaire d'imposer des restrictions motivées par des arguments économiques afin d'obtenir des informations sur la relation économique de long terme. Par exemple une contrainte du type $\beta_{1j} = -\beta_{2j}$ peut permettre d'identifier les colonnes de β . Le problème consiste alors à réaliser des tests de contraintes basés sur la statistique LR afin d'identifier β .

16

LA VALORISATION DES MARCHÉS ACTIONS

1. Introduction
2. Théorie financière et prix d'équilibre des actions
3. La construction d'un indicateur de valorisation
4. Éléments de corrigé

OBJECTIFS DU CHAPITRE

D'un point de vue méthodologique

- Pratique des tests de stationnarité : Dickey-Fuller augmenté
- Estimation d'un modèle à correction d'erreur univarié
- Méthode de Engel et Granger
- Tests de cointégration
- Calcul d'un délai moyen d'ajustement sur un modèle dynamique

D'un point de vue économique

- Présentation d'un modèle de prix d'équilibre pour les marchés actions
- Construction d'un indicateur de valorisation du marché boursier américain

NIVEAU

MASTER, deuxième année

CHAMPS ÉCONOMIQUE TRAITÉ

Macroéconomie financière

1. INTRODUCTION

Le paradigme de référence en finance est celui des marchés efficients (Fama (1970, 1991))¹. Il nous enseigne que les prix d'équilibre sur les marchés financiers intègrent l'ensemble de l'information pertinente disponible. Cela signifie en particulier que l'anticipation d'un événement futur entraîne des variations de prix d'actifs au moment où l'anticipation est formulée, et non pas à l'instant où se réalise cet événement. Il est donc impossible de prévoir l'évolution future des prix d'actifs puisqu'ils dépendent par hypothèse d'événements totalement inconnus au moment de la prévision. Dans ce contexte, les prix d'actifs sont supposés suivre une marche aléatoire, c'est-à-dire un processus stochastique pour lequel l'espérance mathématique du prix futur, c'est-à-dire sa meilleure prévision, est simplement le prix constaté à la période courante.

Les tests empiriques montrent en général que c'est le marché des changes qui correspond le plus aux critères de l'efficacité. Les travaux de Meese et Rogoff (1983)² sont à ce titre restés célèbres dans la littérature. Ils montrent en effet que pour des horizons de prévision courts, jusqu'à 6 mois, de purs modèles de séries temporelles, tels que des marches aléatoires, fournissent de meilleures prévisions que des modèles dits fondamentaux intégrant les principales variables explicatives proposées par la théorie économique des taux de change.

Si ces résultats empiriques n'ont pu être pleinement transposés aux cas des marchés obligataires et boursiers, ils ont néanmoins marqué de manière significative l'orientation générale des travaux de modélisation et de prévision sur les marchés financiers. Très souvent les recherches économétriques renoncent à la construction de modèles de prévision capables de reproduire de manière très précise la dynamique des prix d'actifs. Elles se concentrent plutôt sur l'identification de relations d'équilibre ou de long terme entre les prix d'actifs et leurs déterminants fondamentaux, de manière à pouvoir détecter des déséquilibres notoires ou des anomalies dans la formation de ces variables. On passe en quelque sorte d'une logique de prévision à une logique de valorisation des marchés. Du point de vue pratique qui est celui des gestionnaires de portefeuilles, la question centrale n'est plus alors de savoir si le marché va baisser ou monter mais s'il est éloigné de son équilibre fondamental, c'est-à-dire significativement sous-évalué ou surévalué. Dans ce cas, la nécessité d'un ajustement ou d'un retour à l'équilibre réhabilite l'exercice de la prévision et les deux logiques, de prévision et de valorisation, tendent à se confondre.

Cette réorientation des travaux d'économétrie financière au profit d'une logique de valorisation a trouvé avec l'économétrie des séries non stationnaires un cadre formel particulièrement adapté. En effet

- (i) Les tests de stationnarité (Dickey-Fuller, Phillips-Perron) des séries sont des tests de détection de racines unitaires caractéristiques des processus de marche aléatoire.

¹ Fama E., 1970, « Efficient capital markets : a review of theory and empirical work », *Journal of Finance*, 25. Fama E., 1991, « Efficient capital markets II », *Journal of Finance*, 46.

² Meese R., Rogoff K., 1983, « Empirical exchange rate models in the seventies : do they fit out of sample ? », *Journal of International Economics*, 141, 3-24.

- (ii) La théorie de la cointégration (Engle-Granger (1987)) fournit une interprétation statistique aux notions courantes en économie de relations d'équilibre ou de long-terme entre variables. Elle est donc adaptée à la définition de prix fondamental sur des actifs financiers.
- (iii) Enfin les modèles à correction d'erreur proposent avec la notion de force de rappel une représentation simple de l'articulation entre la dynamique de court terme des séries (portant sur les variations) et leur équilibre de long terme (portant sur les niveaux). Ils permettent ainsi d'appréhender les mécanismes de déviation temporaire entre le prix observé et le prix fondamental.

Ce chapitre propose de combiner la théorie financière de base et l'économétrie des séries non stationnaires pour construire un indicateur de valorisation des marchés boursiers. La section 2 propose de retrouver dans un contexte d'équilibre par arbitrage la formulation traditionnelle de Gordon-Shapiro donnant le prix fondamental des actifs boursiers. La section 3 présente les principes d'exploitation économétrique de ce modèle dans le but de construire l'indicateur de valorisation. Elle propose ensuite une application au cas des marchés américain et français. Les résultats d'estimation sont commentés à la section 4.

2. THÉORIE FINANCIÈRE ET PRIX D'ÉQUILIBRE DES ACTIONS

Le modèle de Gordon-Shapiro (1956)³ constitue une exploitation particulière de la définition de la valeur théorique d'un actif financier comme la somme actualisée des revenus futurs qui lui sont associés. Cette approche s'applique aussi bien aux obligations distribuant des revenus fixes (les coupons) qu'aux actions distribuant des revenus variables et aléatoires (les dividendes). Il convient toutefois dans le cas des actifs boursiers de majorer le taux d'actualisation des dividendes pour prendre en compte le caractère aléatoire de ces revenus. C'est ce que Gordon et Shapiro appellent le taux de rendement requis par les actionnaires.

2.1 Recherche d'un prix d'équilibre à partir d'une relation d'arbitrage

La théorie financière contemporaine⁴ permet de justifier ce type de relation d'équilibre par actualisation sur la base d'un comportement d'arbitrage des investisseurs entre placements risqués et placements non risqués. On parle alors d'équilibre par arbitrage.

³ Gordon M., Shapiro E., 1956, « Capital equipment analysis : the required rate of profit », *Management Science*, 31, 102-110.

⁴ Voir par exemple Cuthbertson K., 2000, *Économie financière quantitative*, De Boeck Université.

Le rendement d'une action entre deux dates t et $t + 1$ s'écrit

$$\rho_t = \frac{P_{t+1} - P_t + D_t}{P_t} \quad (16.1)$$

où P_t désigne le prix ou cours de l'action et D_t le dividende éventuellement versé à l'actionnaire au cours de la période. Ce rendement est aléatoire car P_{t+1} le prix de fin de période est par définition inconnu en début de période, tout comme le dividende.

On suppose que les investisseurs présents sur le marché vont procéder à des opérations d'arbitrage entre ce placement boursier risqué et un placement sans risque de taux de rendement r_t . Ces opérations s'effectuent tant que n'est pas réalisée la condition d'équilibre ou d'absence d'opportunité d'arbitrage (AOA) entre les deux stratégies de placement. Cette condition s'écrit

$$E_t(\rho_t) = r_t + \pi_t \quad (16.2)$$

où $E_t(\cdot)$ est l'opérateur d'espérance mathématique et π_t désigne la prime de risque ou surcroît de rémunération demandée par les investisseurs sur le placement risqué. En développant $E_t(\rho_t)$ la condition d'équilibre devient

$$P_t = \frac{E_t(P_{t+1}) + D_t}{1 + r_t + \pi_t} \quad (16.3)$$

Le prix d'équilibre de l'action dépend ainsi du prix futur anticipé. Il y a là une lecture possible des phénomènes de bulle financière dans lesquels la hausse du prix des titres s'explique plus par des anticipations de hausses futures que par des variations de leur valeur fondamentale (les dividendes).

En remplaçant P_{t+1} par sa valeur théorique au sens de (16.3), ce qui revient à faire une hypothèse d'anticipations rationnelles sur le prix futur, on obtient par itérations successives (en supposant que r_t et π_t sont anticipés comme constants pour le futur ($r_{t+i} = r_t$, $\pi_{t+i} = \pi_t$)).

$$P_t = \frac{D_t}{(1 + r_t + \pi_t)} + \frac{E_t(D_{t+1})}{(1 + r_t + \pi_t)^2} + \dots + \frac{E_t(P_{t+i})}{(1 + r_t + \pi_t)^i} \quad (16.4)$$

Soit pour une suite infinie d'itérations, ce qui revient implicitement à supposer un horizon de réflexion également infini pour les investisseurs

$$P_t = \sum_{i=0}^{\infty} \frac{E_t(D_{t+i})}{(1 + r_t + \pi_t)^{i+1}} + \lim_{i \rightarrow \infty} \frac{E_t(P_{t+i})}{(1 + r_t + \pi_t)^i} \quad (16.5)$$

En supposant que la condition dite de transversalité ou d'absence de solution de bulle

$$\lim_{i \rightarrow \infty} \frac{E_t(P_{t+i})}{(1 + r_t + \pi_t)^i} = 0$$

est vérifiée, on obtient la relation définissant le cours fondamental de l'action ou de l'indice boursier. Il apparaît simplement comme la valeur actualisée des dividendes futurs anticipés.

$$P_t = \sum_{i=0}^{\infty} \frac{E_t(D_{t+i})}{(1+r_t+\pi_t)^{i+1}}$$

Il faut noter toutefois la présence de la prime de risque dans le facteur d'actualisation $(1+r_t+\pi_t)^i$. Elle découle directement de la condition (16.2) d'absence d'opportunité d'arbitrage.

Si l'on suppose en outre comme Gordon et Shapiro que les dividendes croissent à un taux moyen constant g_t , l'écriture du prix fondamental devient

$$P_t = \sum_{i=0}^{\infty} \frac{D_t(1+g_t)^i}{(1+r_t+\pi_t)^{i+1}} \quad (16.6)$$

ce qui donne sous l'hypothèse $g_t < r_t + \pi_t$

$$P_t = \frac{D_t}{(r_t + \pi_t - g_t)} \quad (16.7)$$

Les dividendes constituent la fraction des bénéfices distribués aux actionnaires. En notant k le taux de distribution des bénéfices (E_t), le prix fondamental s'exprime de la manière suivante

$$P_t = \frac{kE_t}{(r_t + \pi_t - g_t)} \quad (16.8)$$

On pose en outre $k = 1$, hypothèse qui traduit le fait qu'à long terme tous les bénéfices sont versés aux actionnaires. On obtient sur cette base deux relations d'équilibre couramment utilisées par les praticiens pour juger de la valorisation des marchés actions

$$\frac{P_t}{E_t} = \frac{1}{(r_t + \pi_t - g_t)} \quad (16.9)$$

ou encore

$$\frac{E_t}{P_t} = r_t + \pi_t - g_t \quad (16.10)$$

La première relation (16.9) définit la valeur d'équilibre du Price Earning Ratio (noté PE) encore appelé le multiple de capitalisation de l'action ou de l'indice boursier. En négligeant le terme $\pi_t - g_t$, le PE apparaît donc comme étant homogène avec l'inverse des taux d'intérêt. Un multiple de capitalisation de 25 pour un indice boursier sera ainsi cohérent ou à l'équilibre vis-à-vis d'un taux d'intérêt de $1/25 = 4\%$.

La seconde relation (16.10) définit plus simplement la valeur d'équilibre de l'inverse du PE comme la somme du taux d'intérêt et de l'écart entre la prime de risque et le taux de croissance (de long terme) des bénéfices. $\frac{E_t}{P_t}$ est appelé le Earning Yield ou taux de rendement de l'indice boursier. C'est un taux de rendement attendu qui mesure le gain potentiel des actionnaires si la totalité des bénéfices courants⁵ des entreprises leur était distribuée. Il peut être qualifié de taux de rendement instantané si on l'oppose à $\frac{E_t}{P_t} + g_t$ qui le corrige de l'évolution de long terme des bénéfices.

2.2 Une première approche de la valorisation

L'utilisation de ces deux relations comme instrument de valorisation des marchés est délicate dans la mesure où elles mettent en jeu deux grandeurs, π_t et g_t , qui ne sont pas directement observables. On est de fait conduit à fixer arbitrairement un de ces paramètres pour calculer sur la base des données observées (PE ou EP , r) la valeur implicite de l'autre paramètre.

En fixant g_t à un niveau raisonnable, on obtient à partir de (16.10) la prime de risque implicite du marché à la date t

$$\pi_t = \left(\frac{E_t}{P_t} + g_t \right) - r_t \quad (16.11)$$

Elle apparaît ici comme l'excès de rendement anticipé à long terme du placement boursier par rapport au taux sans risque. Cette prime de risque peut être ensuite comparée à sa moyenne historique pour juger de la juste évaluation du marché. Comme le calcul des primes historiques doit être réalisé sur la base d'une valeur constante pour g_t , il est envisageable comme le font parfois les praticiens de raisonner sur une valeur nulle de g_t . Une prime de risque supérieure (respectivement inférieure) à sa moyenne pourra traduire alors un rendement trop élevé (respectivement faible) pour le marché boursier et donc une sous-évaluation (respectivement surévaluation).

Alternativement, en fixant π à un niveau raisonnable (5 % par exemple), on obtient l'anticipation implicite du marché à la date t sur la croissance de long terme des bénéfices

$$g_t = r_t + \pi_t - \frac{E_t}{P_t} \quad (16.12)$$

Cette valeur calculée pour g_t doit être comparée à une valeur raisonnable ou encore soutenable par les entreprises. Si l'analyse porte sur une action particulière ou un indice boursier sectoriel, cette croissance soutenable à long terme peut être définie par référence à la croissance passée des bénéfices ou de manière *ad hoc* selon un jugement d'expert sur les perspectives de l'entreprise ou du secteur industriel.

⁵ Très souvent les praticiens raisonnent à partir de Price Earning ou d'Earning Yield normalisés, c'est-à-dire incorporant non pas les bénéfices courants des entreprises mais des consensus de prévisions d'analystes financiers pour les 12 ou 24 mois à venir.

En revanche, si l'évaluation porte sur un indice boursier global, la croissance de long terme des bénéfices est indissociable du partage de la valeur ajoutée à l'échelle macroéconomique. La croissance des bénéfices peut être temporairement supérieure à la croissance économique en valeur en cas de déformation du partage de la valeur ajoutée en faveur des profits. À long terme toutefois, ces deux taux de croissance doivent être identiques. Il est donc possible de juger de la valorisation du marché en comparant la valeur calculée pour g_t à sa moyenne historique, qui sur longue période ne doit pas être différente de la croissance économique en valeur.

Comme précédemment avec la prime de risque, la validité de la démarche repose sur l'hypothèse de fixité de la croissance économique de long terme. Celle-ci peut être remise en cause par une élévation de la croissance potentielle de l'économie.

3. LA CONSTRUCTION D'UN INDICATEUR DE VALORISATION

3.1 Principes

Les deux indicateurs de valorisation précédents présentent finalement plusieurs limites qui peuvent être regroupées autour de trois arguments.

- (i) Ils apparaissent comme les variables inobservables d'une relation d'équilibre entre la valeur boursière des titres, les bénéfices des sociétés et les taux d'intérêt. Il faut ainsi faire une hypothèse sur l'une de ces variables pour lever le degré de liberté de la relation et définir l'indicateur de valorisation.
- (ii) Le jugement de surévaluation ou de sous-évaluation du marché selon la position de l'indicateur par rapport à sa moyenne historique repose sur l'hypothèse que cette moyenne historique est un bon estimateur de sa moyenne théorique, c'est-à-dire de son espérance mathématique. On fait donc une hypothèse de fixité de la moyenne de l'indicateur qui exclut une possible non stationnarité de la variable. Les tests statistiques montrent au contraire que les primes de risque calculées à partir de la relation d'équilibre sont non stationnaires et précisément intégrées d'ordre 1 ($I(1)$). Cette variable de prime de risque possède donc une racine unitaire et doit être différenciée pour être stationnarisée. Cela signifie en d'autres termes que la relation (16.10) ne peut être considérée comme une relation de cointégration, c'est-à-dire comme un équilibre ne pouvant être durablement déformé.
- (iii) On a supposé par souci de simplification que le taux d'intérêt sans risque était constant dans le temps. Les calculs développés à partir de (16.3) peuvent toutefois être généralisés au cas où le taux d'intérêt est au contraire variable. Il est préférable dans ce cas, du fait de la linéarité de

l'opérateur espérance, de supposer un taux d'intérêt variable mais déterministe. Les facteurs d'actualisation se présentent alors comme les produits du taux d'intérêt courant et des taux courts futurs⁶, c'est-à-dire finalement comme des taux d'intérêt à long terme. En d'autres termes la vraie relation d'équilibre porte sur le prix, les bénéfices courants et l'ensemble des taux de rendement constituant la structure par terme des taux d'intérêt. L'approche précédente n'est donc acceptable que dans le cas d'une structure des taux plate, c'est-à-dire dans le cas d'un taux d'intérêt unique quelque soit la maturité des titres.

On propose maintenant de construire un nouvel indicateur de valorisation, en restant dans le même cadre théorique mais en transformant la relation (16.10) en modèle économétrique. Cet indicateur devra s'affranchir des limites précédentes et répondre à plusieurs exigences garantissant *a priori* sa fiabilité. Sa construction est décomposable en trois étapes.

3.1.1 Identification d'une relation de cointégration entre Earning Yield et taux d'intérêt

Sur la base de la relation (16.10), on pose le modèle linéaire suivant où $(\frac{E}{P})_t$ est expliqué par une constante et deux taux d'intérêt de maturités différentes : un taux court r_t^c et un taux long r_t^l .

$$(E/P)_t = \alpha + \beta r_t^c + \gamma r_t^l + u_t \tag{16.13}$$

La différence avec le modèle théorique réside dans la présence des deux taux d'intérêt affectés des coefficients β et γ . La constante α peut être interprétée comme la moyenne de la variable $\pi_t - g_t$.

Les trois variables étant, selon la plupart des études empiriques, non stationnaires de type $I(1)$, l'estimation du modèle (16.13) par les MCO ne fournit des estimateurs convergents des paramètres que si ce modèle décrit une relation de cointégration. Le test de cointégration consiste à tester la stationnarité des résidus \hat{u}_t du modèle estimé par les MCO.

On pratique donc un test *ADF* sur ces résidus en estimant le modèle suivant par les MCO

$$\Delta \hat{u}_t = \gamma \hat{u}_{t-1} + \sum_{i=1}^k \gamma_i \Delta \hat{u}_{t-i} + \varepsilon_t \tag{16.14}$$

⁶ On obtient précisément pour le facteur d'actualisation de la période i (en négligeant la prime de risque) : $\prod_{i=0}^k (1 + r_{t+k}^*)$ où r_{t+k}^* désigne le taux de maturité une période anticipé pour la période $t + k$. La théorie des anticipations de la structure par terme des taux d'intérêt (voir par exemple Shiller (1990)) montre, de nouveau par une logique d'absence d'opportunité d'arbitrage, mais cette fois sur les marchés monétaire et obligataire, que ces produits de taux courts futurs anticipés ne sont rien d'autre que les taux d'intérêt à long terme. On a ainsi pour le taux de maturité k périodes la relation : $(1 + r_t^k)^k = \prod_{i=0}^k (1 + r_{t+i}^*)$,

ce qui donne en passant par les log : $r_t^k \simeq 1/k \sum_{i=0}^k r_{t+i}^*$.

où l'on teste

$$H_0 : \gamma = 0$$

$$H_1 : \gamma < 0$$

On souhaite clairement ici rejeter l'hypothèse nulle et conclure H_1 qui garantit la stationnarité de \widehat{u}_t . Celle-ci traduit l'existence d'une relation de cointégration entre le Earning Yield et les taux d'intérêt, c'est-à-dire l'existence d'un équilibre qui ne peut être durablement déformé.

3.1.2 Retour à l'équilibre

Mécanismes de retour à l'équilibre

Le théorie financière présentée dans ce chapitre nous incite à considérer les variables de taux d'intérêt présentes dans (16.13) comme exogènes et à pratiquer une lecture causale de la relation de cointégration en considérant qu'elle fournit bien une estimation du niveau de long terme (fondamental) du Earning Yield. Soit

$$\left(\widehat{E/P}\right)_t = \widehat{\alpha} + \widehat{\beta}r_t^c + \widehat{\gamma}r_t^l \quad (16.15)$$

La théorie économique des taux d'intérêt nous conforte dans cette position. Le taux d'intérêt à court terme en tant qu'instrument de politique monétaire dépend principalement de l'évolution de l'inflation par rapport à l'objectif d'inflation fixé par les autorités monétaires. Les taux d'intérêt à long terme résultent de l'équilibre du marché obligataire et s'interprètent selon la théorie des anticipations de la structure par terme des taux d'intérêt comme une moyenne pondérée du taux court et des taux courts futurs anticipés (Shiller (1990))⁷.

Cette hypothèse d'exogénéité des taux d'intérêt peut être testée en appliquant le théorème de la représentation de Granger. Si les taux d'intérêt sont exogènes, le modèle à correction d'erreur associé à la relation de cointégration concernera uniquement la dynamique de court terme de $(E/P)_t$. Une spécification simple du MCE est donnée par exemple par

$$\Delta(E/P)_t = \alpha_1 + \beta_1 \Delta r_t^c + \gamma_1 \Delta r_t^l - \lambda_1 \widehat{u}_{t-1} + \varepsilon_{1t} \quad (16.16)$$

En revanche, si les taux d'intérêt ne sont pas véritablement exogènes ils peuvent participer au mécanisme de correction d'erreur ou de retour à l'équilibre de long terme. Deux autres MCE pourront être également associés à la relation de cointégration. Soient par exemple

$$\Delta r_t^c = \alpha_2 + \beta_2 \Delta(E/P)_t + \gamma_2 \Delta r_t^l + \lambda_2 \widehat{u}_{t-1} + \varepsilon_{2t} \quad (16.17)$$

$$\Delta r_t^l = \alpha_3 + \beta_3 \Delta(E/P)_t + \gamma_3 \Delta r_t^c + \lambda_3 \widehat{u}_{t-1} + \varepsilon_{3t} \quad (16.18)$$

⁷ Shiller R., 1990, « The term structure of interest rates », *Handbook of Monetary Economics*, Volume 1, Friedman B.M. et Hahn F.M.

On doit donc estimer les trois MCE précédents et s'assurer que seul λ_1 est significatif. Dans ce cas les taux d'intérêt sont bien exogènes et le retour à l'équilibre de long terme se fait uniquement par des variations du Earning Yield, c'est-à-dire in fine par des variations de l'indice boursier.

Délai moyen de retour à l'équilibre

Pour que le diagnostic sur la valorisation puisse s'inscrire dans une logique de prévision, il faut que les délais de retour à l'équilibre de long terme ne soient pas trop élevés. Le modèle à correction d'erreur permet de calculer le délai moyen de retour à l'équilibre. Le MCE s'écrit aussi

$$(E/P)_t - (E/P)_{t-1} = \alpha_1 + \beta_1 \Delta r_t^c + \gamma_1 \Delta r_t^l - \lambda_1 \left[\left((E/P)_{t-1} - (\widehat{E/P})_{t-1} \right) \right] \quad (16.19)$$

soit encore en utilisant les polynômes de retard

$$(E/P)_t = \Phi_\alpha(L) + \Phi_\beta(L) \Delta r_t^c + \Phi_\gamma(L) \Delta r_t^l + \Phi_\lambda(L) \left(\widehat{E/P} \right)_t \quad (16.20)$$

avec

$$\Phi_\alpha(L) = \frac{\alpha_1}{[1 - (1 - \lambda_1)L]}$$

$$\Phi_\beta(L) = \frac{\beta_1}{[1 - (1 - \lambda_1)L]}$$

$$\Phi_\gamma(L) = \frac{\gamma_1}{[1 - (1 - \lambda_1)L]}$$

et

$$\Phi_\lambda(L) = \frac{\lambda_1 L}{[1 - (1 - \lambda_1)L]}$$

Le délai moyen de retour à l'équilibre est donné par

$$d = \frac{\Phi'_\lambda(1)}{\Phi_\lambda(1)}$$

ce qui donne $d = 1/\lambda_1$. Notons bien que, c'est parce que la dynamique du MCE est ici très simple, que le délai moyen de retour à l'équilibre est directement donné par l'inverse de la force de rappel.

3.1.3 Indicateur de valorisation, niveaux théoriques de l'indice et degré de surévaluation

L'indicateur de valorisation est ensuite défini comme le résidu de la relation de cointégration rapporté à son écart type. On pose

$$IV_t = \frac{\hat{u}_t}{\hat{\sigma}_{\hat{u}_t}} \quad (16.21)$$

$IV_t > 0$ (respectivement < 0) signifie que le Earning Yield du marché est supérieur (inférieur) à sa valeur d'équilibre donnée par le niveau des taux d'intérêt. La valeur de l'indice boursier est donc trop faible compte tenu des bénéfices courants et des taux d'intérêt. Le marché sera en conséquence jugé peu cher (cher) ou sous-évalué (surévalué).

La division par l'écart type estimé des résidus permet de juger de la significativité de l'écart à l'équilibre. Si les résidus \hat{u}_t sont normaux, IV_t est distribué selon une loi de Student et sa distribution asymptotique sera une loi $N(0,1)$. Dans ce cas, les intervalles de confiance à 65 % et 95 % pour IV_t sont $[-1,1]$ et $[-2,2]$. La significativité de la déformation de la relation de long terme pourra être jugée selon la position de l'indicateur par rapport à ces bornes.

Le modèle de long terme fournit la valeur théorique du Earning Yield dont on tire simplement la valeur théorique de l'indice boursier

$$\hat{P}_t = \frac{E_t}{\left(\hat{\alpha} + \hat{\beta}r_t^c + \hat{\gamma}r_t^l\right)} \quad (16.22)$$

En modulant la valeur théorique de $(E/P)_t$ d'un écart type, on obtient une borne inférieure et une borne supérieure pour \hat{P}_t . Soient

$$\hat{P}_t^{\text{inf}} = \frac{E_t}{\left(\hat{\alpha} + \hat{\beta}r_t^c + \hat{\gamma}r_t^l\right) + \hat{\sigma}_{\hat{u}_t}} \quad (16.23)$$

$$\hat{P}_t^{\text{sup}} = \frac{E_t}{\left(\hat{\alpha} + \hat{\beta}r_t^c + \hat{\gamma}r_t^l\right) - \hat{\sigma}_{\hat{u}_t}} \quad (16.24)$$

Enfin le degré de surévaluation (positif ou négatif) du marché est donné par l'écart relatif entre le prix observé et le prix fondamental. Soit

$$d_t = 100 \cdot \frac{P_t - \hat{P}_t}{\hat{P}_t} \quad (16.25)$$

3.2 Application sur les marchés boursiers américain et français

3.2.1 Données et profil des séries

La question centrale concerne ici le choix d'un indice boursier représentatif. Pour conférer à ces évaluations une réelle dimension macroéconomique, il est souhaitable de retenir des indices larges intégrant un grand nombre de sociétés cotées.

Nous retenons pour les deux pays les indices « Total Market » construits et diffusés par le serveur Datastream. Ils ne couvrent pas de manière exhaustive l'ensemble des sociétés cotées mais les 200 plus grosses capitalisations de la place. Les données sont rassemblées dans le fichier chapitre 15.xls. Elles concernent le Price Index, c'est-à-dire la valeur de l'indice boursier en points (PI)⁸ et le Price Earning ou multiple de capitalisation associé à l'indice (PE). Les séries de taux d'intérêt portent sur les taux à 3 mois ($T3M$) et à 10 ans ($T10Y$). Les taux courts sont les taux de rémunération sur les dépôts bancaires en eurodevises tandis que les taux longs désignent les taux de rendement actuariels des obligations d'état de référence (benchmark) à 10 ans. Les séries sont suffixées par le code US pour les données américaines et par FR pour les données françaises. Toutes les séries sont de fréquence mensuelle et construites à partir des valeurs de fin de mois des séries quotidiennes. Elles couvrent la période janvier 1980 - août 2000 à l'exception de la série de taux longs français qui n'est disponible qu'à compter de 1985.

La figure 16.1 ci-après présente pour le marché américain l'évolution depuis 1980 des variables d'intérêt du modèle de valorisation : les deux taux d'intérêt et l'inverse du Price Earning du marché, c'est-à-dire le Earning Yield. Ces trois variables présentent des évolutions de long terme semblables qui semblent bien traduire l'existence d'une relation d'équilibre ou de cointégration entre ces séries. La décline tendancielle du Earning Yield traduit de manière équivalente la progression du multiple de capitalisation, c'est-à-dire une progression de l'indice boursier plus rapide que celle des bénéfices des sociétés. Elle révèle donc un renchérissement continu du marché, qui pourrait être assimilé à une anomalie ou une surévaluation s'il n'était pas associé à une décline comparable des taux d'intérêt.

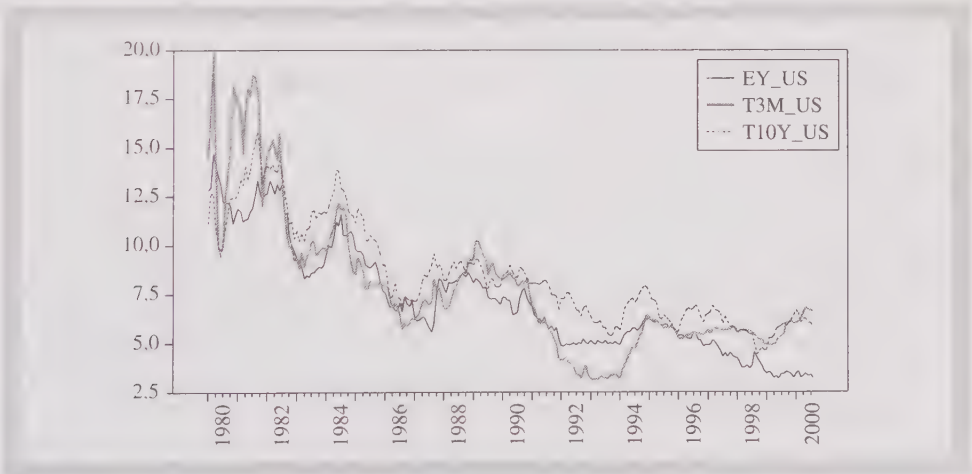


Figure 16.1 Earning Yield et taux d'intérêt — États-Unis

Rappelons enfin quelques phases marquantes de ces évolutions financières. Le début des années 1980 est marqué par la politique monétaire hyper restrictive de

⁸ Noté P dans le modèle théorique.

la réserve fédérale et de son président Paul Volcker. Il s'agissait alors de lutter contre les conséquences inflationnistes du deuxième choc pétrolier. Les hausses brutales de taux courts à près de 20 % ont entraîné une importante déconnexion entre les taux courts et les taux longs qui s'est traduite par une inversion temporaire de la structure par terme des taux d'intérêt (taux longs inférieurs aux taux courts).

Le krach d'octobre 1987 s'est produit dans un contexte de remontée des taux d'intérêt amorcée un an plutôt en 1986. Il s'est traduit par une remontée violente du rendement du marché (Earning Yield) laquelle apparaît finalement comme une correction légitime après la hausse des taux.

La période 1991-1993 est également significative du point de vue de l'évolution des taux d'intérêt. Elle fait suite à un mouvement de faillites bancaires dans le secteur des caisses d'épargne américaines. Les autorités monétaires ont alors rapidement et significativement baissé leur taux d'intervention pour faciliter l'accès des caisses d'épargne à une ressource bon marché. Ces baisses de taux courts ont manifestement été jugées comme temporaires par le marché obligataire qui n'a répercuté que partiellement la détente monétaire. C'est de nouveau une période de déconnexion importante des taux courts et des taux longs, mais cette fois dans un contexte différent de pentification de la courbe des taux. L'évolution du Earning Yield est à cette période intermédiaire entre celles des taux courts et des taux longs.

3.2.2 Tests, estimations et construction de l'indicateur

Problème 16.1

Tests de stationnarité

1. Réaliser les tests de Dickey-Fuller Augmenté sur les variables d'intérêt du modèle de valorisation. On considère donc les variables $T3M$, $T10Y$ et EY en niveau et en différence.

Ces tests seront effectués selon deux étapes. Dans un premier temps, on sélectionne le nombre de retards adéquats pour pratiquer les tests ADF en utilisant la procédure *lagselec.src* disponible sur le site web (extrait du site [http : www.estima.com](http://www.estima.com)). Cette procédure propose de déterminer l'ordre p du processus AR pouvant être associé à chaque variable. Les critères utilisés sont : les critères d'information usuels AIC et BIC , les tests d'autocorrélation de Ljung-Box et du multiplicateur de Lagrange et enfin un test de significativité du dernier retard (General to Simple). Concrètement, l'introduction de ces retards dans les tests a vocation à «blanchir» les résidus des régressions de manière à accroître la précision des estimateurs.

Dans une seconde étape, les tests ADF seront pratiqués selon la stratégie de test séquentiel où l'hypothèse nulle de racine unitaire est successivement testée dans les modèles (1), (2) et (3) suivants avec respectivement un trend et une constante, une constante mais pas de trend, et enfin ni trend ni constante (cf. chapitre 15, Économétrie des séries non stationnaires, p. 340). La procédure de test s'arrête dès qu'est rejetée l'hypothèse nulle de racine unitaire.

$$\text{Modèle 1 : } \Delta x_t = \mu + \beta t + \gamma x_{t-1} + \sum_{i=1}^{p-1} \gamma_i \Delta x_{t-i} + \varepsilon_t \quad (16.26)$$

$$\text{Modèle 2 : } \Delta x_t = \mu + \gamma x_{t-1} + \sum_{i=1}^{p-1} \gamma_i \Delta x_{t-i} + \varepsilon_t \quad (16.27)$$

$$\text{Modèle 3 : } \Delta x_t = \gamma x_{t-1} + \sum_{i=1}^{p-1} \gamma_i \Delta x_{t-i} + \varepsilon_t \quad (16.28)$$

On utilisera la procédure *urauto.src* qui reproduit exactement cette stratégie de test disponible sur le site web et extrait du site [http : www.estima.com](http://www.estima.com).

2. Vérifier que les indices boursiers pris en logarithmes sont des processus intégrés d'ordre $I(1)$ avec dérive.

Problème 16.2

Estimation de la relation de cointégration

1. Estimer par les MCO la relation d'équilibre (16.13) du modèle de valorisation et récupérer la série des résidus estimés.
2. Effectuer le test de stationnarité sur les résidus pour s'assurer que le modèle estimé constitue bien une relation de cointégration.
3. Estimer par les MCO les trois modèles à correction d'erreur qui peuvent être associés à la relation de cointégration. Conclure sur les mécanismes de retour à l'équilibre en cas de déformation de la relation de long terme.

Problème 16.3

Indicateur de valorisation, niveaux théoriques de l'indice et degré de surévaluation

1. Définir l'indicateur de valorisation et représenter son évolution dans le temps encadrée des valeurs ± 1 . Analyser graphiquement les propriétés dynamiques de l'indicateur.
2. Sur la base de la série de Earning Yield théorique, définir le niveau théorique de l'indice ainsi que l'intervalle de confiance obtenu en corrigeant les niveaux théoriques de ± 1 écart type.
3. Définir et représenter graphiquement le degré de surévaluation de l'indice boursier.

4. ÉLÉMENTS DE CORRIGÉ

Les éléments de correction proposés ci-après portent uniquement sur le marché boursier américain.

Problème 16.1

Tests de stationnarité

- Le tableau 16.1 propose les retards p (ordre de l' $AR(p)$) à introduire dans les tests ADF de stationnarité. Les critères de sélection ne fournissent pas on le voit une réponse uniforme. Il y a là l'illustration que les critères AIC et « General to Simple » ont tendance à proposer un nombre de retards plus élevé. On choisit de retenir le critère BIC qui suggère d'introduire 7 retards dans le test ADF sur $T3M$, et 1 seul retard pour les deux autres variables.

Tableau 16.1 Détermination du nombre de décalages

Critères	T3M	T10Y	EY
AIC	8	2	1
BIC	7	1	1
Ljung-Box Test	8	2	1
Lagrange Multiplier Test	3	1	1
General to Simple	7	1	6

Le tableau 16.2 reproduit les résultats de la procédure *urauto.src* sur les trois variables.

Tableau 16.2 Test ADF

	Modèle 1		Modèle 2		Modèle 3	Conclusion
	t_τ $H_0 : \gamma = 0$	Φ_3 $H_0 : \gamma = \beta = 0$	t_μ $H_0 : \gamma = 0$	Φ_1 $H_0 : \gamma = \beta = 0$	t_p $H_0 : \gamma = 0$	
T3M	-1,27	1,48	-1,69	1,93	-1,58	I(1)
T10Y	-2,98	4,51	-1,69	1,91	-1,43	I(1)
EY	-2,89	4,33	-1,61	2,66	-2,12	I(0)
$\Delta T3M$	-12,92	-	-	-	-	I(0)
$\Delta T10Y$	-10,55	-	-	-	-	I(0)
ΔEY	-10,89	-	-	-	-	I(0)
Valeur Critique à 5 %	-3,41	6,25	-2,86	4,59	-1,95	

Ces résultats appellent plusieurs commentaires.

Pour les 3 variables, le modèle pertinent pour réaliser le test de racine unitaire est le modèle 3, c'est-à-dire le modèle sans trend ni constante. La séquence des décisions est la suivante.

Sur la base du modèle 1, on ne peut rejeter l'hypothèse nulle de racine unitaire car $t_\tau > -3.41$, ni l'hypothèse jointe $\gamma = \beta = 0$ car $\Phi_3 < 6.25^9$. Sur la base du modèle 2, on ne peut rejeter l'hypothèse nulle de racine unitaire car $t_\mu > -2.86$, ni l'hypothèse jointe $\gamma = \mu = 0$ car $\Phi_1 < 4.59$.

⁹ On rappelle que ces valeurs critiques fournies par le logiciel sont tirées de Fuller (1976, page 373) pour le t -test et de Dicker-Fuller (1981, pour le F -test). Les seuils des tests ADF correspondent aux seuils des tests DF pour $T = \infty$.

Sur la base du modèle 3, on ne peut rejeter l'hypothèse nulle de racine unitaire pour les deux taux d'intérêt car $t_\rho > -1.95$. En revanche pour la série *EY*, $t_\rho = -2.12$ permet de rejeter l'hypothèse de racine unitaire au seuil de 5 %, mais pas au seuil de 1 % (valeur critique : -2.58).

On peut donc conclure que les processus suivis par les deux taux d'intérêt admettent une racine unitaire mais pas de dérive. Le résultat obtenu sur le *Earning Yield* est inconfortable. Il est fragile car il n'est pas significatif à 1 %. Il s'oppose au résultat obtenu sur les taux d'intérêt alors que les séries ont des profils d'évolution assez semblables. Enfin et surtout, il conduit à considérer la série de *Earning Yield* comme stationnaire autour d'une moyenne nulle, ce qui n'est pas acceptable économiquement. On adoptera donc ici une position prudente en choisissant de ne pas rejeter l'hypothèse nulle de racine unitaire, c'est-à-dire de non stationnarité.

Il est donc nécessaire de réaliser les tests sur les séries de différence première afin de déterminer précisément l'ordre d'intégration des séries. On sait pour l'instant qu'elles sont au moins intégrées d'ordre 1 ($I(1)$). La même procédure est appliquée. On rejette alors sans ambiguïté l'hypothèse nulle de racine unitaire ($t_\tau < -3.41$). Les séries de différence première sont stationnaires. Il est donc permis de conclure que les trois séries sont des processus $I(1)$ sans dérive, ce qui constitue concernant les taux d'intérêt une confirmation d'un résultat largement répandu dans la littérature empirique. L'absence de dérive dans la dynamique des séries signifie que leur non stationnarité est simplement le fait d'une variance croissante avec le temps.

Deux remarques doivent être faites sur ces résultats et la procédure *urauto*.

Cette procédure de test, proposée par Dolado, Jenkinson et Sosvilla-Rivero (1990)¹⁰, n'est véritablement à conseiller que lorsque le processus générateur des séries (*Data generating process*) est inconnu, c'est-à-dire lorsque l'économètre ne dispose pas de schéma théorique *a priori* pour interpréter l'évolution des variables. En pratique, le fait de ne pas écarter au départ l'hypothèse d'un trend, c'est-à-dire d'une tendance déterministe, est surtout adapté au cas des séries macroéconomiques affichant une tendance positive sur longue période, révélatrice de la croissance des économies : c'est en premier lieu le cas des séries de PIB, de consommation, d'investissement... Les tests réalisés sur la base du modèle 1 permettent alors justement de savoir si cette tendance est de type déterministe (série TS, trend stationary) ou au contraire stochastique, c'est-à-dire liée à la présence d'une racine unitaire et d'un terme de dérive (série DS, difference stationary).

La présence d'une tendance déterministe sur les trois variables d'intérêt de notre modèle est ici complètement exclue d'un point de vue économique, et ce en dépit de la tendance négative observée sur la période. Les seules positions tenables du point de vue de la théorie économique et de l'évolution générale de ces séries sont les suivantes : soit les séries sont stationnaires autour d'une moyenne non nulle (la constante du modèle 2) ; soit elles présentent une racine unitaire mais sans terme de dérive : l'espérance de la série est constante et donnée par la valeur initiale du processus, la racine unitaire n'a d'effet que sur la variance qui croît avec le temps.

¹⁰ Dolado J., Jenkinson T., Sosvilla-Rivero S., 1990 « Cointegration and Unit Roots », *Journal of Economic Surveys*, 4, 249-73.

Les modèles 2 et 3 étaient donc *a priori* ici les plus adaptés pour pratiquer les tests de stationnarité.

Notons enfin que le seuil global du test séquentiel est inconnu et supérieur au seuil du test de racine unitaire sur le modèle 3 car ce dernier constitue la dernière étape de plusieurs tests emboîtés.

2. Lorsqu'on applique le test ADF sur la série d'indice de prix en logarithme, on constate que la série en niveau est intégrée d'ordre 1 et que la série en différence est stationnaire.

Problème 16.2

Estimation de la relation de cointégration

1. L'estimation de (16.13) sur la période 1980 : 1 - 2000 : 08 ($T = 248$) donne les résultats suivants

$$(E/P)_t = -0.82 + 0.17 T3M_t + 0.78 T10Y_t$$

$$\begin{matrix} & (-3.46) & (4.42) & (14.58) \end{matrix}$$

$$\bar{R}^2 = 0.87, \hat{\sigma}_{\hat{u}_t} = 0.99, DW = 0.22, Q(36 - 0) = 630$$

2. Le test *ADF* sur les résidus \hat{u}_t est réalisé après avoir déterminé selon la procédure *laselec* le nombre de retards à introduire dans le modèle. Le critère *BIC* suggère $p = 1$. Il peut être directement réalisé à partir du modèle 2 (16.27) qui donne $t_\rho = -3.80$. Les valeurs critiques à 5 % et 1 % sont d'après la table de Phillips-Ouliaris (1990, page 190)¹¹ -3.77 et -4.31 (régression de type (b) avec constante et deux régresseurs ($N = 2$)¹²). Les résidus sont donc stationnaires. La relation estimée est bien une relation de cointégration s'interprétant comme un équilibre qui ne peut être durablement déformé. Elle permet ainsi de définir à chaque période compte tenu des taux d'intérêt un niveau fondamental de Earning Yield pour le marché boursier.

Les statistiques de Durbin-Watson ($DW < 2$) et de Ljung-Box (P-value < 0.05) révèlent une autocorrélation des résidus qui doit être positive étant donnée la faible valeur de *DW*. Cette autocorrélation positive est plutôt naturelle. Elle traduit l'ajustement progressif de la variable d'Earning Yield et donc du niveau de la bourse vers leur valeur d'équilibre.

Il faut noter ici une plus grande sensibilité du Earning Yield au taux d'intérêt à 10 ans et une sensibilité globale aux taux qui est proche de 1 ($0.95 = 0.17 + 0.78$) conformément à la relation d'équilibre du modèle théorique. La constante de la régression peut s'interpréter comme l'excès (en moyenne) de la prime de risque sur la croissance de long terme des bénéfices des sociétés.

Rappelons enfin que si les estimateurs des trois paramètres sont ici superconvergens du fait de la propriété de cointégration, l'inférence statistique usuelle et en particulier les tests de significativité de Student ne sont pas praticables. Les lois suivies par les t-stat ne sont pas standard.

¹¹ Phillips PCB, Ouliaris S., *Asymptotic Properties of residual based tests for cointegration*, *Econometrica*, 1990, 58, 165-193.

¹² Les valeurs critiques du test ne sont pas identiques aux tables de Dickey-Fuller de la section précédente car la série des résidus est obtenue par une estimation MCO.

3. L'estimation des trois modèles à correction d'erreur pouvant être associés à cette relation de cointégration donne les résultats suivants

$$\begin{aligned} \Delta(E/P)_t &= -0.03 + 0.03 \Delta T3M_t + 0.22 \Delta T10Y_t - 0.11 \hat{u}_{t-1} \\ &\quad (-1.57) \quad (1.18) \quad (3.63) \quad (-5.50) \\ \Delta T3M_t &= 0.14 \Delta(E/P)_t + 1.15 \Delta T10Y_t - 0.005 \hat{u}_{t-1} \\ &\quad (1.18) \quad (11.60) \quad (-0.12) \\ \Delta T10Y_t &= -0.07 + 0.42 \Delta(E/P)_t + 0.008 \Delta T3M_t + 0.05 \hat{u}_{t-1} \\ &\quad (-1.11) \quad (5.53) \quad (1.12) \quad (2.16) \end{aligned}$$

Seul le premier modèle décrivant la dynamique de court terme de E/P peut être considéré comme un modèle à correction d'erreur¹³, puisqu'il fait ressortir pour \hat{u}_{t-1} un coefficient négatif (-0.11) et significatif ($t\text{-stat} = -5.50$) qui traduit bien la notion de force de rappel. Ces estimations confirment donc l'exogénéité des taux d'intérêt. Le retour à l'équilibre en cas de déformation de la relation de cointégration se fait donc bien par un mouvement de la bourse et non pas par un mouvement de l'ensemble des trois variables. Le délai moyen de retour à l'équilibre est de $1/0.11 \simeq 9$ mois.

Problème 16.3

Indicateur de valorisation, niveaux théoriques de l'indice et degré de surévaluation

1. L'analyse graphique des propriétés dynamiques de l'indicateur de valorisation calculé selon (16.21)¹⁴ est intéressante (figure 16.2).

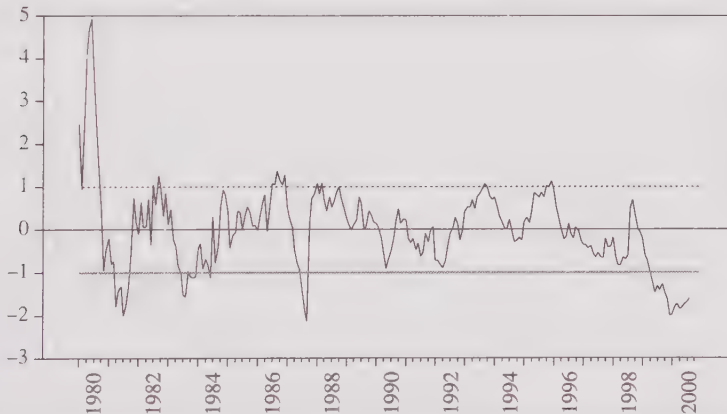


Figure 16.2 Indicateur de valorisation — États-Unis

¹³ Le cas du 3^e modèle est limite, le $t\text{-stat}$ du coefficient de \hat{u}_{t-1} étant égal à 2.16.

¹⁴ Il est calculé à partir des résidus normalisés de l'équation de cointégration estimée dans le problème 16.2, il est distribué asymptotiquement selon une loi normale centrée réduite.

L'autocorrélation positive des résidus révélée par le test de Durbin-Watson est très apparente. Le marché boursier américain passe donc par des phases successives de sous-évaluation et de surévaluation. La propriété la plus intéressante concerne le profil de retour à l'équilibre lorsque la déformation est significative (indicateur proche de -1 ou 1). Le retour vers l'équilibre est dans ce cas très rapide et le plus souvent intégral (indicateur proche de 0) et s'accompagne parfois même d'une déformation significative de sens opposé (krach de 1987). Cette propriété fait que l'indicateur de valorisation peut devenir un instrument de prévision efficace dans ces circonstances particulières de déformation significative de la relation de long terme. Les résultats obtenus sur les modèles à correction d'erreur nous assurent en effet qu'en général la correction s'effectue bien par la bourse.

Le diagnostic du modèle sur la situation de fin de période (1999-2000) est tranché. L'indicateur IV_t proche de -2 traduit une surévaluation très significative du marché.

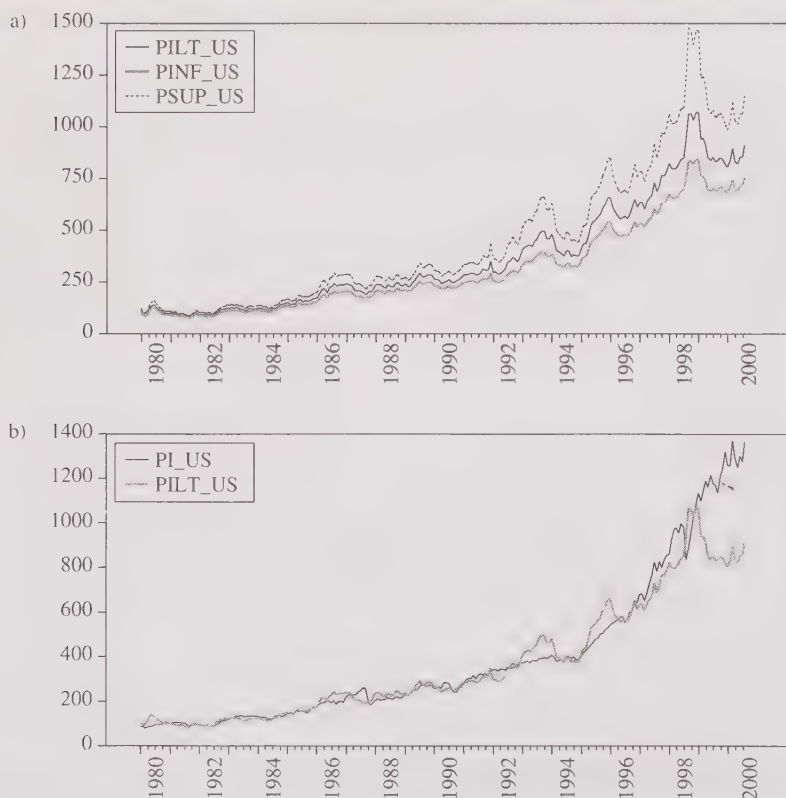


Figure 16.3 *Price index et niveaux théoriques — États-Unis. a) Price index de long terme théorique, b) Price index et niveau théorique*

2. La figure 16.3 représente d'une part l'évolution de la valeur théorique de l'indice boursier selon l'équation (16.22) encadrée par des bornes inférieure et supérieure et, d'autre part, l'évolution comparée de l'indice boursier et de sa valeur théorique. La surévaluation boursière est de nouveau manifeste : la valeur de l'indice en fin de période est largement supérieure à sa valeur théorique de long terme. On peut vérifier en outre qu'elle dépasse amplement le borne supérieure de l'estimation de la valeur théorique.
3. La figure 16.4 représente le degré de surévaluation de l'indice boursier selon la formule (16.25).

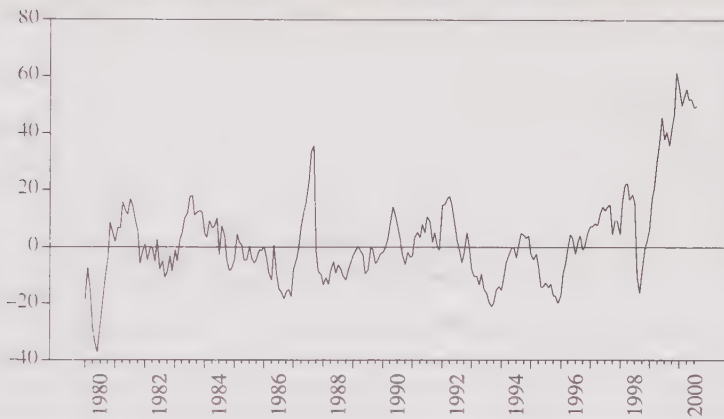


Figure 16.4 Degré de surévaluation de l'indice boursier — États-Unis

Au vu de ces résultats, il est légitime d'anticiper une correction significative du marché boursier américain pour les mois suivants. Les éléments susceptibles de démentir cette conclusion sont, comme toujours en économétrie, à chercher dans une remise en cause plus ou moins radicale des résultats de la modélisation.

Il est tout d'abord permis de considérer que, même si en moyenne les retours à l'équilibre se font par le biais d'un mouvement de la bourse, ils puissent ponctuellement s'opérer par les taux d'intérêt. En particulier ici toute détente monétaire (baisse des taux courts) contribuerait à diminuer la valeur théorique du rendement du marché boursier et ainsi à réduire la surévaluation boursière.

Plus radicalement, il est toujours possible de croire en une instabilité des comportements sur le marché boursier aboutissant à une rupture dans la relation de cointégration entre Earning Yield et taux d'intérêt. L'hypothèse la plus vraisemblable serait alors une diminution de la constante du modèle derrière laquelle se cachent sans doute des niveaux moyens pour la prime de risque et la croissance de long terme des bénéfices des sociétés. La prime de risque est susceptible d'avoir baissé dans la période récente en raison d'une baisse de l'aversion pour le risque des investisseurs épargnants à la suite de plusieurs années de fortes hausses boursières (lien richesse-aversion absolue pour le risque). Parallèlement les anticipations à long terme pour

la croissance des bénéfices ont pu être révisées à la hausse par des investisseurs interprétant la forte croissance de l'économie à la fin de la décennie 1990 comme un phénomène durable, c'est-à-dire comme le signe d'une élévation de la croissance potentielle de l'économie américaine (thème de la «nouvelle économie»).

Une analyse similaire peut-être réalisée à partir des données sur le marché boursier français.

Les évolutions boursières postérieures à la préparation de ce manuel ont montré qu'il n'était pas forcément nécessaire de privilégier l'hypothèse d'une instabilité du modèle économétrique de valorisation...

17

LA RELATION SALAIRES — PRIX À LONG TERME

1. Introduction
2. Cadre analytique
3. Stratégie d'estimation
4. Éléments de corrigé

OBJECTIFS DU CHAPITRE

D'un point de vue méthodologique

- Modèles à corrections d'erreurs
- Estimation de relations de cointégration
- Méthode de Johansen
- Critères d'information : AIC, BIC, ...
- Test de Ljung Box, du multiplicateur de Lagrange et du ratio de vraisemblance
- Test de racines unitaires, Test de cointégration, Test d'exogénéité faible

D'un point de vue économique

- Analyse du marché du travail
- Évaluation du modèle WS-PS (formation des salaires et des prix)

NIVEAU

MASTER, deuxième année

CHAMPS ÉCONOMIQUES TRAITÉS

Macroéconomie réelle

Économie du travail

1. INTRODUCTION

Les piètres performances des économies européennes en matière de chômage ont conduit les théoriciens des marchés du travail à s'interroger sur les facteurs explicatifs des rigidités constatées sur ces marchés, à partir de l'estimation d'équations de salaire. La littérature économique a longtemps privilégié dans cette optique l'estimation d'une courbe de Phillips, qui décrit la relation entre la variation du salaire (réel) et le taux de chômage, variable en niveau. Ce type de spécification souffre néanmoins de deux imperfections.

D'un point de vue théorique, elle est mise en cause pour son manque de fondements microéconomiques. La courbe de Phillips n'est en effet qu'une relation économétrique exprimant la sensibilité des salaires au taux de chômage ou d'autres déterminants des salaires (fiscalité, anticipation des prix...) ajoutés intuitivement de manière *ad hoc*. Les fondements théoriques ne sont jamais explicités clairement, de sorte que la confrontation de la courbe de Phillips et d'une équation de prix, exprimant le comportement des producteurs, aboutit à un taux de chômage d'équilibre au contenu informatif un peu flou et qui explique difficilement l'augmentation du chômage en France. En effet, cette approche conduit par définition à un concept de taux de chômage stationnaire peu compatible avec l'évolution du chômage effectif en France, qui connaît une tendance persistante sur la période étudiée (Cotis, Meary et Sobczak, 1996)¹. La spécification des équations de salaire et de prix en taux de croissance, hormis le niveau de chômage, revient à considérer uniquement les effets transitoires des chocs sur le taux de chômage d'équilibre, contrairement à l'observation des faits.

D'un point de vue méthodologique, l'estimation d'une courbe de Phillips n'est valable que si le taux de chômage suit un processus stationnaire. A défaut, les outils référentiels de l'économétrie classique se révèlent inadaptés pour l'étude des séries non stationnaires.

Pour répondre à ces critiques les travaux empiriques se sont donc naturellement tournés vers l'estimation simultanée d'une quasi offre et d'une quasi demande de travail spécifiées en niveau. Ce système est connu sous le nom de modèle WS-PS (Wage Schedule — Price Schedule). La première relation est dérivée des comportements microéconomiques des offreurs de travail dans le cadre des nouvelles théories du marché du travail (notamment la théorie des négociations salariales). La deuxième relation décrit la formation des prix par une entreprise dont l'output est offert sur un marché des biens imparfaitement concurrentiel. La confrontation des deux équations structurelles détermine le taux de chômage d'équilibre, influencé par les chocs affectant les déterminants des prix et des salaires.

L'estimation de ces deux relations est rendue possible par le développement parallèle de l'économétrie des séries non stationnaires. Deux approches différentes sont envisageables. La première consiste à estimer séparément les deux relations par les MCO et à tester ensuite la stationnarité des résidus (procédure de Engel et Granger). Cette technique est ici inadaptée car elle suppose l'existence d'une seule

¹ Cotis, J.-P., Meary R., Sobczak N., 1996, « Le chômage d'équilibre en France : une évaluation », document de travail n° 96-14, décembre, Direction de la Prévision.

relation de cointégration entre les variables et ne permet pas, de fait, d'identifier les formes structurelles de salaire et de prix. La deuxième approche requiert l'utilisation de la procédure multivariée de Johansen, qui permet de tester le nombre de relations de cointégration, et d'estimer simultanément les équations WS et PS.

Ce chapitre reprend le cadre d'analyse de L'Horty et Sobczak (1997)². La section 2 présente succinctement la construction du modèle WS-PS et les problèmes liés à l'identification du modèle. La section 3 formule la stratégie d'estimation du modèle. *

2. CADRE ANALYTIQUE

2.1 Le modèle WS-PS

La représentation des marchés du travail sous la forme d'un modèle WS-PS a été popularisée par les travaux de Layard, Nickell et Jackman (1991)³. La plupart de la recherche empirique actuelle utilise en effet couramment cette structure, en y apportant des modifications marginales.

Ces auteurs analysent la formation des prix et des salaires à partir de deux relations structurelles formalisant les comportements d'offre et de demande de travail. Sans entrer dans une formalisation poussée de ces relations, il convient toutefois d'en rappeler les fondements.

2.1.1 La courbe PS

Soit $Q = F(K, L)$ la fonction de production avec K et L les niveaux de capital et de travail utilisés.

La courbe PS traduit une quasi demande de travail. Elle est déterminée à partir du programme du producteur qui maximise son profit dans le cadre d'un marché des biens imparfaitement concurrentiel. Le producteur est confronté à une fonction de demande inverse du bien, supposée décroissante et qui est définie par

$$P = P(Q), \quad \frac{dP}{dQ} < 0 \quad (17.1)$$

L'élasticité du prix par rapport à la production est notée

$$\eta = \frac{\partial P}{\partial Q} \frac{Q}{P} \quad (17.2)$$

² L'Horty Y., Sobczak N., 1997, « Identification de la courbe de salaire et déterminants du chômage d'équilibre dans un modèle de négociation salariale », *Économie et Prévision*, n° 27, 1, pp. 101-117.

³ Layard R., Nickell S., Jackman R., 1991, *Unemployment : macroeconomic performance and the labour market*, Oxford University Press.

À capital constant, la demande de travail est déterminée à partir du programme

$$\begin{aligned} \underset{L}{\text{Max}} \pi &= P(Q) \cdot Q - W_n \cdot L - C_k \cdot K \\ &\text{avec } Q = F(K, L), K \text{ fixé} \end{aligned} \quad (17.3)$$

où W_n et C_k désignent respectivement le salaire nominal et le coût d'une unité de capital. La solution du programme est donnée par les conditions du premier ordre, soit

$$\begin{aligned} \frac{\partial \pi}{\partial L} &= 0 \\ \Leftrightarrow \frac{\partial P}{\partial Q} \frac{\partial Q}{\partial L} Q + P \frac{\partial Q}{\partial L} - W_n &= 0 \\ \Leftrightarrow P \frac{\partial Q}{\partial L} \left(\underbrace{\frac{\partial P}{\partial Q} \frac{Q}{P}}_{\eta} + 1 \right) - W_n &= 0 \end{aligned} \quad (17.4)$$

La demande de travail est donc définie par

$$\frac{\partial Q}{\partial L} = \frac{1}{1 + \eta} \frac{W_n}{P} \quad (17.5)$$

À l'équilibre, le producteur choisit un niveau d'emploi tel que la productivité marginale du travail égale le salaire réel majoré d'un taux de marge.

Considérons par exemple une fonction CES (Constant Elasticity of Substitution) à rendements constants. Elle s'écrit

$$Q = [\alpha K^{-\rho} + (1 - \alpha)L^{-\rho}]^{-\frac{1}{\rho}} \quad (17.6)$$

avec $0 < \alpha < 1$ et $-1 < \rho < \infty$.

Dans ce cas, la condition du premier ordre donne

$$\frac{W_n}{P} = (1 + \eta) [Q^{1+\rho} (1 - \alpha)L^{-(1+\rho)}] \quad (17.7)$$

Prise en logarithme, cette expression devient

$$\log \left(\frac{W_n}{P} \right) = \log(1 + \eta) + (1 + \rho) \log Q + \log(1 - \alpha) - (1 + \rho) \log L \quad (17.8)$$

c'est-à-dire

$$w_n - p = \frac{1}{\sigma}(q - l) + C \quad (17.9)$$

où $\sigma = \frac{1}{1+\rho}$ est l'élasticité de substitution des deux facteurs de production, w_n , p , q et l sont respectivement les logarithmes du salaire nominal, du prix du bien, de la production et du facteur travail. C rassemble les termes constants $\log(1 + \eta)$ et $\log(1 - \alpha)$. Le chômage n'apparaît dans PS qu'en décomposant le terme $q - l$. L'équation peut être en effet réécrite sous la forme

$$w_n - p = \frac{1}{\sigma}(q - l + n - n) + C \quad (17.10)$$

où n désigne le logarithme de la population active N . On a finalement⁴

$$w_n - p = \frac{1}{\sigma}(q - n + U) + C \quad (17.11)$$

avec U le taux de chômage.

2.1.2 La courbe WS

La courbe WS peut être formulée à partir de plusieurs théories des marchés du travail, principalement celle des négociations salariales. La principale conclusion de ce schéma théorique est que le salaire issu des négociations entre une firme et un syndicat est fixé à un niveau supérieur au salaire concurrentiel. L'offre de travail n'est plus individuelle comme dans le cas de la représentation néoclassique de ces marchés, mais est remplacée par une offre de travail syndicale. Celle-ci est déterminée à partir d'une négociation formalisée par le critère de Nash généralisé, du type

$$\underset{w}{Max} (V - V^*)^\gamma (\pi - \pi^*)^{1-\gamma}$$

où V et π sont respectivement l'utilité du syndicat et le profit de la firme, et V^* et π^* sont les situations de *statu quo* correspondant à l'échec des négociations. L'objectif du syndicat est de capter un revenu supplémentaire qui est fonction croissante de

⁴ On définit le taux de chômage U par

$$U = 1 - \frac{L}{N} \Rightarrow \log l - \log N = l - n = \log(1 - U)$$

au voisinage de 0 on peut utiliser l'approximation

$$\log(1 - U) \simeq -U$$

d'où

$$n - l \simeq U$$

son pouvoir de négociation γ . Dans la mesure où cette variable n'est pas directement observable, on suggère d'utiliser le taux de chômage comme approximation du pouvoir de négociation. Le mécanisme sous-jacent est le suivant : une baisse du chômage renforce le pouvoir de négociation du syndicat qui accroît ses prétentions salariales. La relation WS traduit donc une relation décroissante entre salaire réel et chômage. Enfin, dans la mesure où la formation des salaires résulte d'une négociation entre employés et employeur, cela implique que le salaire négocié dépend à la fois des paramètres des fonctions objectifs des deux joueurs. Toute variable introduite dans PS doit donc figurer dans WS.

2.1.3 Le modèle WS-PS

Dans sa version la plus simple, le modèle WS-PS peut être représenté par le système suivant

$$\begin{aligned} W &= a_{21}Y - a_{31}U + Z_w, & \text{WS} \\ W &= a_{22}Y + a_{32}U + C, & \text{PS} \end{aligned} \quad (17.12)$$

avec $W = w_n - p$ le salaire réel et $Y = (q - n)$ la productivité, exprimés en logarithme⁵. Les paramètres sont positifs. La variable Z_w rassemble tous les autres déterminants des salaires. Certains travaux considèrent par exemple l'impact de la pression fiscale, des termes de l'échange, du degré de centralisation des négociations, entre autres, sur le processus de formation des salaires. Néanmoins ces variables sont écartées de l'analyse pour plusieurs raisons. D'une part, l'indisponibilité de certaines séries oblige le recours à la construction de proxy dans la fréquence de données souhaitée. D'autre part, la prise en compte d'un nombre important de variables complique considérablement la recherche de relations structurelles dans le cadre d'une modélisation multivariée à la Johansen. Dans un pur souci de simplification, l'exposé est concentré sur l'étude des interdépendances des salaires réels, de la productivité et du chômage.

2.2 Identification du modèle théorique

Avant de procéder à l'estimation du système, il convient tout d'abord de déterminer les contraintes du modèle. Le problème est ici évident. Le chômage et la productivité apparaissent dans les deux équations. Aucune variable d'une relation n'est exclue de l'autre. L'identification du modèle repose alors sur deux hypothèses.

1. On suppose en général que la productivité n'apparaît pas dans WS. Cette hypothèse a été formulée par Manning (1993)⁶, selon lequel la productivité n'intervient pas directement dans l'équation structurelle des salaires, mais influence ces derniers via les déplacements exogènes de PS.

⁵ Les variables sont exprimées en logarithme hormis le taux de chômage.

⁶ Manning A., 1993, « Wage Bargaining and the Phillips Curve: the Identification and the Specification of Aggregate Wage Equations », *The Economic Journal*, vol. 103, pp. 98-118.

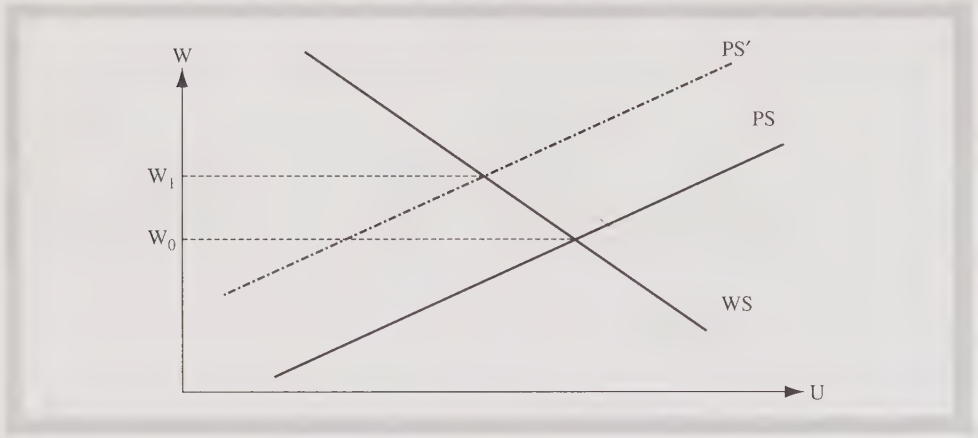


Figure 17.1 Le modèle WS-PS

La figure 17.1 représente le modèle WS-PS. La croissance de la productivité déplace PS vers la gauche en PS', vers un niveau de salaire $W_1 > W_0$.

Dans la pratique, on pose $a_{21} = 0$ et on introduit un trend linéaire (noté t) qui capte les gains de productivité tendancielle de l'économie sur lesquels s'indexent partiellement les revendications salariales.

- La relation PS est identifiée par une contrainte théorique issue de la construction même de cette relation. Dans (17.11) les coefficients associés au chômage et à la productivité de plein emploi sont égaux $a_{22} = a_{32} = \frac{1}{\sigma}$ (cf. 17.11).

Le modèle exactement identifié qui sera estimé est donc de la forme

$$\begin{array}{ll} W = -a_{31}U + a_{41}t, & \text{WS} \\ W = a_{22}[Y + U] + C, & \text{PS} \end{array} \quad (17.13)$$

Les endogènes du système sont les salaires réels et le taux de chômage. La productivité est supposée exogène.

3. STRATÉGIE D'ESTIMATION

L'estimation du modèle WS-PS par la procédure de Johansen suppose de considérer successivement un certain nombre d'étapes. Il s'agit tout d'abord de définir l'ordre d'intégration des séries. Si les séries sont stationnaires, alors il est possible d'estimer le modèle à partir du système VAR (17.14). Sa principale caractéristique est de prendre en compte l'ensemble des interactions des variables sans imposer de contraintes théoriques entre celles-ci. Formellement, le modèle s'écrit

$$\begin{pmatrix} W \\ Y \\ U \end{pmatrix}_t = A_1 \begin{pmatrix} W \\ Y \\ U \end{pmatrix}_{t-1} + \dots + A_k \begin{pmatrix} W \\ Y \\ U \end{pmatrix}_{t-k} + \Psi \begin{pmatrix} (1) \\ du80 \\ du82 \\ du83 \end{pmatrix} + \underbrace{\begin{pmatrix} \varepsilon_w \\ \varepsilon_y \\ \varepsilon_u \end{pmatrix}}_{\varepsilon_t}$$

$\varepsilon_t \sim N(0, \Sigma)$

(17.14)

où W , Y et U désignent respectivement le logarithme du salaire réel, le logarithme de la productivité de plein emploi et le taux de chômage. D est un vecteur qui intègre la constante (1) et d'éventuelles variables indicatrices, qui servent à rendre compte des points aberrants d'une série ou d'une rupture dans la tendance de celle-ci. Elles sont justifiées par l'occurrence d'un choc économique (choc pétrolier en 1974 par exemple) ou le renversement des orientations de politiques économiques poursuivies par un gouvernement. Elles peuvent être enfin introduites pour améliorer les propriétés statistiques des résidus du système. Dans l'application, les données sont trimestrielles et couvrent la période 1977 : 1-1995 : 4. Trois variables indicatrices sont considérées : $DU80$ qui prend la valeur 1 sur la période 1980 : 1 -1985 : 4, $DU82$ qui prend la valeur 1 en 1982 : 2 et 1982 : 3 et $DU83$ qui prend la valeur 1 en 1983 : 1. La première indicatrice est intégrée pour améliorer les propriétés des résidus. Les deux suivantes correspondent respectivement aux politiques menées sur le marché du travail français (blocage des salaires et lois Auroux) et au changement d'orientation de la politique économique. ε_t est un vecteur d'aléas distribués selon une loi normale.

Si les séries sont intégrées d'ordre 1, l'estimation du modèle WS-PS passe par la réécriture du système multivarié (17.14) sous une forme VECM (modèle vectoriel à correction d'erreur)

$$\begin{pmatrix} \Delta W \\ \Delta Y \\ \Delta U \end{pmatrix}_t = \Gamma_1 \begin{pmatrix} \Delta W \\ \Delta Y \\ \Delta U \end{pmatrix}_{t-1} + \dots + \Gamma_k \begin{pmatrix} \Delta W \\ \Delta Y \\ \Delta U \end{pmatrix}_{t-k+1} + \Pi \begin{pmatrix} W \\ Y \\ U \end{pmatrix}_{t-1} + \Psi \begin{pmatrix} (1) \\ du80 \\ du82 \\ du83 \end{pmatrix} + \varepsilon_t$$
(17.15)

avec $\Gamma_i = \sum_{j=i+1}^k -A_j$, $i = 1, \dots, k-1$ et $\Pi = A_1 + \dots + A_k - I$. Δ est l'opérateur différence. Ce modèle contient à la fois des informations sur la dynamique de court terme des variables via les matrices Γ_i et sur l'ajustement de long terme via la matrice Π . Celle-ci renseigne à la fois sur la vitesse d'ajustement des variables

vers l'équilibre de long terme (matrice α), et sur les paramètres des relations de cointégration (matrice β) qui déterminent des combinaisons linéaires stationnaires des variables en niveau du système. La matrice Π peut être en effet réécrite sous la forme suivante dans le cas de deux relations de cointégration⁷

$$\Pi Z_{t-1} = \underbrace{\begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{bmatrix}}_{\alpha} \underbrace{\begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} \\ \beta_{12} & \beta_{22} & \beta_{32} \end{bmatrix}}_{\beta'} \underbrace{\begin{pmatrix} W \\ Y \\ U \end{pmatrix}}_{Z_{t-1}} \quad (17.16)$$

On s'intéresse ici précisément à la valeur des coefficients β , dans la mesure où ils permettent d'obtenir les paramètres structurels recherchés des relations WS et PS. En normalisant (17.16) par rapport à W , et en posant

$$\begin{aligned} a_{21} &= -\frac{\beta_{21}}{\beta_{11}}, & a_{31} &= \frac{\beta_{31}}{\beta_{11}} \\ a_{22} &= -\frac{\beta_{22}}{\beta_{12}}, & a_{32} &= -\frac{\beta_{32}}{\beta_{12}} \end{aligned} \quad (17.17)$$

$\beta' Z$ donne le système (17.12).

Les paramètres α indiquent comment les déviations de l'équilibre de long terme affectent la dynamique de court terme des variables du système pour retrouver cet équilibre. Par exemple, en cas de déséquilibre, le paramètre α_{11} donne la vitesse à laquelle le salaire réel s'ajuste lorsque la relation de cointégration est donnée par $(\beta_{11}W + \beta_{21}Y + \beta_{31}U)$. En d'autres termes les relations de cointégration, relations d'équilibre de long terme, agissent comme des forces de rappel sur la dynamique de court terme des variables, et sont ainsi qualifiées de mécanismes correcteur d'erreurs par rapport à l'équilibre de long terme.

Le cadre analytique étant posé, les différents problèmes rencontrés (qui constituent en même temps les étapes de l'analyse) dans l'utilisation de la procédure de Johansen sont les suivants

- Problème 17.1 : tester l'ordre d'intégration de chaque variable du VECM.
- Problème 17.2 : déterminer le nombre de retards du modèle VAR.
- Problème 17.3 : déterminer le nombre de relations de cointégration.
- Problème 17.4 : tester différentes hypothèses sur les relations de cointégration.
- Problème 17.5 : tester l'exogénéité faible des variables, et construire un VECM partiel conditionné par une variable faiblement exogène.

⁷ Des termes déterministes peuvent également être introduits dans les relations de cointégration mais sont ici volontairement omis. Ce point est revu au problème 17.3.

4. ÉLÉMENTS DE CORRIGÉ

L'application est menée sur données trimestrielles françaises et couvre la période 1977 : 1-1995 : 4 ($T = 72$ observations). Les séries utilisées sont des données corrigées des variations saisonnières de l'OCDE et du FMI. Le taux de salaire nominal correspond aux gains horaires perçus par les salariés de l'industrie manufacturière. L'emploi L et la production Q sont mesurés par un indice de l'emploi et de la production industriels. L'indice des prix utilisé est le déflateur du PIB. Enfin, le chômage est mesuré par le taux de chômage standardisé en % (U). On utilise l'approximation $U \simeq -\log(1 - U/100)$. De même, le logarithme de la productivité de plein emploi est défini en utilisant l'approximation du logarithme de plein emploi $\log N \simeq U + \log L$.

Le fichier chapitre 17.xls contient les séries brutes nécessaires à l'estimation du modèle. On note, emp : l'emploi (L), prod : la production (Q), sal : le salaire nominal (w_n), defpib : le prix (P) et cho : le taux de chômage (U).

Problème 17.1

Ordre d'intégration des séries

L'ordre d'intégration des séries est testé à partir des tests usuels de Dickey-Fuller augmentés. Pour chaque série, on teste successivement l'hypothèse nulle de racine unité en envisageant l'existence éventuelle d'une constante et d'un trend, puis uniquement une constante, puis sans constante ni trend. Cette démarche suppose tout d'abord de déterminer le nombre p de décalages à introduire pour le test. Les critères d'information AIC et BIC sont souvent utilisés dans cette tâche, mais ils peuvent être complétés par les tests de Ljung-Box (LB), du Multiplicateur de Lagrange (ML) et un test de significativité du dernier retard introduit de la variable étudiée (GS pour General to Simple). En général, les critères BIC, LB et LM donnent la même structure de retards et sont de fait privilégiés aux critères AIC et GS qui ont tendance à retenir un nombre trop important de retards. Ces tests sont programmés dans le fichier **lagslec.src**, disponible sur le **site d'Estima**⁸. Les résultats sont consignés dans les tableaux suivants, tout en privilégiant l'analyse des critères BIC, LB et ML pour les raisons évoquées ci-dessus. Le tableau 17.1 précise la méthode de décision du retard au regard du critère BIC. Le tableau 17.2 détaille les tests de non-autocorrélation des résidus selon les tests LB et ML. Le tableau 17.3 résume les retards sélectionnés par la procédure **lagslec.src**.

Dans le tableau 17.1 le critère d'information est calculé pour chaque retard p . Il s'agit de sélectionner le retard pour lequel la valeur du critère est la plus faible (1 retard pour les séries W et Y, 2 retards pour U).

⁸ <http://www.estima.com/>

Tableau 17.1 Critère BIC

AR(p)	W	Y	U
0	-6.11	-4.127	-7.826
1	-10.753	-8.768	-12.115
2	-10.751	-8.706	-12.654
3	-10.690	-8.679	-12.620
4	-10.644	-8.617	-12.570
5	-10.587	-8.560	-12.547
6	-10.637	-8.507	-12.508
7	-10.576	-8.451	-12.549
8	-10.523	-8.99	-12.491

Tableau 17.2 Tests de Ljung Box et du Multiplicateur de Lagrange

AR(P)	LB(18)*			ML(18)		
	H ₀ : non autocorrélation des résidus			H ₀ : non autocorrélation des résidus		
	W	Y	U	W	Y	U
0	996.78 (0.00)	1234.32 (0.00)	929.81 (0.00)	57.56 (0.00)	57.70 (0.00)	57.67 (0.00)
1	35.99 (0.00)	14.69 (0.54)	129.33 (0.00)	18.28 (0.44)	19.88 (0.34)	39.48 (0.00)
2	24.54 (0.05)		33.36 (0.00)			23.67 (0.17)

*l'ordre du test est fixé en divisant le nombre d'observations par 4.
Les chiffres entre parenthèses sont les P-values des tests.

Le tableau 17.2 indique que l'hypothèse nulle de non-autocorrélation des résidus est rejetée pour un AR(0) quel que soit le test utilisé et la série étudiée. En revanche, l'introduction d'un terme autorégressif modifie les résultats pour la productivité. On accepte l'hypothèse nulle de non-autocorrélation avec un niveau de significativité de 54 % d'après le test de Ljung Box, et de 34 % d'après le test du multiplicateur de Lagrange. La non-autocorrélation des résidus est également acceptée à 44 % pour les salaires réels d'après ce même test. L'intégration d'un terme autorégressif d'ordre 2 pour la série de taux de chômage permet de ne pas rejeter finalement l'hypothèse nulle de non-autocorrélation de cette série.

Tableau 17.3 Synthèse des tests : nombre de retards p

	AIC	BIC	LB	ML	GS
W	6	1	2	1	6
Y	1	1	1	1	1
U	7	2	8	2	7

D'après les tests, BIC et ML, nous retenons un retard pour les séries de salaire réel et de productivité et deux retards pour la série de chômage.

Les tests de racine unité ADF (Dickey Fuller Augmenté) peuvent être désormais appliqués. On utilise la stratégie de test séquentiel (cf. chapitre 15, Économétrie des séries non stationnaires), à l'aide de la procédure **urauto.src**⁹. L'hypothèse nulle de racine unité est testée successivement sur les modèles 1 à 3. La stratégie de test est interrompue dès que l'on rejette l'hypothèse H_0 .

$$\begin{aligned} \text{Modèle 1 : } \Delta x_t &= \mu + \beta t + \gamma^* x_{t-1} + \sum_{i=1}^{p-1} \gamma_i^* \Delta x_{t-i} + \varepsilon_t \\ \text{Modèle 2 : } \Delta x_t &= \mu + \gamma^* x_{t-1} + \sum_{i=1}^{p-1} \gamma_i^* \Delta x_{t-i} + \varepsilon_t \\ \text{Modèle 3 : } \Delta x_t &= \gamma^* x_{t-1} + \sum_{i=1}^{p-1} \gamma_i^* \Delta x_{t-i} + \varepsilon_t \end{aligned}$$

Le tableau 17.4 résume les statistiques calculées.

Tableau 17.4 Tests de Dickey-Fuller augmenté

	Modèle 1		Modèle 2		Modèle 3
	t_τ $H_0 : \gamma^* = 0$	Φ_3 $H_0 : \gamma^* = \gamma = 0$	t_μ $H_0 : \gamma^* = 0$	Φ_1 $H_0 : \gamma^* = \mu = 0$	t_ρ $H_0 : \gamma^* = 0$
W	- 2,49	5,78	- 3,11		
Stationnaire autour d'une moyenne non nulle					
U	- 2,21	2,55	1,40	1,93	0,97
Racine Unité					
Y	- 1,91	1,82	- 0,34	6,84	
Racine Unité avec dérive					
ΔW	- 5,25				
Stationnaire					
ΔU	- 3,06	4,68	- 3,00		
Stationnaire autour d'une moyenne non nulle					
ΔY	- 5,10				
Stationnaire					
Valeurs tabulées (5 %)					
	- 3,41	6,25	- 2,86	4,59	- 1,95

⁹ Cette procédure nécessite l'introduction d'au moins un retard pour l'étude de l'ordre d'intégration des séries différenciées. Or, on devrait théoriquement prendre 0 retard [(p-1) retards] pour les séries W et Y. On conserve par conséquent les retards suggérés (ie. p retards) pour les séries en niveau par la procédure **lagselec.src**.

L'ordre d'intégration de la série de productivité est commentée car son interprétation peut paraître plus délicate

1. On teste dans le modèle 1 l'hypothèse de racine unité suivant la statistique t_τ . L'hypothèse nulle n'est pas rejetée puisque nous avons $|-1.91| < |-3.41|$.
2. Le programme réalise ensuite un test joint de l'hypothèse de racine unité et de nullité du coefficient associé au trend. Ce test est réalisé avec la statistique Φ_3 . On ne peut rejeter à nouveau l'hypothèse nulle.
3. On passe alors à l'estimation du modèle 2. L'hypothèse de racine unité est à nouveau testée à l'aide de la statistique t_μ et également acceptée puisque $|-0.34| < |-2.86|$. Un test joint (Φ_1) vérifie alors l'opportunité d'introduire une constante conjointement à l'hypothèse de racine unité. Cette fois-ci, l'hypothèse nulle est rejetée ($|6.84| > |4.59|$).
4. Il est alors nécessaire d'approfondir la stratégie de tests. On va tester la nullité de la constante conditionnellement à l'hypothèse de racine unité. Le processus étant de fait différencié, les statistiques suivent des lois standards. Lorsqu'on estime la relation suivante

$$\Delta x_t = \mu + \sum_{i=1}^{p-1} \gamma_i^* \Delta x_{t-i} + \varepsilon_t$$

la statistique $\frac{\hat{\mu}}{\hat{\sigma}_\mu}$ vaut 3.71, ce qui confirme la significativité de la constante. Le programme reprend alors le test de racine unité dans le modèle 2, l'hypothèse $H_0 : \gamma^* = 0$ étant acceptée car $|-0.34| < |-2.86|$.

On peut donc conclure que la série de productivité est intégrée d'ordre 1 avec dérive (le processus étant différencié, la constante devient en effet une tendance). Enfin, on vérifie la stationnarité de la série différenciée en réitérant la stratégie de test. D'après le modèle 1, l'hypothèse de racine unité est rejetée ($|-5.10| > |-3.41|$). Il est inutile de poursuivre. La série de productivité est bien non stationnaire.

D'après le tableau 17.4, alors que la productivité et le chômage sont intégrées d'ordre 1, la série de taux de salaire réel apparaît stationnaire autour d'une constante. Ce résultat paraît surprenant, et doit être interprété avec prudence, les séries économiques étant souvent intégrées, notamment la série des salaires réels est usuellement intégrée d'ordre 1. On considère pour la suite que les variables du modèle sont I(1).

Problème 17.2

Nombre de décalages du modèle VAR et analyse résiduelle

Détermination du nombre de retard

Les séries étant supposées intégrées d'ordre 1, le modèle WS-PS doit être estimé par le biais du VECM (17.15). Au préalable, il faut cependant déterminer le nombre de décalage k du modèle VAR (17.14).

Le nombre de retards est déterminé à partir d'une séquence de tests emboîtés de type ratio de vraisemblance. Le modèle VAR (17.14) est successivement estimé pour un nombre de décalage k compris entre 1 et 6 (valeur maximale déterminée par le critère d'information BIC). À chaque étape, l'hypothèse de nullité des coefficients associés au dernier décalage est testée. En réduisant progressivement l'ordre du VAR, le nombre de retards k^* retenu est donné par le premier décalage pour lequel l'hypothèse nulle est rejetée. Enfin, dans le cas où il existe un retard k inférieur à k^* pour lequel l'hypothèse nulle est rejetée, on effectue un dernier test de l'apport significatif des $(k^* - k)$ retards. Dans cet exemple, on teste concrètement l'apport significatif du 6^e retard par rapport à un VAR d'ordre 5, puis l'apport du 5^e retard contre 4 etc. Cette démarche conduit à estimer à chaque étape deux VAR, puis à construire la statistique suivante à partir des matrices des covariances des résidus

$$\chi_c^2 = (T - h) [\log(\det(\Sigma_0)) - \log(\det(\Sigma_1))]$$

où T est le nombre d'observations, h^{10} un paramètre de correction de la statistique, Σ_0 la matrice des covariances résiduelles sous H_0 (modèle contraint à $k-1$ retards) et Σ_1 la matrice des covariances résiduelles sous H_1 (modèle non contraint à k retards). La statistique calculée suit un Chi deux à 9 degrés de liberté (nombre de contraintes multiplié par le nombre d'équations).

En partant du bas du tableau 17.5, l'hypothèse de nullité de la matrice associée au dernier décalage du modèle ne peut pas être rejetée, jusqu'à un modèle VAR d'ordre 2. L'hypothèse de nullité des coefficients associés au 2^e retard est alors indubitablement rejetée. Un nombre de deux retards est donc choisi pour ce modèle. On note que l'hypothèse nulle d'un VAR à 4 retards est acceptée à 7 % seulement, de sorte que la stratégie du test aurait pu s'interrompre à ce stade et conclure à un VAR d'ordre 5. Une lecture approfondie du tableau est alors nécessaire pour nuancer et conforter le choix du nombre de retards lorsque les statistiques calculées sont proches des seuils théoriques usuels. La validité de l'hypothèse d'un VAR à 5 retards est tout d'abord mise en doute par les résultats des tests effectués pour des retards inférieurs. Les hypothèses de nullité des coefficients associés aux 3^e et 4^e retards sont clairement acceptées, ce qui rend possible la détermination d'un VAR d'ordre inférieur à 5.

¹⁰ h est le nombre de paramètres d'une équation du modèle non contraint, c'est-à-dire le nombre de variables multiplié par le nombre de retards plus la constante.

Tableau 17.5 Résultats des tests emboîtés

Test	Statistique	P-value	Décision
$H_0 : A_2 = A_3 = A_4 = A_5 = A_6 = 0$ (VAR à 1 retard) $H_1 : A_3 = A_4 = A_5 = A_6 = 0$ (VAR à 2 retards)	$\chi_c^2 = 37,02$	$2,61 \times 10^{-5}$	H_1
$H_0 : A_3 = A_4 = A_5 = A_6 = 0$ (VAR à 2 retards) $H_1 : A_4 = A_5 = A_6 = 0$ (VAR à 3 retards)	$\chi_c^2 = 11,51$	0,24	H_0
$H_0 : A_4 = A_5 = A_6 = 0$ (VAR à 3 retards) $H_1 : A_5 = A_6 = 0$ (VAR à 4 retards)	$\chi_c^2 = 8,94$	0,44	H_0
$H_0 : A_5 = A_6 = 0$ (VAR à 4 retards) $H_1 : A_6 = 0$ (VAR à 5 retards)	$\chi_c^2 = 15,90$	0,07	H_0
$H_0 : A_6 = 0$ (VAR à 5 retards) $H_1 : A_6 \neq 0$ (VAR à 6 retards)	$\chi_c^2 = 12,41$	0,19	H_0

↑
S
E
N
S

D
E

L
E
C
T
U
R
E

Enfin, le choix d'un VAR à 2 retards peut être conforté par un dernier test. On teste la nullité des retards 2 à 5.

$$H_0 : A_3 = A_4 = A_5 = 0 \quad (\text{VAR à deux retards})$$

$$H_1 : A_3, A_4 \text{ et } A_5 \text{ non nulles simultanément}$$

La statistique suit cette fois-ci un Chi deux à 27 degrés de liberté (produit des 3 retards avec 3 variables et 3 équations). Elle vaut 35.07. Le niveau de significativité étant de 14 %, on ne rejette pas H_0 . Le nombre de retards finalement retenu est de 2.

Analyse résiduelle

L'analyse des résidus fournit *in fine* une indication de la pertinence de la spécification adoptée. Elle peut être conduite simplement à partir de **Cats in Rats**. La commande **@cats** permet de lancer ce programme. On estime le VECM (17.15) avec un nombre de décalage égal à 2. Une rubrique **Cats** apparaît alors dans la barre de menu de **Winrats**. L'étude des résidus est effectuée en choisissant l'option « **residual analysis** » du menu déroulant **Cats**.

Apparaissent alors trois séries de résultats

- Le premier tableau présente la matrice des corrélations des résidus ainsi que les écarts types.
- Le deuxième fournit des statistiques sur l'ensemble du VECM, des critères d'information (utiles dans la comparaison de modèles concurrents), des tests d'autocorrélation et de normalité.
- Le troisième tableau donne les propriétés des résidus pour chaque équation (dans l'ordre W, Y, U).

Les commentaires sont centrés uniquement sur les tests d'autocorrélation et de normalité des résidus. Le logiciel réalise trois types de test d'autocorrélation. Le premier test est le test de Ljung-Box (L-B) d'ordre $T/4$. La statistique suit un Chi deux à $N^2[(T/4) - k + 1] - Nr$ degrés de liberté, où N est le nombre de variables, k est le nombre de retards introduit dans le VAR, r le nombre de relations de cointégration (ici égal à 3 puisque le test du rang de cointégration n'a pas encore été effectué). Les deux autres tests sont des tests du multiplicateur de Lagrange (LM) respectivement d'ordre 1 et 4. Les statistiques de ces tests suivent un Chi deux à N^2 degrés de liberté. Enfin le test de normalité des résidus est asymptotiquement distribué suivant un Chi deux à $2N$ degrés de liberté.

Tableau 17.6 Test d'autocorrélation et de normalité

Test d'autocorrélation		
L-B(18)	CHISQ(144) = 164,060	p-val = 0,12
LM(1)	CHISQ(9) = 17,457	p-val = 0,04
LM(4)	CHISQ(9) = 13,359	p-val = 0,15
Test de normalité	CHISQ(6) = 5,726	p-val = 0,45

Globalement, les résultats statistiques sont satisfaisants. L'autocorrélation des résidus est rejetée dans deux cas et la normalité des résidus n'est pas rejetée à 45 %.

Problème 17.3

Nombre de relations de cointégration

Une fois l'ordre d'intégration et le nombre de retards du VAR déterminés, l'estimation du modèle identifié WS-PS (17.13) passe par la recherche du nombre de relations de cointégration par le biais de l'étude du rang de la matrice Π . Pour être en conformité avec le modèle théorique, on doit avoir deux vecteurs de cointégration correspondant à deux relations stationnaires entre les variables données par (17.13).

Avant de se lancer dans les tests du rang de cointégration, il convient de s'interroger tout d'abord sur l'éventualité de l'introduction de termes déterministes (constante ou trend) dans les relations de cointégration. Les résultats des tests sont en effet conditionnés par la présence de ces variables. **Cats in Rats** distingue quatre cas de figure.

1. L'option **NONE** de la commande `@cats(det =)` spécifie un modèle sans composantes déterministes. Dans la pratique, ce cas fait figure d'exception. Il n'y a pas de constante et pas de trend dans le VECM.
2. L'option, **CIMEAN** restreint la constante dans l'espace de cointégration, c'est-à-dire que la constante n'apparaît que dans les relations de cointégration données par $\beta' Z_{t-1}$ (cf. équation (17.16)). Ceci implique que le modèle exclut la présence d'une tendance dans les données.
3. L'option **DRIFT** inclut une constante dans le modèle non contraint. Il y a une constante dans le VECM indiquant que les variables contiennent un trend linéaire mais le trend n'est pas présent dans la relation de cointégration.
4. Enfin, l'option **CIDRIFT** permet d'introduire un trend dans l'espace de cointégration. Cette option est privilégiée lorsque l'on soupçonne une variable d'être stationnaire en tendance, ou que le trend est justifié dans la relation d'équilibre (cas de WS). Donc, il y a une constante dans le VECM et un trend dans la relation de cointégration.

Un premier test est donc mené en considérant les options **CIMEAN** et **DRIFT**.

Tableau 17.7 Nombre de relations de cointégration (constante restreinte dans l'espace de cointégration)

Eigenv.	L-max	Trace	H0 : r	L-max90	Trace90
0.3821	35.63	77.34	0	14.09	31.88
0.2898	25.33	41.71	1	10.29	17.79
0.1986	16.39	16.39	2	7.50	7.50

Tableau 17.8 Nombre de relations de cointégration (constante non contrainte)

Eigenv.	L-max	Trace	H0 : r	L-max90	Trace90
0.3178	28.29	52.08	0	13.39	26.70
0.1987	16.39	23.78	1	10.60	13.31
0.0950	7.39	7.39	2	2.71	2.71

La première colonne des tableaux 17.7 et 17.8 donne les valeurs propres associées aux vecteurs de cointégration. Les deux colonnes suivantes présentent respectivement les statistiques de la valeur propre maximale et de la trace, tandis que les deux dernières colonnes donnent les valeurs tabulées de ces statistiques pour un niveau de significativité de 10 %. Il s'agit donc de comparer les statistiques calculées aux seuils tabulés. Le tableau 17.9 rappelle la construction des tests des statistiques L-max et Trace tests.

Tableau 17.9 Tests du L-max et de la trace

L-Max		Trace	
H ₀	H ₁	H ₀	H ₁
r=0	r=1	r=0	r=3
r=1	r=2	r=1	r=3
r=2	r=3	r=2	r=3

On s'aperçoit que dans les tableaux 17.7 et 17.8, on a L-max > L-max₉₀ et Trace > Trace₉₀ quel que soit la ligne. L'hypothèse H₀ est donc rejetée à chaque fois. Il paraît clair que ces tests ne sont pas acceptables. Ils concluent en effet à la présence de trois relations de cointégration, ce qui impliquerait la stationnarité des trois variables contrairement aux conclusions des tests de racine unité. Par conséquent, il convient de vérifier si la présence d'un trend dans l'espace de cointégration permet d'aboutir à de meilleures conclusions. Ce trend est justifié en outre par la contrainte d'identification théorique de la courbe WS (équation (17.13)).

Tableau 17.10 Nombre de relations de cointégration (trend contraint dans l'espace cointégré)

Eigenv.	L-max	Trace	H0 : r	L-max90	Trace90
0.3710	34.30	66.15	0	16.13	39.08
0.2511	21.40	31.85	1	12.39	22.95
0.1317	10.45*	10.45*	2	10.56	10.56

*H₀ non rejetée à 10 %

Les tests du rang de l'espace cointégrant conduisent cette fois-ci à retenir deux relations de cointégration. Au total, nous avons pour l'instant un modèle VECM à trois variables, deux retards, deux relations de cointégration et un trend dans l'espace cointégrant. Le système s'écrit donc

$$\begin{aligned}
 \begin{pmatrix} \Delta W \\ \Delta Y \\ \Delta U \end{pmatrix}_t &= \Gamma_1 \begin{pmatrix} \Delta W \\ \Delta Y \\ \Delta U \end{pmatrix}_{t-1} \\
 &+ \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} & \beta_{41} \\ \beta_{12} & \beta_{22} & \beta_{32} & \beta_{42} \end{bmatrix} \begin{pmatrix} W \\ Y \\ U \\ t \end{pmatrix}_{t-1} \\
 &+ \Psi \begin{pmatrix} (1) \\ du80 \\ du82 \\ du83 \end{pmatrix} + \varepsilon_t
 \end{aligned} \tag{17.18}$$

Avant de procéder aux tests d'identification des courbes WS-PS, il convient enfin de conduire deux derniers tests à l'aide de l'option **proc=tsprop**. Le premier test

vérifie l'appartenance des variables à l'espace de cointégration. L'acceptation de l'hypothèse nulle d'exclusion d'une variable impliquerait théoriquement le rejet de celle-ci des 2 relations de cointégration. Ce test est particulièrement utile dans la mesure où les contraintes d'identification du modèle supposent précisément l'existence des variables contraintes. Le deuxième teste l'exogénéité faible des variables. Une variable faiblement exogène n'est pas affectée dans ce cas par le mécanisme correcteur d'erreur exercé par les relations de cointégration. Il est alors possible, sans perte d'information, de réduire le modèle initial en un modèle VECM conditionné par la variable exogène communément appelé VECM partiel. Les résultats sont présentés dans le tableau 17.11, pour un risque de première espèce de 5 %.

Tableau 17.11 *Test d'exclusion et d'exogénéité faible*

Test d'exclusion : LR TEST CHISQ(r)

r	DGF	CHISQ_5 %	W	Y	U	TREND
1	1	3.84	0.07	3.39	11.54	6.01
2	2	5.99	1.83*	12.28	21.53	11.01

Test d'exogénéité faible : LR TEST CHISQ(r)

r	DGF	CHISQ_5 %	W	Y	U
1	1	3.84	7.05	1.12	8.22
2	2	5.99	13.58	4.61*	16.33

*H₀ non rejetée à 5 %

La première colonne du tableau 17.11 donne le nombre de relations de cointégration, la deuxième le degré de liberté de la statistique du Chi deux, la troisième les valeurs tabulées à 5 % puis les statistiques calculées pour chaque variable. Dans la mesure où deux relations cointégrantes sont retenues, seule la deuxième ligne des tableaux est examinée ($r = 2$).

Hormis le salaire réel, l'exclusion des variables est refusée pour toutes les variables, ce qui permet d'appliquer les contraintes identifiantes du modèle. L'acceptation de l'hypothèse nulle pour le salaire réel est présente également dans les travaux de L'Horty et Sobczak, et semble imputable à l'inclusion du trend dans la relation de cointégration. Le lecteur intéressé peut vérifier en effet le refus de l'hypothèse d'exclusion de cette variable en réitérant le test sans trend. Cette variable est cependant conservée pour poursuivre l'estimation du modèle WS-PS.

Enfin, le dernier test indique que la productivité peut être considérée comme faiblement exogène, conformément au modèle théorique. Un modèle VECM partiel composé de deux variables endogènes (les salaires réels et le chômage) et conditionné par une variable exogène, la productivité, pourrait donc être estimé. Ceci n'implique pas néanmoins l'exclusion de la productivité des relations de cointégration. Ce point sera considéré dans le problème 17.5.

Problème 17.4

Identification et contraintes linéaires sur les relations de cointégration

Identification de WS et PS

Les étapes préalables permettent à présent de se consacrer à l'estimation de la forme identifiée du modèle WS-PS. L'exécution de la commande **@cats** affiche une fois de plus le menu **Cats** dans la barre de menu de **Winrats**. Dans ce menu on choisit la rubrique **Rank of Pi**. La boîte de dialogue demande le nombre de relations de cointégration. Taper 2. Puis le logiciel demande à partir de quelle variable il doit normaliser les vecteurs de cointégration. Taper 1 à chaque fois afin de normaliser les vecteurs par rapport au salaire réel. L'écran donne les résultats présentés dans le tableau 17.12. La première matrice fournit les coefficients de la transposée de la matrice β . La seconde matrice donne les valeurs estimées des vitesses d'ajustement (matrice α) et les statistiques de significativité correspondantes. Enfin, la troisième fournit les coefficients de la matrice $\Pi = \alpha\beta'$, qui indiquent dans quelle proportion les variables en niveau W , Y , U et t affectent la dynamique des variables du système ΔW , ΔY , ΔU .

Tableau 17.12 Identification de WS-PS — étape 1

BETA (transposed)				
W	Y	U	TREND	
1	7.043	32.254	- 0.071	
1	- 1.645	- 1.649	0.008	

ALPHA			T-Values for ALPHA	
ΔW	- 0.007	- 0.03	- 4.156	- 2.553
ΔY	0.008	0.10	1.440	2.503
ΔU	- 0.004	0.014	- 4.752	2.592

PI				
	W	Y	U	TREND
ΔW	- 0.037	- 0.001	- 0.182	0
ΔY	0.108	- 0.105	0.106	0
ΔU	0.010	- 0.048	- 0.139	0

T-Values for PI				
	- 3.127	- 0.037	- 3.080	1.654
	2.685	- 1.359	0.533	0.437
	1.877	- 4.716	- 5.342	5.331

Il convient de noter au préalable que les deux relations de cointégration présentées n'ont pas de sens économique précis. Il s'agit uniquement de deux vecteurs de l'espace de cointégration. Il existe donc une infinité de combinaisons linéaires de ces

deux vecteurs qui constituent également des relations de cointégration. Par conséquent, il faut être également prudent sur l'interprétation des statistiques associées, tant que les vecteurs n'ont pas une signification économique. Néanmoins, l'analyse de ces premiers vecteurs et des vitesses d'ajustement α fournit souvent une information précieuse. Le tableau 17.12 donne deux relations de cointégration

$$\begin{aligned} W + 7.034Y + 32.254U - 0.071t &= 0 \\ W - 1.645Y - 1.649U + 0.008t &= 0 \end{aligned} \quad (17.19)$$

D'emblée, on peut remarquer dans la deuxième relation de cointégration que les coefficients de la productivité et du chômage sont de même signe et très proches. Par analogie au modèle théorique, cette deuxième relation ressemble fortement à PS. Par ailleurs le coefficient du chômage dans la première relation est de signe inversé, ce qui suggère probablement l'existence de WS.

À ce stade, le modèle n'est pas identifié puisque aucune des variables d'une relation n'est exclue de l'autre. Dans cette optique les contraintes d'identification théoriques présentées auparavant (équation (17.13)) sont reprises. WS est identifiée en supprimant nul le coefficient associé à la productivité. PS est identifiée en posant l'égalité des coefficients de productivité et de chômage. Comme le soulignent L'Horty et Sobsczak, il ne s'agit pas d'un test mais uniquement de contraintes théoriques destinées à l'identification des deux équations structurelles PS et WS. La démarche requise par le logiciel pour la construction des tests ou l'imposition de contraintes identifiantes est toutefois similaire. Elle consiste à définir des matrices H_1 et H_2 de contraintes sur les coefficients.

Formellement les contraintes s'écrivent

$$\beta = (H_1\varphi_1, H_2\varphi_2) \quad (17.20)$$

où H_1 et H_2 sont des matrices de taille respective (N, s_1) , (N, s_2) . s_1 et s_2 sont le nombre de paramètres non contraints dans la première et la deuxième relation de cointégration. Les matrices H_1 et H_2 , de paramètres fixés par l'utilisateur, traduisent les contraintes d'identification du modèle. φ_1 et φ_2 , de tailles (s_1, r_1) et (s_2, r_2) sont des matrices de coefficients à estimer¹¹.

Ainsi pour le premier vecteur on a

$$\begin{bmatrix} \beta_{11} \\ \beta_{21} \\ \beta_{31} \\ \beta_{41} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{H_1} \underbrace{\begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix}}_{\varphi_1} = \begin{bmatrix} a_1 \\ 0 \\ b_1 \\ c_1 \end{bmatrix} \quad (17.21)$$

¹¹ r_1 et r_2 sont le nombre de vecteurs contraints par les matrices H_1 et H_2 respectivement. On a donc ici $r_1 = r_2 = 1$ et $r_1 + r_2 = r$, où r est le nombre de relations de cointégration ($r = 2$).

En normalisant par rapport aux salaires réels, on obtient

$$\begin{bmatrix} \beta_{11} \\ \beta_{21} \\ \beta_{31} \\ \beta_{41} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ b_1/a_1 \\ c_1/a_1 \end{bmatrix} \quad (17.22)$$

Cette hypothèse économique correspond bien à la contrainte d'identification de WS ($a_{21} = 0$) dans l'équation (17.13).

De même, pour le deuxième vecteur, on a

$$\begin{bmatrix} \beta_{12} \\ \beta_{22} \\ \beta_{32} \\ \beta_{42} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{H_2} \underbrace{\begin{bmatrix} a_2 \\ b_2 \\ c_2 \end{bmatrix}}_{\varphi_2} = \begin{bmatrix} a_2 \\ b_2 \\ b_2 \\ c_2 \end{bmatrix} \quad (17.23)$$

et en normalisant par rapport aux salaires réels, on obtient

$$\begin{bmatrix} \beta_{12} \\ \beta_{22} \\ \beta_{32} \\ \beta_{42} \end{bmatrix} = \begin{bmatrix} 1 \\ b_2/a_2 \\ b_2/a_2 \\ c_2/a_2 \end{bmatrix} \quad (17.24)$$

La contrainte d'égalité des coefficients de productivité et de chômage ($a_{22} = a_{32} = b_2/a_2$) est imposée afin d'obtenir la relation PS dans l'équation (17.13).

Le modèle identifié est obtenu en procédant comme suit¹².

Étape 1: Choisir l'option **Restrictions on each of the betavectors** du menu **Cats**.

Étape 2: Le logiciel demande le nombre de restrictions pour le premier vecteur. Taper 1, puis remplir la matrice $H' (= H'_1)$ de la manière suivante

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

La deuxième colonne correspond à la contrainte de nullité du coefficient de productivité dans WS.

¹² À ce stade lorsqu'on relance la commande **cats** dans **Rats** il faut indiquer à nouveau, avec la rubrique **Rank of Pi**, le nombre de relations de cointégration. La boîte de dialogue demande le nombre de relations de cointégration. Taper 2. Puis le logiciel demande à partir de quelle variable il doit normaliser les vecteurs de cointégration. Taper 1 à chaque fois afin de normaliser les vecteurs par rapport au salaire réel.

Étape 3: L'identification de PS est assurée en imposant une restriction. La matrice H' ($= H_2$) se présente dans ce cas sous la forme

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Les résultats sont présentés dans le tableau 17.13.

Tableau 17.13 Identification de WS-PS – étape 2

BETA (transposed)				
W	Y	U	TREND	
1	0	4.777	-0.007	
1	-1.644	-1.644	0.008	
Standard Errors for BETA (transposed)				
0	0	0.765	0.001	
0	0.33	0.33	0.002	
ALPHA			T-Values for ALPHA	
ΔW	-0.038	0.001	-4.153	0.037
ΔY	0.044	0.064	1.437	1.359
ΔU	-0.019	0.029	-4.755	4.716
PI				
	W	Y	U	TREND
ΔW	-0.037	-0.001	-0.182	0
ΔY	0.108	-0.105	0.106	0
ΔU	0.010	-0.048	-0.139	0
T-Values for PI				
	-3.127	-0.037	-3.080	1.654
	2.685	-1.359	0.533	0.437
	1.877	-4.716	-5.342	5.331

Les signes des coefficients β sont conformes à la théorie et toutes les variables sont significatives (écarts types entre parenthèses). Les deux relations de cointégration sont

$$W + 4.777U - 0.007t = 0 \quad (\text{WS})$$

$$W - 1.644U - 1.644Y + 0.008t = 0 \quad (\text{PS})$$

Le chômage affecte négativement les salaires réels dans WS conformément à la théorie des négociations salariales. Le chômage réduit le pouvoir de négociation syndical et

modère de fait les revendications salariales. Par ailleurs, l'élasticité de la demande de travail mesurée par $-\frac{\partial U}{\partial W} = -\frac{1}{1.64}$ ¹³ dans PS est proche de -0.6.

Les valeurs des vitesses d'ajustement (α) ainsi que les statistiques calculées (T-values) sont également conformes aux valeurs attendues. Théoriquement en effet, on doit avoir un signe négatif de la vitesse d'ajustement associée à la relation WS dans la première équation du VECM. Ceci indique qu'une valeur du salaire réel différente de sa cible de long terme implique une variation des salaires réels à la période suivante de sens opposé afin de converger vers l'équilibre. De même, le mécanisme correcteur d'erreurs associé à la relation PS dans la troisième équation du VECM doit agir négativement sur la dynamique du taux de chômage. On note que seule la première relation de cointégration est significative dans l'équation de salaire et agit bien comme une force de rappel sur la dynamique de court terme des salaires réels. Toutes choses égales par ailleurs, une hausse des salaires réels crée une situation de déséquilibre qui ramène à la baisse les salaires réels les périodes suivantes (vitesse d'ajustement négative égale à -0.038 , avec une T-value de -4.153). En revanche, aucune des relations structurelles n'est significative dans l'équation de productivité, ce qui confirme l'hypothèse d'exogénéité faible de la productivité. Enfin, les deux relations sont significatives dans l'équation de chômage. La relation PS agit bien également comme une force de rappel. Le lecteur ne doit pas être trompé par le signe positif associé à ce mécanisme correcteur d'erreurs ($\alpha_{32} = 0.029$). La normalisation par rapport à U du deuxième vecteur de cointégration aurait amené en effet à un signe négatif de ce coefficient. Cela signifie qu'un déséquilibre sur le taux de chômage tend à se résorber dans le temps. En revanche, le fait qu'un déséquilibre sur les salaires ait un effet négatif sur le chômage est plus difficile à justifier. Il signifie en effet qu'un niveau de salaire réel supérieur à son niveau de long terme implique une baisse du chômage.

Test de restriction sur les coefficients

Un dernier point doit être traité : le coefficient de productivité dans PS est élevé et implique à long terme une tendance croissante de la part des salaires dans la valeur ajoutée, ce qui est peu réaliste. L'hypothèse d'homogénéité des salaires réels à la productivité est donc testée, tout en conservant les contraintes identifiantes. Le test s'écrit sous la forme de l'équation (17.20). La contrainte d'identification de WS ($a_{21} = 0$) est donnée par l'équation (17.22) tandis que les contraintes sur PS sont telles que

$$\begin{bmatrix} \beta_{12} \\ \beta_{22} \\ \beta_{32} \\ \beta_{42} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 \\ -1 & 0 \\ -1 & 0 \\ 0 & 1 \end{bmatrix}}_{H_2} \underbrace{\begin{bmatrix} a_3 \\ b_3 \end{bmatrix}}_{\varphi_2} = \begin{bmatrix} a_3 \\ -a_3 \\ -a_3 \\ b_3 \end{bmatrix} \quad (17.25)$$

et en normalisant par rapport aux salaires réels, on obtient

$$\begin{bmatrix} \beta_{12} \\ \beta_{22} \\ \beta_{32} \\ \beta_{42} \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ -1 \\ b_3/a_3 \end{bmatrix} \quad (17.26)$$

¹³ Ce résultat vient de la relation $\log L - \log N = \log(1 - U) \simeq -U$.

H_2 exprimant les contraintes d'identification de PS et l'hypothèse d'homogénéité des salaires réels à la productivité ($a_{22} = a_{32} = -1$)

Ce test est effectué en reprenant la procédure utilisée pour l'identification du modèle. Les étapes 1 et 2 sont identiques.

Étape 1 : Choisir l'option **Restrictions on each of the betavectors** du menu Cats.

Étape 2 : Le logiciel demande le nombre de restrictions pour le premier vecteur. Taper 1, puis remplir la matrice $H' (= H'_1)$ de la manière suivante

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

La deuxième colonne correspond à la contrainte de nullité du coefficient de productivité dans WS (on impose $a_{21} = 0$).

Étape 3 : Le test sur la relation PS est assuré en imposant maintenant 2 restrictions, la matrice $H' (= H'_2)$ se présente dans ce cas sous la forme

$$\begin{bmatrix} 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

On impose l'égalité des coefficients associés à la productivité, au chômage et au salaire réel ($a_{22} = a_{32} = -1$). Dans ce cas, le ratio de vraisemblance suit un Chi deux à $(N - r)r_2$ degré de liberté¹⁴.

Les résultats du tests ainsi que l'estimation du modèle contraint sont présentés dans le tableau 17.14.

Tableau 17.14 Test de restriction et estimation du modèle contraint

LR Test	CHISQR(1) = 0.30			p-value = 0.59
BETA (transposed)				
W	Y	U	TREND	
1	0	3.08	- 0.005	
1	- 1	- 1	0.004	
Standard Errors for BETA (transposed)				
0	0	0.484	0.001	
0	0	0	0	
ALPHA			T-Values for ALPHA	
ΔW	- 0.056	0.003	- 3.973	0.157
ΔY	0.060	0.066	1.254	0.920
ΔU	- 0.031	0.047	- 4.967	5.011

¹⁴ N est le nombre de variables (3), r est le nombre de relations de cointégrations (2) et r_2 correspond au nombre de vecteurs contraints lorsqu'on teste H'_2 (1).

Le test est accepté largement avec une p-value de 59 %. On ne peut rejeter l'hypothèse de semi-élasticité¹⁵ unitaire des salaires réels au taux de chômage et d'élasticité¹⁶ unitaire des salaires réels à la productivité¹⁷. Côté offre de travail, une hausse du taux de chômage de 1 point entraîne une modération salariale de 3.08 % de la part des syndicats sur longue période. La forme contrainte du modèle WS-PS pour la France est donc

$$\begin{aligned} W + 3.08U - 0.005t &= 0 & (\text{WS}) \\ W - U - Y - 0.004t &= 0 & (\text{PS}) \end{aligned} \quad (17.27)$$

Problème 17.5

Exogénéité faible et VECM partiel

À deux reprises, l'hypothèse d'exogénéité faible de la productivité a été suggérée. Les tests réalisés dans le problème 17.3 ainsi que les valeurs des t de Student associées aux mécanismes correcteurs d'erreurs laissent penser que les relations de long terme n'agissent pas sur la dynamique de la productivité. Il est possible de confirmer cette hypothèse à l'aide d'un dernier test tout en conservant les restrictions effectuées sur les coefficients des équations WS-PS. Il s'agit de tester simplement la significativité des vitesses d'ajustement associées aux mécanismes correcteurs d'erreurs dans l'équation de productivité du VECM. Formellement l'hypothèse jointe suivante est testée

$$\beta = (H_1\varphi_1, H_2\varphi_2) \text{ et } B'\alpha = 0 \quad (17.28)$$

B est une matrice de taille $(N(N-m))$ où $(N-m)$ est le nombre de lignes contraintes. Ici, $N = 3$ et $(N-m) = 1$. Il faut donc définir la matrice B afin de tester la nullité de la deuxième ligne de la matrice α . On a

$$B'\alpha = 0 \iff \underbrace{\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}}_{B'} \underbrace{\begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{bmatrix}}_{\alpha} = [\alpha_{21} \quad \alpha_{22}] = [0 \quad 0] \quad (17.29)$$

¹⁵ Dans une relation de la forme $\log y = \alpha x$, α représente la semi-élasticité. Lorsque la variable x est une variable en taux, comme le taux de chômage, la relation est interprétée de la manière suivante : si x augmente de un point de pourcentage, y varie de $\alpha\%$.

¹⁶ Dans une relation de la forme $\log y = \alpha \log x$, α représente l'élasticité. La relation est interprétée de la manière suivante : si x augmente de 1 %, y varie de $\alpha\%$.

¹⁷ Ceci implique une élasticité de substitution unitaire d'après la formation de PS ($a_{22} = a_{32} = \frac{1}{\sigma} = 1$).

L'exogénéité faible est testée en procédant comme suit.

Étape 1 : Choisir l'option **Restrictions on alpha** du menu **Cats**¹⁸.

Étape 2 : Le logiciel demande le nombre de restrictions. Taper 1, puis remplir la matrice B' telle qu'elle est définie dans l'équation (17.29).

Étape 3 : Trois options sont proposées, selon que le test est réalisé en conservant ou non des restrictions sur les vecteurs de cointégration. Choisir ici **Current restrictions**.

Les résultats sont présentés dans le tableau 17.15.

Tableau 17.15 Test d'exogénéité faible de la productivité

LR Test	CHISQR(3) = 5.04			p-value = 0.17
BETA (transposed)				
W	Y	U	TREND	
1	0	2.426	-0.004	
1	-1	-1	0.004	
Standard Errors for BETA (transposed)				
0	0	0.408	0.001	
0	0	0	0	
ALPHA			T-Values for ALPHA	
ΔW	-0.064	0.007	-3.882	0.343
ΔY	0	0	0	0
ΔU	-0.036	0.051	-4.843	5.342
PI				
	W	Y	U	TREND
ΔW	-0.057	-0.007	-0.164	0
ΔY	0	0	0	0
ΔU	0.015	-0.051	-0.139	0
T-Values for PI				
	-4.018	-0.343	-2.838	2.033
	NA	NA	NA	NA
	2.386	-5.566	-5.561	5.715

L'hypothèse d'exogénéité faible, jointe aux tests sur les β n'est pas rejetée. Le niveau de significativité du test est de 17 %¹⁹.

¹⁸ Au préalable, dans le menu **Cats** il faut refaire l'option **Rank of PI** puis l'option **Restrictions on each of the betavectors** (on impose 1 contrainte pour WS et 2 contraintes pour PS).

¹⁹ Comme le note Harris, R., 1995, *Using Cointegration Analysis in Econometric Modelling*, Prentice Hall, le calcul du degré de liberté dans le cas d'un test joint n'est pas évident. Néanmoins, il correspond ici à la somme des degrés de libertés des tests individuels sur les β et sur l'exogénéité faible réalisés au problème 17.3

Il est donc possible, sur la base de ce résultat, de réestimer le modèle VECM initial sous la forme d'un modèle VECM partiel conditionné par la productivité. Le VECM (équation(17.18)) devient donc

$$\begin{aligned} \begin{pmatrix} \Delta W \\ \Delta U \end{pmatrix}_t &= \Gamma_0 \Delta Y_t + \Gamma_1 \begin{pmatrix} \Delta W \\ \Delta Y \\ \Delta U \end{pmatrix}_{t-1} + \\ &\begin{bmatrix} \alpha_{11} & \alpha_{12} \\ 0 & 0 \\ \alpha_{31} & \alpha_{32} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} & \beta_{41} \\ \beta_{12} & \beta_{22} & \beta_{32} & \beta_{42} \end{bmatrix} \begin{pmatrix} W \\ Y \\ U \\ t \end{pmatrix}_{t-1} + \\ &\Psi \begin{pmatrix} (1) \\ du80 \\ du82 \\ du83 \end{pmatrix} + \varepsilon_t \end{aligned} \quad (17.30)$$

Ce système est estimé aisément en précisant l'option **exo** de la commande @cats et en renouvelant les opérations des contraintes d'identification sur les deux vecteurs de cointégration.

Attention : dans le programme rats, l'ordre des séries est inversé car la variable Y est considérée faiblement exogène. L'ordre affiché dans les tableaux de résultats rats est le suivant : $W U Y TREN D$. De ce fait, pour retrouver les résultats du tableau 17.15 à partir de l'option **exo** on doit choisir la rubrique **Rank of Pi**. La boîte de dialogue demande le nombre de relations de cointégration. Taper 2. Puis le logiciel demande à partir de quelle variable il doit normaliser les vecteurs de cointégration. Taper 1 à chaque fois afin de normaliser les vecteurs par rapport au salaire réel. Ensuite il faut imposer les contraintes comme suit

Étape 1 : Choisir l'option **Restrictions on each of the betavectors** du menu Cats.

Étape 2 : Le logiciel demande le nombre de restrictions pour le premier vecteur. Taper 1, puis remplir la matrice $H' (= H'_1)$ de la manière suivante

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

par rapport à l'analyse précédente les colonnes 2 et 3 sont interverties en raison du changement de place des variables U et Y

Étape 3 : Le test sur la relation PS est assuré en imposant 2 restrictions, la matrice $H' (= H'_2)$ se présente dans ce cas sous la forme

$$\begin{bmatrix} 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Les colonnes 2 et 3 étant identiques, on peut donc reprendre la même matrice H'_2 .

Ces étapes aboutissent finalement à l'estimation du système WS-PS (entre les parenthèses figurent les écarts types)²⁰.

$$\begin{aligned} W + 2.426U - 0.004t &= 0 & \text{(WS)} \\ W - U - Y - 0.004t &= 0 & \text{(PS)} \end{aligned} \quad (17.31)$$

Nous retrouvons des résultats proches des travaux de L'Horty et Sobczak qui trouvent une semi-élasticité des salaires réels au taux de chômage de -2.2 %. L'hypothèse d'égalité à cette valeur du coefficient du chômage ne peut d'ailleurs pas être rejetée par un test de Wald. Il est en effet possible de construire un tel test à partir des écarts types estimés de la matrice β (« standard errors for BETA ») en calculant la statistique : $\left(\frac{2.426 - 2.2}{0.437}\right)^2 = 0.27$. L'hypothèse nulle ne peut être rejetée. Le niveau de significativité est en effet de 60 %.

Enfin, en choisissant l'option **Short-run matrices** du menu Cats, on retrouve les paramètres des matrices Γ_0 , Γ_1 et Ψ (Tstat entre parenthèses), la matrice α étant obtenue en même temps que la matrice des coefficients BETA.

$$\Gamma_0 = \begin{bmatrix} \overbrace{0.015}^{\Delta Y_t} \\ (0.449) \\ -0.035 \\ (-2.458) \end{bmatrix}, \quad \Gamma_1 = \begin{bmatrix} \overbrace{0.103}^{\Delta W_{t-1}} & \overbrace{-0.017}^{\Delta Y_{t-1}} & \overbrace{-0.135}^{\Delta U_{t-1}} \\ (1.109) & (-0.437) & (-0.701) \\ -0.043 & 0.034 & 0.421 \\ (-1.068) & (2.233) & (5.057) \end{bmatrix}$$

$$\alpha = \begin{bmatrix} -0.064 & 0.007 \\ (-3.84) & (0.34) \\ -0.035 & 0.051 \\ (-5.036) & (5.55) \end{bmatrix}, \quad \Psi = \begin{bmatrix} (1) & du80 & du82 & du83 \\ 0.002 & 0.003 & 0.01 & -0.012 \\ (0.336) & (2.57) & (4.024) & (-3.034) \\ -0.011 & 0 & -0.001 & -0.004 \\ (-5.429) & (-0.26) & (-0.502) & (-2.897) \end{bmatrix}$$

Globalement, la spécification de la dynamique de court terme n'est pas très satisfaisante. Les variables sont faiblement significatives, en particulier dans l'équation de salaire. Ce résultat n'a rien de surprenant dans la mesure où l'estimation des relations de long terme entre les variables a été privilégiée, au détriment de la spécification de la dynamique de court terme. L'estimation du modèle WS-PS passe en effet par l'étude des relations entre les salaires réels et le taux de chômage, et suppose donc une indexation unitaire des salaires nominaux aux prix. Si cette propriété peut être vérifiée à long terme, des rigidités nominales liées à la nature des contrats salariaux ou à de mauvaises anticipations des agents peuvent à court terme infléchir cette homogénéité des salaires aux prix. Ainsi, la modélisation des rigidités nominales permettrait d'enrichir l'analyse de la formation des prix et des salaires nominaux à court terme et d'améliorer la dynamique du VECM.

²⁰ Les résultats du système sont extraits des tableaux de résultats rats BETA (transposed) et Standard Errors for BETA (transposed).

18

ÉCONOMÉTRIE DES VARIABLES QUALITATIVES

1. Modèles à variables qualitatives binaires
2. Modèles à variables dépendantes limitées

OBJECTIFS DU CHAPITRE

Introduire l'économétrie des variables qualitatives

- Modèles de choix binaire : probit et logit
- Modèle de régression tronqué
- Modèle de régression censuré (tobit)
- Modèle de sélection de l'échantillon

NIVEAU

MASTER, première et deuxième années

Dans ce chapitre on présente deux types de modèles relevant de l'économétrie des variables qualitatives et très souvent utilisés : les modèles à variables qualitatives binaires (probit, logit) et les modèles à variables dépendantes limitées. Les premiers cherchent à modéliser une décision représentée par une variable qualitative binaire. Les seconds cherchent à estimer un modèle dont les données sont tronquées ou censurées (tobit).

1. MODÈLES À VARIABLES QUALITATIVES BINAIRES

1.1 Présentation des modèles à variables qualitatives binaires

Dans ces modèles on cherche à modéliser le comportement d'agents. La variable expliquée est une variable binaire qui peut prendre la valeur 0 ou 1.

Exemple 18.1

Soit un échantillon composé de 1000 étudiants ayant une licence. On cherche à modéliser la probabilité qu'un individu poursuive ses études ou pas. On observe pour les 1000 individus la variable y

$$\begin{aligned} y_i &= 1 \text{ si l'individu } i \text{ choisit de poursuivre ses études} \\ y_i &= 0 \text{ sinon} \end{aligned} \quad (18.1)$$

Soit U_{i1} l'utilité de l'individu i s'il choisit de poursuivre ses études et U_{i0} son utilité sinon. Ces niveaux d'utilité ne sont pas directement observés, ils dépendent des caractéristiques socio-économiques telles que le sexe, l'âge, le lieu de résidence et des caractéristiques spécifiques à chaque alternative. Pour simplifier, supposons que les niveaux d'utilité U_{i1} et U_{i0} sont expliqués par un modèle linéaire, les variables explicatives sont x_{i1} et x_{i2} , deux variables socio-économiques

$$\begin{aligned} U_{i1} &= a_1 + a_{11}x_{i1} + a_{21}x_{i2} + \varepsilon_{i1} \\ U_{i0} &= a_0 + a_{10}x_{i1} + a_{20}x_{i2} + \varepsilon_{i0} \end{aligned} \quad (18.2)$$

ε_{i1} et ε_{i0} sont des aléas.

Les niveaux d'utilité ne sont pas des variables observables, et l'écart entre U_{i0} et U_{i1} est une variable latente¹. La règle de décision est la suivante : l'individu choisit de poursuivre ses études si $U_{i0} < U_{i1}$.

Soit P_i la probabilité que l'individu i choisisse de poursuivre ses études

$$\begin{aligned} P_i &= Pr(y_i = 1) = Pr(U_{i0} < U_{i1}) \\ &= Pr(\varepsilon_{i0} - \varepsilon_{i1} < (a_1 - a_0) + x_{i1}(a_{11} - a_{10}) + x_{i2}(a_{21} - a_{20})) \end{aligned}$$

¹ Une variable latente est une variable en fonction de laquelle la décision est prise, elle peut être non observée, comme le niveau d'utilité $U_{i0} - U_{i1}$.

$$P_i = Pr(\varepsilon_i \leq x'_i \beta) = F(x'_i \beta) \quad (18.3)$$

avec F la fonction de répartition, ε_i un aléa égal à $\varepsilon_{i0} - \varepsilon_{i1}$, $x'_i = (1 \quad x_{i1} \quad x_{i2})$ le vecteur des variables explicatives et β le vecteur des paramètres à estimer

$$\beta = \begin{pmatrix} a_1 - a_0 \\ a_{11} - a_{10} \\ a_{21} - a_{20} \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

De manière générale dans un modèle de choix binaire, on considère des individus qui ont le choix entre deux alternatives et on cherche à modéliser la probabilité qu'un individu choisit une des alternatives. On observe une décision représentée par la variable y_i qui peut prendre deux valeurs 0 ou 1 et on note P_i la probabilité que y_i soit égale à 1.

$$P_i = Pr(y_i = 1) = F(x'_i \beta) \quad (18.4)$$

Une probabilité est comprise entre 0 et 1, par conséquent le modèle doit vérifier les deux conditions suivantes

$$\begin{aligned} \lim_{x'_i \beta \rightarrow +\infty} Pr(y_i = 1) &= 1 \\ \lim_{x'_i \beta \rightarrow -\infty} Pr(y_i = 1) &= 0 \end{aligned} \quad (18.5)$$

La fonction F est une fonction de répartition et on note f sa fonction de densité. Dans la pratique, on utilise souvent deux spécifications qui conduisent à des résultats très proches : le modèle probit et le modèle logit.

1.2 Les modèles probit et logit

Dans un modèle probit la probabilité P_i est définie de la manière suivante

$$P_i = F(x'_i \beta) = \int_{-\infty}^{x'_i \beta} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-t^2/2}}_{\phi(t)} dt = \Phi(x'_i \beta) \quad (18.6)$$

Dans un modèle probit on suppose que la fonction de répartition F est une loi normale. On note Φ la fonction de répartition et ϕ la fonction de densité de la loi normale centrée réduite $N(0,1)$.

Dans le modèle logit, on suppose que la fonction de répartition est une loi logistique notée Λ avec

$$P_i = F(x'_i\beta) = \frac{e^{x'_i\beta}}{1 + e^{x'_i\beta}} = \frac{1}{1 + e^{-x'_i\beta}} = \Lambda(x'_i\beta) \tag{18.7}$$

la fonction de densité de la loi logistique est donnée par la dérivée de la fonction Λ et elle est égale à $\Lambda(1 - \Lambda)$

$$\Lambda(x'_i\beta)(1 - \Lambda(x'_i\beta)) = \frac{e^{-x'_i\beta}}{(1 + e^{-x'_i\beta})^2} \tag{18.8}$$

De manière générale, on note que les fonctions normale et logistique sont très proches. En effet considérons une variable aléatoire X distribuée selon une loi logistique avec $E(x) = \gamma$ et $var(x) = \frac{(\pi\delta)^2}{3}$

$$X \sim Lo_\lambda \left(\gamma, \frac{(\pi\delta)^2}{3} \right)$$

avec

$$\Lambda(x) = \frac{1}{1 + e^{-\lambda \frac{x - \gamma}{\pi\delta/\sqrt{3}}}}$$

Pour une valeur des paramètres $\lambda = \frac{\pi}{\sqrt{3}} \simeq 1.8$, $\gamma = 0$ et $\delta = 1$, la fonction de répartition logistique s'écrit

$$\Lambda(x) = \frac{1}{1 + e^{-x}}$$

Dans ce cas, les distributions logistique et normale centrée réduite sont très proches (Figure 18.1).

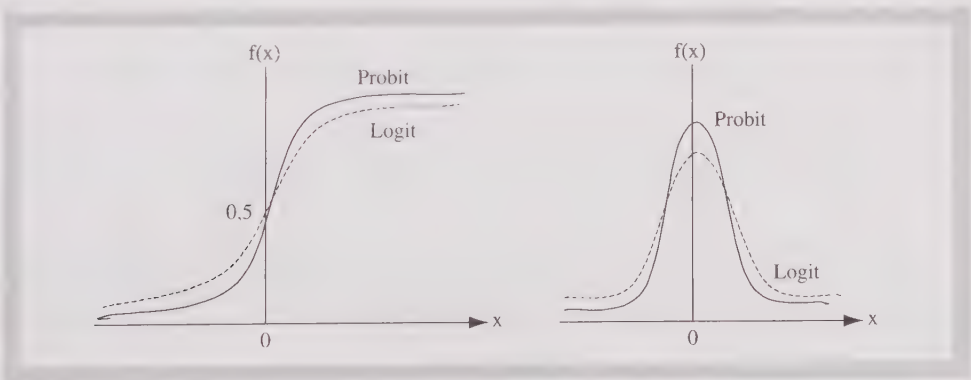


Figure 18.1 Fonctions de répartition et de densité

On note que les deux fonctions de répartition ont une valeur identique pour une valeur de $x = 0$, $\Phi(x) = \Lambda(x) = 0.5$

Donc, on peut représenter la probabilité de réaliser un choix donné soit par une fonction normale (18.6), soit par une fonction logistique (18.7). Les résultats des deux modèles sont très similaires. On observe le choix de l'individu et le vecteur des variables explicatives x_i . Le problème consiste alors à estimer les paramètres β du modèle probit ou logit.

1.3 Estimation des modèles probit et logit

1.3.1 La méthode du maximum de vraisemblance

Pour estimer le vecteur de paramètres β on applique l'estimateur du maximum de vraisemblance. Soit N le nombre d'individus. La fonction de vraisemblance est donnée par

$$L = Pr(Y_1 = y_1, Y_2 = y_2, \dots, Y_N = y_N) \quad (18.9)$$

Les variables aléatoires Y_i étant distribuées de manière indépendantes, la fonction de vraisemblance s'écrit

$$L = \prod_{i=1}^N (P_i)^{y_i} (1 - P_i)^{1-y_i} \quad (18.10)$$

car y_i prend la valeur 1 avec une probabilité P_i et la valeur 0 avec une probabilité $1 - P_i$.

L'expression logarithmique de la fonction de vraisemblance est donnée par

$$\boxed{\text{Log}L = \sum_{i=1}^N y_i \text{Log}P_i + (1 - y_i) \text{Log}(1 - P_i)} \quad (18.11)$$

P_i étant égale à $F(x'_i\beta)$ la fonction devient

$$\text{Log}L = \sum_{i=1}^N y_i \text{Log}F(x'_i\beta) + (1 - y_i) \text{Log}(1 - F(x'_i\beta))$$

Le vecteur des paramètres β est alors obtenu en maximisant la fonction $\text{Log}L$ par rapport à β . Pour obtenir l'estimateur du maximum de vraisemblance $\hat{\beta}_{MV}$ on doit résoudre le système d'équations non linéaire suivant

$$\frac{\partial \text{Log}L}{\partial \beta} = \sum_{i=1}^N \left(\frac{y_i - F(x'_i\beta)}{F(x'_i\beta)(1 - F(x'_i\beta))} \right) f(x'_i\beta) x_i = 0 \quad (18.12)$$

et la matrice de variance-covariance asymptotique peut être estimée en utilisant l'inverse de la matrice Hessienne évaluée en $\widehat{\beta}_{MV}$

$$E \left(-\frac{\partial^2 \text{Log}L}{\partial \beta \partial \beta'} \right)^{-1} = \left(\sum_{i=1}^N \frac{f(x'_i \beta)^2}{F(x'_i \beta) (1 - F(x'_i \beta))} x_i x'_i \right)^{-1} \quad (18.13)$$

L'estimation de β permet d'évaluer la probabilité P_i avec

$$\widehat{P}_i = F(x'_i \widehat{\beta}_{MV}) \quad (18.14)$$

- Dans le modèle probit $F = \Phi$ et $f = \phi$ la condition du premier ordre s'écrit

$$\frac{\partial \text{Log}L}{\partial \beta} = \sum_{y_i=1} \frac{\phi(x'_i \beta)}{\Phi(x'_i \beta)} x_i + \sum_{y_i=0} \frac{-\phi(x'_i \beta)}{1 - \Phi(x'_i \beta)} x_i \quad (18.15)$$

et la matrice de variance-covariance devient

$$E \left(-\frac{\partial^2 \text{Log}L}{\partial \beta \partial \beta'} \right)^{-1} = \left(\sum_{i=1}^N \frac{\phi(x'_i \beta)^2}{\Phi(x'_i \beta) (1 - \Phi(x'_i \beta))} x_i x'_i \right)^{-1} \quad (18.16)$$

La probabilité estimée pour chaque individu est donnée par

$$\widehat{P}_i = \Phi(x'_i \widehat{\beta}_{Probit}) \quad (18.17)$$

avec $\widehat{\beta}_{Probit}$ le vecteur de paramètres du modèle probit estimé avec la méthode du maximum de vraisemblance.

- Dans le modèle logit $F = \Lambda$ et $f = \Lambda(1 - \Lambda)$. Les conditions du premier ordre et la matrice de variance-covariance s'écrivent

$$\frac{\partial \text{Log}L}{\partial \beta} = \sum_{i=1}^N (y_i - \Lambda(x'_i \beta)) x_i \quad (18.18)$$

$$E \left(-\frac{\partial^2 \text{Log}L}{\partial \beta \partial \beta'} \right)^{-1} = \left(\sum_{i=1}^N \Lambda(x'_i \beta) (1 - \Lambda(x'_i \beta)) x_i x'_i \right)^{-1} \quad (18.19)$$

Dans le modèle logit on évalue la probabilité P_i avec

$$\widehat{P}_i = \Lambda(x'_i \widehat{\beta}_{Logit}) \quad (18.20)$$

avec $\widehat{\beta}_{Logit}$ le vecteur de paramètres du modèle logit estimé avec la méthode du maximum de vraisemblance.

1.3.2 Tests statistiques et mesure de la qualité de l'ajustement

Ayant évalué le vecteur de paramètres β et sa matrice de variance-covariance avec un modèle probit ou logit on peut appliquer les tests standards de contraintes sur les paramètres. En particulier si on veut tester la significativité du modèle c'est-à-dire l'hypothèse $H_0 : \beta_2 = \beta_3 = \dots = \beta_K = 0$ dans un modèle comportant K paramètres² on peut appliquer le test du ratio de vraisemblance

$$LR = -2 \left(\text{Log}L(\widehat{\beta}_{MV}^*) - \text{Log}L(\widehat{\beta}_{MV}) \right) \quad (18.21)$$

où $\text{Log}L(\widehat{\beta}_{MV})$ est la valeur de la fonction $\text{Log}L$ lorsque les paramètres β_k , $k = 2, \dots, K$ sont non nuls.

$\text{Log}L(\widehat{\beta}_{MV}^*)$ est la valeur du logarithme de la fonction de vraisemblance sous l'hypothèse H_0

$$\begin{aligned} \text{Log}L(\widehat{\beta}_{MV}^*) &= \sum_{i=1}^N y_i \text{Log}F(\widehat{\beta}_1) + (1 - y_i) \text{Log}(1 - F(\widehat{\beta}_1)) \\ &= \mu (P \text{Log}P + (1 - P) \text{Log}(1 - P)) \end{aligned} \quad (18.22)$$

Sous l'hypothèse H_0 , $P_i = F(\beta_1) = P$. La probabilité P_i est identique pour tous les individus et elle est donnée par la proportion P d'individus pour lesquels on observe $y_i = 1$.

Si la statistique LR est inférieure au $\chi_{\alpha}^2(K - 1)$ on accepte l'hypothèse H_0 .

Par ailleurs, pour mesurer la qualité de l'ajustement, il existe plusieurs possibilités. Une manière simple consiste à calculer le nombre de mauvaises prédictions. L'estimateur du maximum de vraisemblance donne \widehat{P}_i . On suppose que si \widehat{P}_i est inférieure à 0.5 alors la valeur estimée de y_i notée \widehat{y}_i est égale à 0 sinon elle est égale à 1. On peut évaluer ainsi le nombre de mauvaises prédictions en calculant

$$\sum_{i=1}^N (y_i - \widehat{y}_i)^2 \quad (18.23)$$

On peut également utiliser le coefficient de détermination (*pseudo* R^2) suggéré par Mc Fadden R_{MF}^2

$$R_{MF}^2 = 1 - \frac{\text{Log}L(\widehat{\beta}_{MV})}{\text{Log}L(\widehat{\beta}_{MV}^*)} \quad (18.24)$$

² $x_i' \beta = \beta_1 + \beta_2 x_{2i} + \dots + \beta_K x_{Ki}$

1.4 Interprétation des résultats

L'estimation du modèle probit ou logit permet d'évaluer la probabilité P_i et l'influence des différentes variables explicatives sur cette probabilité. Dans le modèle linéaire les paramètres estimés mesurent les effets marginaux, c'est-à-dire l'impact de la variation des variables explicatives sur la variable endogène. Dans les modèles probit et logit les effets marginaux ne sont pas égaux aux valeurs estimées des paramètres.

1.4.1 Les effets marginaux lorsque les variables explicatives sont continues

Pour mesurer l'impact d'une variation d'une variable explicative sur la probabilité p_i on doit calculer

$$\frac{\partial P_i}{\partial x_i} = \beta f(x_i' \beta) \quad (18.25)$$

Si le modèle estimé est un modèle probit

$$\frac{\partial \hat{P}_i}{\partial x_i} = \hat{\beta}_{Probit} \phi(x_i' \hat{\beta}_{Probit}) \quad (18.26)$$

Si c'est un modèle logit

$$\frac{\partial \hat{P}_i}{\partial x_i} = \hat{\beta}_{Logit} \Lambda(x_i' \hat{\beta}_{Logit}) (1 - \Lambda(x_i' \hat{\beta}_{Logit})) \quad (18.27)$$

Ces effets marginaux peuvent être calculés pour chaque individu. Cependant, pour simplifier la présentation des résultats, on évalue souvent ces effets marginaux au point moyen \bar{x} , ou pour un niveau de probabilité initial donné.

- Dans l'exemple 17.1, si $x_{1i} = \bar{x}_1$ et $x_{2i} = \bar{x}_2$, l'impact d'une variation de la variable x_{i1} sur la probabilité de poursuivre ses études est donné par

$$\begin{aligned} \frac{\partial \hat{P}_i}{\partial x_{i1}} &= \hat{\beta}_2 f(\hat{\beta}_1 + \hat{\beta}_2 \bar{x}_1 + \hat{\beta}_3 \bar{x}_2) \\ &= \hat{\beta}_{2Probit} \phi(\bar{x}' \hat{\beta}) \text{ pour le modèle probit} \\ &= \hat{\beta}_{2Logit} \Lambda(\bar{x}' \hat{\beta}) (1 - \Lambda(\bar{x}' \hat{\beta})) \text{ pour le modèle logit} \end{aligned}$$

- Pour un niveau de probabilité initial égal à 0.2 par exemple ($P_i = 0.2$) on mesure l'impact d'une variation de la variable x_{i1} sur la probabilité de poursuivre ses études en procédant de la manière suivante

→ dans le cas du modèle probit on cherche dans un premier temps la valeur de $x'_i\beta$ correspondant à $\Phi(x'_i\beta) = 0.2$ dans la table de la fonction de répartition de la loi normale, on trouve $x'_i\beta = -0.85$. Dans un second temps, on obtient, avec la table de la fonction de densité, $\phi(-0.85) = 0.278$, d'où

$$\frac{\partial \widehat{P}_i}{\partial x_{i1}} = \widehat{\beta}_{2Probit}(0.278)$$

→ dans le cas du modèle logit la procédure est simple $p_i = \Lambda(x'_i\beta) = 0.2$ d'où

$$\frac{\partial \widehat{P}_i}{\partial x_{i1}} = \widehat{\beta}_{2Logit}(0.2)(1 - 0.2)$$

1.4.2 Les effets marginaux lorsque les variables explicatives sont qualitatives

Si dans le modèle (18.1) il y a une variable explicative qualitative, on ne peut pas appliquer la procédure précédente pour mesurer son impact sur la probabilité P_i . Par exemple, si on fait intervenir la variable explicative x_3 indiquant le sexe de l'individu. Cette variable prend deux valeurs

$$\begin{aligned} x_{3i} &= 1 \text{ si l'individu } i \text{ est de sexe masculin} \\ x_{3i} &= 0 \text{ si l'individu } i \text{ est de sexe féminin} \end{aligned}$$

Le modèle s'écrit alors

$$P_i = F(\beta_1 + \beta_2 x_{1i} + \beta_3 x_{2i} + \beta_4 x_{3i}) \quad (18.28)$$

et son estimation avec la méthode du maximum de vraisemblance donne

$$\widehat{P}_i = F(\widehat{\beta}_1 + \widehat{\beta}_2 x_{1i} + \widehat{\beta}_3 x_{2i} + \widehat{\beta}_4 x_{3i}) \quad (18.29)$$

Au point moyen, $x_{1i} = \bar{x}_1$ et $x_{2i} = \bar{x}_2$, pour évaluer l'impact de la variable x_3 sur la probabilité de poursuivre ses études on mesure, dans un premier temps, la probabilité qu'un individu de sexe masculin choisit de poursuivre ses études, on la note \widehat{P}_M . Dans un second temps, on calcule cette probabilité pour un individu de sexe féminin, on la note \widehat{P}_F . Dans le cas d'un modèle probit

$$\begin{aligned} \widehat{P}_M &= \Phi\left(\widehat{\beta}_{1Probit} + \widehat{\beta}_{2Probit}\bar{x}_1 + \widehat{\beta}_{3Probit}\bar{x}_2 + \widehat{\beta}_4\right) \\ \widehat{P}_F &= \Phi\left(\widehat{\beta}_{1Probit} + \widehat{\beta}_{2Probit}\bar{x}_1 + \widehat{\beta}_{3Probit}\bar{x}_2\right) \end{aligned}$$

L'effet du sexe sur le choix de l'individu est alors mesuré par l'écart entre les deux probabilités : $\widehat{P}_M - \widehat{P}_F$. Si le modèle est un modèle logit, on remplace les paramètres $\widehat{\beta}_{Probit}$ par $\widehat{\beta}_{Logit}$ et la fonction Φ par la fonction Λ .

1.5 Comparaison des résultats des modèles probit et logit

Pour modéliser un problème de choix binaire on peut donc utiliser un modèle probit ou logit. Les deux modèles donnent des résultats très proches. Pour comparer les résultats il faut ramener à une échelle identique la valeur des paramètres estimés. Sur la figure 18.1 on a observé que les fonctions de répartition des deux modèles se coupent en $x = 0$ et que $\Phi(0) = \Lambda(0) = 0.5$. Par conséquent, les effets marginaux obtenus dans les modèles (18.26) et (18.27), doivent vérifier la relation suivante en ce point

$$\widehat{\beta}_{Probit} \phi(0) = \widehat{\beta}_{Logit} \Lambda(0)(1 - \Lambda(0))$$

avec $\phi(0) = 0.3889$ et $\Lambda(0)(1 - \Lambda(0)) = 0.25$

D'où la relation suivante entre les coefficients

$$0.25\beta_{kLogit} = 0.3889\beta_{kProbit} \Leftrightarrow \beta_{kLogit} = 1.6\beta_{kProbit} \quad (18.30)$$

La valeur de 1.6 correspond à une évaluation empirique de la relation proposée par Amemiya. Théoriquement la valeur est donnée par le paramètre λ de la fonction logistique égal à environ $1.8 = \frac{\pi}{\sqrt{3}}$.

2. MODÈLES À VARIABLES DÉPENDANTES LIMITÉES

On présente ici des modèles dans lesquels une variable est tronquée ou censurée. Dans les modèles tronqués on sélectionne l'échantillon de données, par exemple on s'intéresse aux revenus d'individus sachant que leurs revenus sont supérieurs à une valeur seuil. Dans le modèle censuré on ne dispose pas de toutes les données. Par exemple, si on veut modéliser la demande de ticket pour le match de football France Italie en finale de coupe d'Europe, on observera le nombre de tickets vendus. Il est égal à la capacité d'accueil du stade de Rotterdam si le nombre de tickets demandés est supérieur à la capacité d'accueil du stade. Il est égal à la demande sinon. Ces deux types de modèles sont présentés successivement.

2.1 Modèle de régression tronqué

Dans un premier temps on présente la distribution d'une variable aléatoire tronquée et ses caractéristiques. Dans un second temps, le modèle de régression tronqué est examiné.

2.1.1 Distribution tronquée

Une distribution tronquée correspond à une partie d'une distribution non tronquée, en dessus ou en dessous d'une valeur spécifiée.

Soit X une variable aléatoire continue de fonction de répartition $F(x)$ et de densité $f(x)$, sa distribution tronquée, pour une troncature vers le bas³, est donnée par

$$f(x/x > a) = \frac{f(x)}{\Pr(x > a)} = \frac{f(x)}{1 - F(a)} \quad (18.31)$$

Dans le cas d'une variable aléatoire normale d'espérance μ et de variance σ^2

$$f(x/x > a) = \frac{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{1 - \Phi\left(\frac{a-\mu}{\sigma}\right)} = \frac{\frac{1}{\sigma}\phi(z)}{1 - \Phi(\alpha)} \quad (18.32)$$

avec $z = \frac{x-\mu}{\sigma}$, $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ et $\alpha = \frac{a-\mu}{\sigma}$. Pour une valeur de $\mu = 0$ et $\sigma = 1$, la distribution normale tronquée vers le bas au point $x = -0.5$ est représentée figure 18.2.

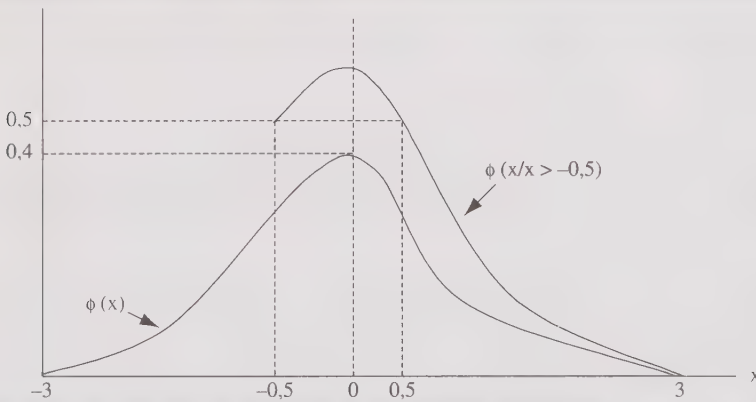


Figure 18.2 Distribution normale tronquée

Les valeurs de la fonction de densité sont données par

$$\begin{aligned} \phi(x/x > -0.5) &= \frac{\phi(x)}{1 - \Phi(-0.5)} = \frac{\phi(x)}{\Phi(0.5)} \\ &= \frac{\phi(x)}{0.6915} \end{aligned}$$

³ Pour une troncature vers le haut

$$f(x/x < a) = \frac{f(x)}{F(a)}$$

Pour une valeur de $x = 0.5$, $\phi(x) = 0.3521$, d'où

$$\rho(0.5/0.5 > -0.5) = \frac{0.3521}{0.6915} \simeq 0.5$$

Les moments d'une variable aléatoire tronquée sont donnés par

$$\begin{aligned} E(x/x > a) &= \int_a^{+\infty} x f(x/x > a) dx \\ \text{Var}(x/x > a) &= \int_a^{+\infty} (x - E(x/x > a))^2 f(x/x > a) dx \end{aligned} \quad (18.33)$$

Dans le cas d'une variable aléatoire distribuée selon une loi normale d'espérance μ et de variance σ^2 , on obtient

$$\boxed{E(x/x > a) = \sigma \frac{\phi(\alpha)}{1 - \Phi(\alpha)} + \mu} \quad (18.34)$$

et

$$\boxed{\text{Var}(x/x > a) = \sigma^2 \left(1 - \left(\left(\frac{\phi(\alpha)}{1 - \Phi(\alpha)} \right)^2 + \frac{\alpha \phi(\alpha)}{1 - \Phi(\alpha)} \right) \right)} \quad (18.35)$$

avec $\alpha = \frac{a - \mu}{\sigma}$.

2.1.2 Le modèle de régression tronqué

Dans le modèle linéaire standard

$$\begin{aligned} y_i &= \beta_1 + \beta_2 x_{2i} + \dots + \beta_K x_{Ki} + \varepsilon_i = x'_i \beta + \varepsilon_i \\ &= E(y_i) + \varepsilon_i \end{aligned} \quad (18.36)$$

Si les aléas sont distribués selon une loi normale avec une espérance nulle et une variance égale à σ^2 alors la variable y_i est distribuée selon une loi normale avec une espérance égale à $x'_i \beta$ et une variance égale à σ^2 .

$$y_i \sim N(x'_i \beta, \sigma^2) \quad (18.37)$$

Le modèle de régression tronqué correspondant au modèle linéaire standard (18.36) s'écrit

$$\boxed{y_i/y_i > a = E(y_i/y_i > a) + u_i} \quad (18.38)$$

où u_i représente un aléa, d'après (18.34)

$$E(y_i/y_i > a) = \sigma \frac{\phi(\alpha_i)}{1 - \Phi(\alpha_i)} + x_i' \beta \quad (18.39)$$

avec $\alpha_i = \frac{a - x_i' \beta}{\sigma}$.

Soit $\lambda_i = \frac{\phi(\alpha_i)}{1 - \Phi(\alpha_i)}$, appelé inverse du ratio de Mills, le problème consiste à évaluer le modèle tronqué suivant

$$y_i/y_i > a = x_i' \beta + \sigma \lambda_i + u_i \quad (18.40)$$

avec $E(u_i) = 0$ et $Var(u_i/u_i > a) = \sigma^2 (1 - \lambda_i^2 + \alpha_i \lambda_i)$ d'après (18.35).

Ce modèle est constitué d'une partie linéaire $x_i' \beta$ et d'une partie non linéaire $\sigma \lambda_i$ par rapport aux variables explicatives.

2.1.3 La fonction de vraisemblance

L'estimateur du maximum de vraisemblance donne des estimateurs convergents asymptotiquement distribués selon une loi normale. La fonction de vraisemblance s'écrit

$$L = \prod_{i=1}^N f(y_i/y_i > a) = \prod_{i=1}^N \frac{\frac{1}{\sigma} \phi\left(\frac{y_i - x_i' \beta}{\sigma}\right)}{1 - \Phi\left(\frac{a - x_i' \beta}{\sigma}\right)} \quad (18.41)$$

et le logarithme de la fonction de vraisemblance est donné par

$$\begin{aligned} \text{Log} L &= \sum_{i=1}^N \text{Log} \left(\frac{1}{\sigma} \phi \left(\frac{y_i - x_i' \beta}{\sigma} \right) \right) - \text{Log} \left(1 - \Phi \left(\frac{a - x_i' \beta}{\sigma} \right) \right) \\ &= -\frac{N}{2} (\text{Log} 2\pi + \text{Log} \sigma^2) - \sum_{i=1}^N \frac{1}{2\sigma^2} (y_i - x_i' \beta)^2 \\ &\quad - \sum_{i=1}^N \text{Log} \left(1 - \Phi \left(\frac{a - x_i' \beta}{\sigma} \right) \right) \end{aligned} \quad (18.42)$$

Les estimateurs sont obtenus en maximisant $\text{Log} L$ par rapport à β et σ^2 . Les valeurs estimées du modèle de régression tronqué (18.40) sont données par

$$\hat{y}_i/y_i > a = x_i' \hat{\beta}_{MV} + \hat{\sigma}_{MV} \hat{\lambda}_i, MV \quad (18.43)$$

2.1.4 Interprétation des résultats

Afin d'interpréter les résultats on calcule les effets marginaux donnés théoriquement par l'expression suivante

$$\frac{\partial E(y_i | y_i > a)}{\partial x_i} = \beta (1 - \lambda_i^2 + \alpha_i \lambda_i) \quad (18.44)$$

avec $\alpha_i = \frac{a - \beta' x_i}{\sigma}$ et $\lambda_i = \frac{\phi(\alpha_i)}{1 - \Phi(\alpha_i)}$

2.2 Modèle de régression censuré — Modèle Tobit

Comme pour le modèle de régression tronqué, on présente dans un premier temps la distribution d'une variable censurée et dans un second temps le modèle de régression (le modèle Tobit).

2.2.1 Distribution d'une variable censurée

On considère une variable aléatoire distribuée selon un loi normale avec une espérance μ et une variance σ^2 .

Par exemple, soit y^* le nombre de tickets demandés pour un match de football

$$y^* \sim N(\mu, \sigma^2)$$

On note y le nombre de tickets vendus et c le nombre de places disponibles. On observe la relation suivante

$$\begin{cases} y = y^* & \text{si } y^* \leq c \\ y = c & \text{si } y^* > c \end{cases} \quad (18.45)$$

La contrainte de capacité donnée par c correspond à un point de censure.

La variable y est une variable censurée dont la distribution est composée de deux parties : une partie continue et une partie discrète. La partie continue de la distribution de y est donnée par la distribution de la variable y^* lorsque $y^* \leq c$

$$\begin{aligned} Pr(y = y^*) &= Pr(y^* \leq c) \\ &= Pr\left(\frac{y^* - \mu}{\sigma} \leq \frac{c - \mu}{\sigma}\right) \end{aligned} \quad (18.46)$$

$$Pr(y = y^*) = \Phi\left(\frac{c - \mu}{\sigma}\right) \quad (18.47)$$

La partie discrète est donnée par la probabilité que le nombre de tickets vendus soit identique au nombre de places disponibles lorsque $y^* > c$

$$\begin{aligned} Pr(y = c) &= Pr(y^* > c) = 1 - Pr(y^* \leq c) \\ &= 1 - Pr\left(\frac{y^* - \mu}{\sigma} \leq \frac{c - \mu}{\sigma}\right) \end{aligned}$$

$$\boxed{Pr(y = c) = 1 - \Phi\left(\frac{c - \mu}{\sigma}\right)} \quad (18.48)$$

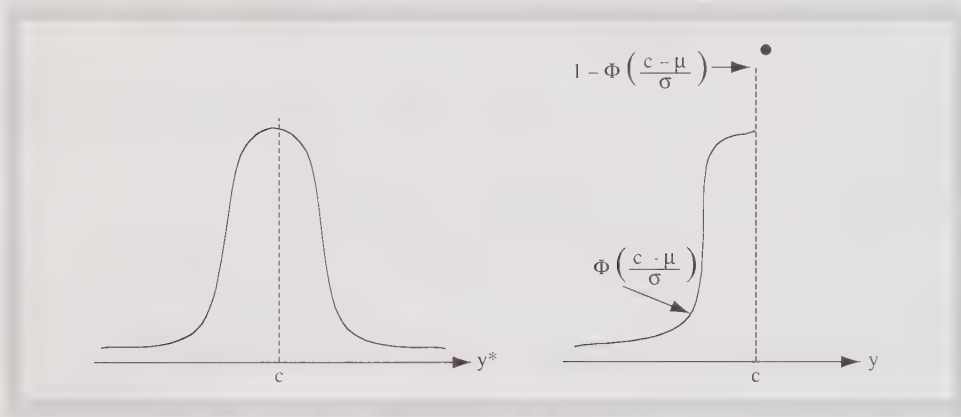


Figure 18.3 Distribution de la variable non censurée et censurée.

La figure 18.3 donne la distribution de la variable y^* . La probabilité que la variable y^* soit supérieure à c est égale à $1 - \Phi\left(\frac{c - \mu}{\sigma}\right)$ et la probabilité que y^* soit inférieure ou égale à c est donnée par $\Phi\left(\frac{c - \mu}{\sigma}\right)$. Elle représente également la distribution de la variable y , elle est censurée au point c . La probabilité que y soit égale à c est identique à la probabilité que y^* soit supérieure à c et la probabilité que y soit inférieure ou égale à c est identique à la probabilité que y^* soit inférieure ou égale à c . Au point de censure on attribue la probabilité totale de la région censurée.

Le point de censure peut être vers le bas ou vers le haut. Un autre exemple souvent cité est le nombre d'heures de travail. Dans un échantillon, certains individus sont caractérisés par un nombre d'heures de travail positif et d'autres par un nombre d'heures de travail nul. Cette variable est tronquée vers le bas, le point de censure étant égal à 0.

De manière plus générale, on considère le modèle suivant

$$\boxed{\begin{aligned} y &= a && \text{si } y^* \leq a \\ y &= y^* && \text{si } y^* > a \end{aligned}} \quad (18.49)$$

où a est le point de censure et y^* une variable aléatoire distribuée selon une loi normale

$$y^* \sim N(\mu, \sigma^2)$$

La distribution de la variable y est identique à celle de y^* lorsque $y^* > a$. Au point de censure, la probabilité que $y = a$ est donnée par la probabilité que $y^* \leq a$

$$\Pr(y = a) = \Pr(y^* \leq a) = \Phi\left(\frac{a - \mu}{\sigma}\right) \quad (18.50)$$

L'espérance de la variable censurée est donnée par

$$\begin{aligned} E(y) &= \Pr(y = a) \cdot E(y_{/y=a}) + \Pr(y = y^*) E(y_{/y=y^*}) \\ &= \Pr(y^* \leq a) \cdot a + \Pr(y^* > a) \cdot E(y_{/y^* > a}) \end{aligned} \quad (18.51)$$

et on montre qu'elle s'écrit

$$E(y) = \Phi(\alpha) \cdot a + (1 - \Phi(\alpha)) \left(\sigma \frac{\phi(\alpha)}{1 - \Phi(\alpha)} + \mu \right) \quad (18.52)$$

avec $\alpha = \frac{a - \mu}{\sigma}$.

Par ailleurs la variance de y est la somme de la variance de l'espérance conditionnelle de y et de l'espérance de la variance conditionnelle de y . On montre que

$$\begin{aligned} \text{var}(y) &= \sigma^2 (1 - \Phi(\alpha)) \cdot \\ &\left[\Phi(\alpha) \left(\alpha \frac{\phi(\alpha)}{1 - \Phi(\alpha)} \right)^2 + \left(1 - \left(\frac{\phi(\alpha)}{1 - \Phi(\alpha)} \right)^2 + \frac{\alpha \phi(\alpha)}{1 - \Phi(\alpha)} \right) \right] \end{aligned} \quad (18.53)$$

2.2.2 Le modèle de régression censuré : le modèle Tobit simple

Soit y_i^* une variable distribuée selon une loi normale telle que

$$\begin{aligned} y_i^* &= \beta_1 + \beta_2 x_{2i} + \dots + \beta_K x_{Ki} + \varepsilon_i = x_i' \beta + \varepsilon_i \\ &= E(y_i^*) + \varepsilon_i \end{aligned} \quad (18.54)$$

$$y_i^* \sim N(x_i' \beta, \sigma^2)$$

le modèle Tobit est composé de deux parties : une partie continue correspondant au modèle de régression linéaire et une partie discrète liée au point de censure supposé égal à 0. Pour simplifier

$$\begin{cases} y_i = 0 & \text{si } y_i^* \leq 0 \\ y_i = y_i^* & \text{si } y_i^* > 0 \end{cases} \quad (18.55)$$

La probabilité que y_i^* prenne une valeur négative ou nulle est donnée par

$$Pr(y_i^* \leq 0) = \Phi\left(-\frac{x'_i\beta}{\sigma}\right) = 1 - \Phi\left(\frac{x'_i\beta}{\sigma}\right) \quad (18.56)$$

et la probabilité que y_i^* prenne une valeur positive s'écrit

$$Pr(y_i^* > 0) = 1 - \Phi\left(-\frac{x'_i\beta}{\sigma}\right) = \Phi\left(\frac{x'_i\beta}{\sigma}\right) \quad (18.57)$$

En utilisant le résultat (18.52) on obtient l'espérance de y

$$\begin{aligned} E(y_i) &= Pr(y_i^* \leq 0) \cdot 0 + Pr(y_i^* > 0) \cdot E(y_i^* | y_i^* > 0) \\ &= \left(\Phi\left(\frac{x'_i\beta}{\sigma}\right)\right) \left(\sigma \frac{\phi\left(-\frac{x'_i\beta}{\sigma}\right)}{\Phi\left(\frac{x'_i\beta}{\sigma}\right)} + x'_i\beta\right) \\ &= \phi\left(\frac{x'_i\beta}{\sigma}\right) \sigma + \Phi\left(\frac{x'_i\beta}{\sigma}\right) x'_i\beta \end{aligned} \quad (18.58)$$

On rappelle que la distribution étant normale $\phi\left(-\frac{x'_i\beta}{\sigma}\right) = \phi\left(\frac{x'_i\beta}{\sigma}\right)$.

De la même manière, à partir de l'équation (18.53), on peut retrouver la variance de y

$$\begin{aligned} var(y) &= \sigma^2 \left(\Phi\left(\frac{x'_i\beta}{\sigma}\right)\right) \cdot \\ &\left(\Phi\left(-\frac{x'_i\beta}{\sigma}\right) \left(\alpha - \frac{\phi\left(\frac{x'_i\beta}{\sigma}\right)}{\Phi\left(\frac{x'_i\beta}{\sigma}\right)}\right)^2 + \left(1 - \frac{\phi\left(\frac{x'_i\beta}{\sigma}\right)}{\Phi\left(\frac{x'_i\beta}{\sigma}\right)} + \frac{x'_i\beta \phi\left(\frac{x'_i\beta}{\sigma}\right)}{\Phi\left(\frac{x'_i\beta}{\sigma}\right)}\right)\right) \end{aligned} \quad (18.59)$$

Le modèle est non linéaire et on utilise le maximum de vraisemblance pour évaluer les paramètres.

2.2.3 Évaluation des paramètres du modèle

Pour évaluer les paramètres on peut appliquer l'estimateur du maximum de vraisemblance. La fonction de vraisemblance du modèle se compose de deux parties : le modèle de régression linéaire et la partie censurée. La fonction s'écrit

$$L = \prod_{y_i > 0} \underbrace{\left(\frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{y_i - \beta' x_i}{\sigma} \right)^2} \right)}_{f(y_i) = f(y_i^*)} \prod_{y_i \leq 0} \underbrace{\left(1 - \Phi \left(\frac{x_i' \beta}{\sigma} \right) \right)}_{Pr(y=0)} \quad (18.60)$$

Pour obtenir les paramètres du modèle, on maximise le logarithme de la fonction de vraisemblance

$$\begin{aligned} \text{Log}L &= \sum_{y_i > 0} - \left(\frac{1}{2} \text{Log}2\pi + \frac{1}{2} \text{Log}\sigma^2 + \frac{1}{2\sigma^2} (y_i - x_i' \beta)^2 \right) + \\ &\sum_{y_i = 0} \text{Log} \left(1 - \Phi \left(\frac{x_i' \beta}{\sigma} \right) \right) \end{aligned} \quad (18.61)$$

Les conditions du premier ordre permettent de calculer les paramètres β et σ^2 . Il s'agit alors de résoudre un système d'équations non linéaire.

2.2.4 Interprétation des résultats

Lorsque le point de censure est égal à 0

$$E(y_i) = Pr(y_i^* > 0) \cdot E(y_i^* / y_i^* > 0) \quad (18.62)$$

Les effets marginaux sont mesurés en calculant

$$\frac{\partial E(y_i)}{\partial x_i} = \frac{\partial Pr(y_i^* > 0)}{\partial x_i} \cdot E(y_i^* / y_i^* > 0) + Pr(y_i^* > 0) \cdot \frac{\partial E(y_i^* / y_i^* > 0)}{\partial x_i} \quad (18.63)$$

Une variation d'une variable explicative induit une variation de la probabilité $Pr(y_i^* > 0)$ et de la moyenne de y_i^* dans la partie positive de sa distribution. On montre que les effets marginaux sont donnés par

$$\boxed{\frac{\partial E(y_i)}{\partial x_i} = \beta \Phi \left(\frac{x_i' \beta}{\sigma} \right)} \quad (18.64)$$

2.3 *Modèle de sélection*

Dans le modèle tronqué et le modèle censuré, la troncature ou la censure est liée à une variable latente correspondant à la variable expliquée. Dans le modèle de sélection, elle est liée à une condition sur une variable autre que la variable latente. On présente ici le cas du modèle avec troncature auxiliaire⁴.

Dans l'exemple sur le nombre d'heures travaillées, certains individus ont un nombre d'heures égal à 0 car ils ne travaillent pas. Une manière d'expliquer ce fait consiste à dire que le salaire proposé par le marché à ces individus est plus faible que leurs salaires de réserve. On fait intervenir dans ce cas deux variables : le nombre d'heures de travail et l'écart entre le salaire du marché et le salaire de réserve. Pour un individu donné, on observe un nombre d'heures positif si l'écart entre les deux salaires est positif, si cet écart est négatif le nombre d'heures de travail est égal à 0.

Plus généralement, on note y la variable étudiée et z la variable qui conditionne la valeur de y . On a

$$E(y_i/y_i \text{ est observée}) = E(y_i/z_i > 0) \quad (18.65)$$

On suppose que

$$y_i = x_i' \beta + \varepsilon_i \quad (18.66)$$

et

$$z_i = w_i' \delta + u_i \quad (18.67)$$

avec x_i et w_i des vecteurs de variables explicatives, β et δ des vecteurs de paramètres, u_i un aléa d'espérance nulle et de variance σ_u^2 et ε_i un aléa d'espérance nulle et de variance σ_ε^2 . Les deux aléas sont distribués selon une loi normale et on note ρ le coefficient de corrélation entre les deux aléas. Si l'équation (18.66) est estimée par les Moindres Carrés Ordinaires on aura un biais dans les estimations lié à la sélection de l'échantillon. On doit estimer un modèle qui tient compte de la sélection de l'échantillon qui dépend de la variable z_i .

z_i prend une valeur positive si u_i est supérieur à $-w_i' \delta$ d'où

$$E(y_i/z_i > 0) = E(y_i/u_i > -w_i' \delta) \quad (18.68)$$

et on montre que

$$E(y_i/z_i > 0) = x_i' \beta + \rho \sigma_\varepsilon \lambda_i(\alpha_u) \quad (18.69)$$

$$\text{avec } \alpha_u = -\frac{w_i' \delta}{\sigma_u} \text{ et } \lambda_i(\alpha_u) = \frac{\phi\left(-\frac{w_i' \delta}{\sigma_u}\right)}{1 - \Phi\left(-\frac{w_i' \delta}{\sigma_u}\right)} = \frac{\phi\left(\frac{w_i' \delta}{\sigma_u}\right)}{\Phi\left(\frac{w_i' \delta}{\sigma_u}\right)}$$

⁴ « *Incidental Truncation* » en anglais.

Le modèle de sélection s'écrit

$$y_i/z_i > 0 = x_i'\beta + \rho\sigma_\varepsilon\lambda_i(\alpha_u) + v_i \quad (18.70)$$

avec

$$\text{var}(v_i) = \sigma_\varepsilon^2(1 - \rho^2[\lambda_i(\alpha_u)(\lambda_i(\alpha_u) - \alpha_u)]) \quad (18.71)$$

Le modèle (18.70) est non linéaire. Pour estimer ce modèle qui prend en compte le problème de sélection de l'échantillon on peut utiliser la procédure d'Heckman. Cette méthode est plus facile à appliquer que la méthode du maximum de vraisemblance mais elle donne des estimateurs moins précis.

La procédure d'Heckman comprend deux étapes

1. Dans la première étape on estime le modèle de sélection avec l'ensemble des données.

Si le processus de sélection est décrit par un modèle linéaire de la forme

$$z_i = w_i'\delta + u_i$$

on applique une méthode d'estimation standard afin d'estimer δ .

Si le processus de sélection est de la forme suivante

$$\begin{aligned} z_i &= w_i'\delta + u_i \\ z_i^* &= 1 \quad \text{si } z_i > 0 \\ z_i^* &= 0 \quad \text{si } z_i < 0 \end{aligned} \quad (18.72)$$

où z_i est une variable latente non observée et z_i^* est une variable binaire. On estime la probabilité que la variable z_i^* prend la valeur 1 avec un modèle de choix binaire de type probit. Le modèle (18.72) est alors estimé avec la méthode du maximum de vraisemblance selon la procédure décrite en début de chapitre.

Au terme de cette première étape on évalue $\lambda_i(\alpha_u)$

$$\lambda_i(\widehat{\alpha}_u) = \frac{\varphi\left(\frac{w_i'\widehat{\delta}}{\widehat{\sigma}_u}\right)}{\Phi\left(\frac{w_i'\widehat{\delta}}{\widehat{\sigma}_u}\right)}$$

2. Dans la seconde étape, on retient le sous-échantillon des individus pour lesquels la variable z_i est positive et on estime le modèle linéaire suivant en tenant compte du problème d'hétéroscédasticité

$$y_i/z_i > 0 = x_i'\beta + \rho\sigma_\varepsilon\lambda_i(\widehat{\alpha}_u) + v_i \quad (18.73)$$

Enfin, pour interpréter les résultats on doit calculer les effets marginaux en dérivant par rapport à x_{ki} , variable supposée présente simultanément dans les vecteurs x_i et w_i , l'espérance de l'expression (18.70). On obtient

$$\frac{\partial E(y_i/z_i > 0)}{\partial x_{ik}} = \beta_k - \delta_k \frac{\rho\sigma_\varepsilon}{\sigma_u} \lambda_i(\alpha_u) (\lambda_i(\alpha_u) - \alpha_u) \quad (18.74)$$

Pour aller au delà de cette présentation on peut se reporter à Thomas A. (2000)⁵.

⁵ Thomas A., 2000, *Économétrie des variables qualitatives*, Dunod.

19

DÉTERMINANTS DES DÉCISIONS D'OFFRE DE TRAVAIL

1. Introduction
2. Théorie microéconomique de la participation au marché du travail
3. Approche économétrique de l'offre de travail
4. Analyse empirique
5. Éléments de corrigé

OBJECTIFS DU CHAPITRE

D'un point de vue méthodologique

- Estimer un modèle logit/probit
- Estimer un modèle tobit
- Estimer un modèle de sélection
- Calculer des effets marginaux

D'un point de vue économique

- Modéliser la décision individuelle de travailler
- Modéliser une équation de salaire
- Modéliser l'offre de travail

NIVEAU

MASTER, première et deuxième années

CHAMPS ÉCONOMIQUES TRAITÉS

Microéconomie

Offre de travail

1. INTRODUCTION

Les études appliquées visant à expliquer la décision de participation des individus à une activité professionnelle et la caractérisation de leurs fonctions d'offre de travail reposent généralement sur le modèle néo-classique d'arbitrage entre travail rémunéré et loisir. Dans sa version la plus simple, le programme de l'individu consiste à maximiser sa fonction d'utilité, qui dépend du temps de travail et des biens de consommation, sous les contraintes de revenu et de temps disponible. La résolution de ce programme fournit la condition de participation au marché du travail : l'individu exerce une activité rémunérée si son salaire de marché est supérieur à la valeur marginale du temps consacré au loisir (ou salaire de réserve, voir par exemple, Killingsworth, 1983)¹. Ce modèle de base peut être enrichi pour tenir compte du contexte familial, des coûts inévitables liés à la participation au marché du travail, des caractéristiques du système fiscal, de la distinction entre travail domestique et loisir stricto-sensu etc. Dans ce chapitre, on présente le modèle néo-classique de base pour expliquer comment la spécification économétrique est dérivée.

On présente dans la seconde section le modèle de base utilisé. La troisième section, porte sur la spécification économétrique et les principes d'estimation. La quatrième section est consacrée à l'analyse empirique.

2. THÉORIE MICROÉCONOMIQUE DE LA PARTICIPATION AU MARCHÉ DU TRAVAIL

Dans le premier paragraphe on détermine de manière analytique les conditions de participation au marché du travail. Ces conditions sont interprétées graphiquement dans un second paragraphe :

2.1 *Les déterminants de l'arbitrage entre travail et loisir*

2.1.1 Le programme de comportement

Dans sa version la plus simple, le programme de l'individu consiste à maximiser sa fonction d'utilité, qui dépend du temps consacré au loisir et des biens de consommation, sous les contraintes de revenu et de temps disponible. Ce programme s'écrit

$$\max_{lei, C, L} U(lei, C)$$

¹ Killingsworth M.R. (1983). *Labor Supply*, Cambridge University Press, London.

sous les contraintes

$$lei + L = T \quad (19.1)$$

$$p_C C = wL + V \quad (19.2)$$

$$lei \geq 0 \quad (19.3)$$

$$C \geq 0 \quad (19.4)$$

où U définit la fonction d'utilité de l'individu qui dépend du temps consacré à du loisir lei , du vecteur des consommations C , p_C le prix associé. L définit le temps consacré à du travail, T le temps maximal disponible. La fonction U est supposée croissante, concave et doublement différentiable en lei et en C .

La première contrainte décrit la répartition du temps total disponible entre le loisir et le travail.

La contrainte budgétaire (équation 18.2) égalise la consommation en valeur au revenu disponible du ménage, ici défini comme la somme du revenu lié à l'exercice d'une activité professionnelle et des revenus non salariaux. Le revenu obtenu grâce à l'exercice d'une activité professionnelle salariée est simplement le produit du nombre d'heures de travail (L) par le salaire correspondant (w). Les revenus non salariaux sont notés V .

Les deux autres contraintes définissent les contraintes de positivité concernant les variables endogènes.

La contrainte budgétaire peut se réécrire

$$p_C C = w(T - lei) + V \quad (19.5)$$

soit

$$p_C C + wlei = wT + V \quad (19.6)$$

Le nouveau programme de maximisation défini à partir des variables endogènes C et lei s'écrit

$$\max_{lei, C} U(lei, C) \quad (19.7)$$

sous les contraintes

$$T - lei \geq 0 \quad (19.8)$$

$$p_C C + wlei = wT + V \quad (19.9)$$

$$lei \geq 0 \quad (19.10)$$

$$C \geq 0 \quad (19.11)$$

Les conditions du premier ordre associées au programme (19.7) permettent de définir la condition de participation de l'individu au marché du travail.

2.1.2 Interprétation des conditions du premier ordre

Le Lagrangien associé au programme de maximisation de l'utilité s'écrit

$$\begin{aligned} \mathfrak{L}(lei, C, \lambda, \mu_1, \mu_2) = & U(lei, C) + \eta(T - lei) \\ & + \lambda(wT + V - p_C C - wlei) + \mu_1 lei + \mu_2 C \end{aligned} \quad (19.12)$$

où η , λ représentent respectivement les multiplicateurs associés aux contraintes (19.8), (19.9), μ_1 et μ_2 les multiplicateurs associés aux contraintes de positivité des variables endogènes.

Les conditions du premier ordre s'écrivent

$$\frac{\partial \mathfrak{L}}{\partial lei} = \frac{\partial U}{\partial lei} - \eta - \lambda w + \mu_1 = 0 \quad (19.13)$$

$$\frac{\partial \mathfrak{L}}{\partial C} = \frac{\partial U}{\partial C} - \lambda p_C + \mu_2 = 0 \quad (19.14)$$

$$\eta(T - lei) = 0 \quad (19.15)$$

$$\lambda(wT + V - p_C C - wlei) = 0 \quad (19.16)$$

$$\mu_1 lei = 0 \quad (19.17)$$

$$\mu_2 C = 0 \quad (19.18)$$

Sans perte de généralités, on suppose que le niveau de consommation et le temps consacré au loisir sont strictement positifs. Les multiplicateurs μ_1 et μ_2 sont égaux à 0.

L'équation (19.13) s'écrit

$$\frac{\partial U}{\partial lei} = +\eta + \lambda w \quad (19.19)$$

L'équation (19.14) s'écrit

$$\frac{\partial U}{\partial C} = \lambda p_C \quad (19.20)$$

Si on fait le rapport entre les deux équations précédentes, on peut définir une relation expliquant les variations du taux marginal de substitution de la consommation au loisir

$$\frac{\partial U / \partial lei}{\partial U / \partial C} = \frac{\eta}{\lambda p_C} + \frac{w}{p_C} \quad (19.21)$$

$$\frac{\partial U / \partial lei}{\partial U / \partial C} = \frac{w}{p_C} + \frac{\eta}{\partial U / \partial C} \quad (19.22)$$

Le terme de gauche de cette égalité définit le taux marginal de substitution de la consommation au loisir, c'est-à-dire le nombre d'unités de consommation nécessaire pour compenser une réduction unitaire du temps de loisir, et maintenir la satisfaction de l'individu.

L'équation (19.22) a une interprétation intéressante en termes de participation au marché du travail.

En effet, si le multiplicateur η est nul alors on a une solution intérieure pour le loisir (le temps consacré à du loisir noté lei^* est strictement positif et donc le niveau d'offre de travail est positif) et à l'équilibre on a

$$\frac{\partial U / \partial lei}{\partial U / \partial C}_{lei=lei^*} = \frac{w}{p_C} \quad (19.23)$$

Dans le cas d'une solution intérieure (stricte positivité du temps de loisir et donc de l'offre de travail), il y a égalité entre le taux marginal de substitution et le salaire réel.

Par contre, si le multiplicateur η est strictement positif alors la contrainte (19.8) est saturée, on a une solution en coin pour le loisir (le temps de loisir est égal au temps maximal disponible et l'individu ne travaille pas) et à l'équilibre on a

$$\frac{\partial U / \partial lei}{\partial U / \partial C}_{lei=T} > \frac{w}{p_C} \quad (19.24)$$

Le taux marginal de substitution de la consommation au loisir évalué au point où l'individu ne travaille pas est supérieur au salaire réel. En d'autres termes, le salaire n'est pas suffisant pour compenser une réduction unitaire du temps de loisir. Le terme de gauche de cette inégalité définit le salaire de réserve de l'individu ou le coût d'opportunité du travail. Il représente l'augmentation de la consommation évaluée en termes réels nécessaire pour compenser une réduction unitaire du temps de loisir, la satisfaction de l'individu étant inchangée.

2.2 Critère de participation au marché du travail : illustration graphique

Pour illustrer le critère de participation au marché du travail, les équilibres associés à trois valeurs différentes de salaires sont représentés sur la figure 19.1.

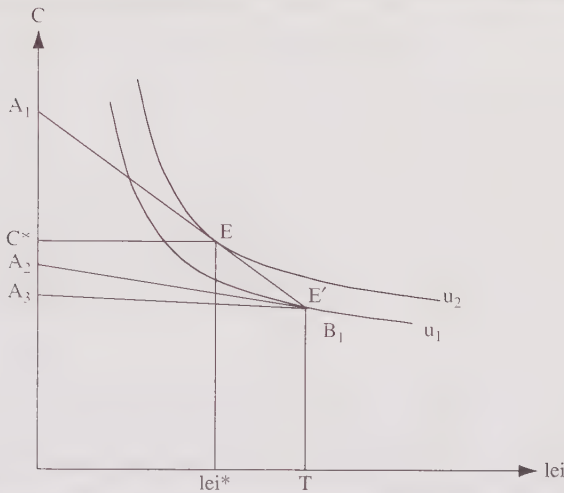


Figure 19.1 Représentation des équilibres

En abscisse, on représente le temps consacré à du loisir (noté lei), en ordonnée le niveau de consommation (noté C). La ligne coudée A_1B_1T représente la droite de budget pour un salaire égal à w_1 . De A_1 à B_1 , la pente de la droite est égale en valeur absolue au salaire réel (w_1/p_C), à partir du point B_1 , la droite de budget devient verticale car en raison de la contrainte de temps, l'individu peut consacrer au maximum T heures (abscisse du point B_1) à du loisir. Le niveau maximal de consommation est égal à $(w_1T + V/p_C)$ (ordonnée du point A_1). Par ailleurs, les courbes u_1 et u_2 représentent des courbes d'indifférence.

Dans le cas d'un salaire réel égal à w_1/p_C , vu la forme des courbes d'indifférence, la solution optimale (point E) est une solution intérieure à la fois pour le niveau de consommation et pour l'offre de travail.

Si le salaire nominal diminue de w_1 à w_2 , la nouvelle droite de budget est la droite A_2B_1T . Dans l'exemple représenté sur la figure 19.1, le temps consacré à du loisir augmente (et donc l'offre de travail diminue). Le nouvel équilibre se situe en E' et caractérise une solution en coin pour l'offre de travail. En effet, pour ce point l'individu consacre tout son temps disponible à du loisir (l'offre de travail est donc nulle). Cet équilibre est particulier dans le sens où il y a toujours égalité entre le taux marginal de substitution et le salaire réel.

$$\frac{\partial U / \partial lei}{\partial U / \partial C} \Big|_{L=0} = \frac{w_2}{p_C} \quad (19.25)$$

Cette valeur du salaire correspond à un point limite. En effet, si l'on diminue encore le salaire (si le salaire nominal diminue de w_2 à w_3), la nouvelle droite de budget est la droite A_3B_1T . L'équilibre se situe toujours en E' . Pour ce point, il

n'y a plus tangence entre la courbe d'indifférence et la nouvelle droite de budget de pente w_3/p_C . On a en effet

$$\frac{\partial U / \partial lei}{\partial U / \partial C} \Big|_{L=0} > \frac{w_3}{p_C} \quad (19.26)$$

Le terme de gauche de l'inégalité (19.26) est le taux marginal de substitution des biens de consommation au travail évalué pour un niveau de travail nul (c'est-à-dire la pente de la courbe d'indifférence évaluée au point où l'individu ne travaille pas). Ce terme définit le salaire de réserve de l'individu. Au point E' , l'individu ne travaille pas car le salaire perçu pour une activité professionnelle est inférieur au salaire de réserve.

Dans l'exemple représenté sur la figure 19.1, une diminution du salaire conduit à une diminution de l'offre de travail. Ce résultat est dépendant de la forme des courbes d'indifférence représentées sur la figure (et donc de la fonction d'utilité). De manière générale, l'impact d'une variation du salaire sur l'offre de travail est indéterminé. Cette variation a deux effets opposés. Le premier effet, l'effet substitution, conduit à une baisse du loisir : l'augmentation du salaire provoque un renchérissement du prix du loisir (car le salaire définit le coût d'opportunité du loisir). Cet effet substitution diminue le loisir et augmente la consommation. Le second effet, l'effet revenu, augmente la « consommation » de loisir sous l'hypothèse que le loisir est un bien normal. Les deux effets sont donc opposés et l'effet total est indéterminé. Par contre, une augmentation du revenu exogène se traduit sur la figure précédente par un déplacement parallèle de la droite de budget. Sous l'hypothèse que le loisir est un bien normal, cette augmentation se traduit par une hausse du temps de loisir et par une réduction de l'offre de travail. L'effet revenu induit par la hausse du revenu exogène a donc un effet négatif sur l'offre de travail (voir Picard, 1990)².

2.3 Illustration paramétrique

Le cadre simplifié que l'on vient de présenter permet d'expliquer la décision de participer au marché du travail et le nombre d'heures consacrées au travail. Pour illustrer comment les modèles économétriques peuvent être dérivés à partir de ce cadre, on retient une forme fonctionnelle particulière pour la fonction d'utilité : une forme Cobb-Douglas. La fonction d'utilité s'écrit dans ce cas

$$U(le_i, C) = lei^\alpha C^\beta \quad (19.27)$$

avec α et $\beta \geq 0$ et $\alpha + \beta \leq 1$.

Les contraintes budgétaire et de temps sont

$$p_C C = wL + V \quad (19.28)$$

$$lei + L = T \quad (19.29)$$

² Picard P. (1990) *Éléments de microéconomie*, 2^e édition, Montchrestien.

Elles peuvent être réécrites

$$C = \frac{wL + V}{p_C} \quad (19.30)$$

$$lei = T - L \quad (19.31)$$

Pour déterminer le salaire de réserve et donc le critère de participation, le taux marginal de substitution doit être calculé.

$$TMS = \frac{\partial U / \partial lei}{\partial U / \partial C} = \frac{\alpha C}{\beta lei} = \frac{\alpha}{\beta} \left(\frac{wL + V}{p_C(T - L)} \right) \quad (19.32)$$

Le salaire de réserve noté w_r est calculé en évaluant le taux marginal de substitution au point où $L = 0$ et $lei = T$.

$$w_r = \frac{\alpha}{\beta} \frac{V}{p_C T} \quad (19.33)$$

On en déduit le modèle permettant d'expliquer la participation au marché du travail.

$$\begin{array}{ll} L = 0 & \text{si } w \leq w_r \\ L > 0 & \text{si } w > w_r \end{array}$$

Il reste à définir le niveau d'offre de travail quand le salaire est supérieur au salaire de réserve. La forme empirique de la fonction d'offre de travail (sachant que le nombre d'heures est strictement positif) est déterminée en appliquant la condition (19.23) (condition pour une solution intérieure).

$$\frac{\alpha}{\beta} \left(\frac{wL + V}{p_C(T - L)} \right) = \frac{w}{p_C} \quad (19.34)$$

$$L = \frac{\beta T}{\alpha + \beta} - \frac{V}{w} \frac{\alpha}{(\alpha + \beta)} \quad (19.35)$$

Le modèle d'offre de travail s'écrit

$$\begin{array}{ll} L = 0 & \text{si } w \leq \frac{\alpha}{\beta} \frac{V}{p_C T} \\ L = \frac{\beta T}{\alpha + \beta} - \frac{V}{w} \frac{\alpha}{(\alpha + \beta)} & \text{sinon} \end{array}$$

Une extension possible consiste à introduire l'effet de caractéristiques individuelles (âge, formation) et l'effet des caractéristiques familiales (nombre d'enfants

par exemple) sur les décisions de travail. Ces variables exogènes ont vraisemblablement une influence sur la définition du salaire de réserve de l'individu. Dans ce cas, l'utilité de l'individu est supposée dépendre de caractéristiques individuelles et familiales respectivement notées E et A . Ainsi la fonction d'utilité s'écrit

$$U(lei, C, E, A)$$

Ce cadre permet de montrer comment déterminer simultanément, à partir d'une forme paramétrique, pour la fonction d'utilité, le salaire de réserve et la fonction d'offre de travail. Dans un modèle d'offre de travail il y a donc, au minimum, deux variables à expliquer : la décision d'exercer une activité professionnelle et le nombre d'heures offertes. Dans le modèle statique considéré, les décisions de participation et d'offre de travail (nombre d'heures de travail) sont simultanées.

3. APPROCHE ÉCONOMÉTRIQUE DE L'OFFRE DE TRAVAIL

On spécifie dans cette section le modèle économétrique à estimer.

3.1 Définition du salaire de réserve

La détermination du salaire de réserve nécessite la détermination du taux marginal de substitution. De manière très générale, on suppose que le taux marginal de substitution pour l'individu i se définit de manière linéaire par rapport aux variables explicatives

$$TMS_i = X_i' \theta + \varepsilon_{tmsi} \quad (19.36)$$

où TMS_i représente le taux marginal de substitution pour l'individu i , X_i définit le vecteur des variables explicatives (vecteur ligne), θ le vecteur des coefficients à estimer et ε_{tmsi} l'aléa de l'équation. X_i intègre toutes les variables exogènes du modèle (caractéristiques individuelles, caractéristiques familiales,...) le nombre d'heures de travail (voir l'équation 18.32) et le niveau de salaire.

L'équation (19.36) est résolue en $L = 0$ pour déterminer la relation définissant le salaire de réserve.

Le salaire de réserve se définit de manière linéaire par rapport aux variables explicatives

$$w_{ri} = X_i^* \theta + \varepsilon_{ri} \quad (19.37)$$

où X_i^* correspond au vecteur X_i sans les variables L (heures de travail), w (niveau de salaire), ε_{ri} l'aléa de l'équation.

Le modèle de participation s'écrit de manière générale

$$\begin{aligned} L &= 0 & \text{si } w_i &\leq X_i^{*'}\theta + \varepsilon_{ri} \\ L_i &> 0 & \text{si } w_i &> X_i^{*'}\theta + \varepsilon_{ri} \end{aligned} \tag{19.38}$$

3.2 Définition de l'offre de travail

Pour définir la fonction d'offre de travail, on suppose que le nombre d'heures de travail est proportionnel à la différence entre le salaire et le salaire de réserve

$$L_i = b(w_i - w_{ri}) \tag{19.39}$$

où b est le coefficient de proportionnalité³.

Ainsi, le modèle d'offre de travail s'écrit

$$\begin{aligned} L_i &= 0 & \text{si } w_i &\leq w_{ri} \\ L_i &= b(w_i - w_{ri}) = bw_i - bX_i^{*'}\theta - b\varepsilon_{ri} & \text{si } w_i &> w_{ri} \end{aligned} \tag{19.40}$$

3.3 Équation de salaire

Dans l'analyse empirique, on endogénéise une troisième variable à savoir le niveau de salaire. En effet cette variable est uniquement observée pour les gens qui travaillent.

Une équation expliquant les salaires est définie

$$w_i = Z_i'\Gamma + \varepsilon_{wi} \tag{19.41}$$

où Z_i définit le vecteur des variables explicatives du salaire (caractéristiques individuelles et caractéristiques locales du marché du travail), Γ le vecteur des paramètres estimés et ε_{wi} l'aléa.

3.4 Modèle de participation

Dans un premier temps, on écrit le modèle économétrique qui explique la probabilité pour l'individu i d'exercer une activité professionnelle. Cette probabilité (notée Pr) est égale à la probabilité que la rémunération soit supérieure au salaire de réserve. De manière analytique, on peut écrire

$$Pr(i \text{ travaille}) = Pr(w_i > w_{ri}) \tag{19.42}$$

³ Dans la partie économétrique, on verra que ce coefficient dépend des paramètres estimés.

En remplaçant le salaire et le salaire de réserve par leurs expressions, on obtient

$$\begin{aligned}\Pr(i \text{ travaille}) &= \Pr(Z_i'\Gamma + \varepsilon_{wi} > X_i^*\theta + \varepsilon_{ri}) & (19.43) \\ &= \Pr(\varepsilon_{wi} - \varepsilon_{ri} > X_i^*\theta - Z_i'\Gamma) \\ &= \Pr(\varepsilon_{ti} > S_i)\end{aligned}$$

avec $\varepsilon_{ti} \doteq \varepsilon_{wi} - \varepsilon_{ri}$ de moyenne nulle, de variance égale à $\sigma_w^2 + \sigma_r^2 + 2\sigma_{wr}$ où σ_{wr} note la covariance entre ε_{ri} et ε_{wi} , $S_i = X_i^*\theta - Z_i'\Gamma$. Suivant le choix de la loi pour l'aléa, on utilise un modèle probit ou un modèle logit pour estimer la probabilité. Ainsi si on suppose que ε_{ti} suit une loi normale de moyenne nulle et de variance notée σ_t , la probabilité s'écrit

$$\begin{aligned}\Pr(i \text{ travaille}) &= \Pr\left(\frac{\varepsilon_{ti}}{\sigma_t} > \frac{S_i}{\sigma_t}\right) & (19.44) \\ &= 1 - \Phi\left(-\frac{S_i}{\sigma_t}\right)\end{aligned}$$

où Φ définit la fonction de répartition de la loi normale centrée réduite. Pour estimer cette probabilité on utilise la méthode du maximum de vraisemblance. On détermine dans ce cas un modèle probit standard.

4. ANALYSE EMPIRIQUE

On dispose des résultats d'une enquête sur l'activité professionnelle des femmes mariées pour l'année 1995. L'échantillon comprend 311 individus et on dispose de différentes variables : L le nombre d'heures de travail réalisées par la femme en 1995, $ENF6$ le nombre d'enfants de moins de 6 ans du ménage, $ENF618$ le nombre d'enfants entre 6 et 18 ans du ménage, $AGEF$ l'âge de la femme, $FORM$ le nombre d'années de formation de la femme, w le salaire horaire de la femme en francs 1995, $REVENU$ le revenu du ménage en 1995 et $CHOMAGE$ le taux de chômage dans la région.

Problème 19.1

Construction des variables et analyse descriptive

1. Créer une variable indicatrice notée lfp qui est égale à 1 quand la femme a une activité professionnelle, 0 sinon. Quelle est la proportion de femmes qui ne travaille pas ? Calculer le revenu exogène, c'est-à-dire le revenu non salarial (cette variable est notée V). Créer une variable indicatrice notée f à partir de la variable $FORM$: f est égale à 1 si l'individu a suivi des études supérieures ($FORM > 12$), 0 sinon. Enfin, calculer la variable âge au carré (notée $A2$).
2. Calculer les moyennes et écarts types des variables pour tout l'échantillon, puis pour l'échantillon des femmes qui travaillent, qui ne travaillent pas. Quels enseignements peut-on retirer des statistiques ?

Problème 19.2

Modèle de participation

1. Expliquer la probabilité que la variable lfp soit égale à un en utilisant un modèle logit puis un modèle probit. Les variables explicatives sont les variables : $ENF6$, $ENF618$, $AGEF$, $A2$, f , V et $CHOMAGE$. Commenter les résultats.
2. Pour le modèle probit, est-ce que la variable âge a un effet constant ? Interpréter l'effet de l'âge sur la probabilité de participation, en particulier calculer l'effet maximal de l'âge.
3. Pour le modèle probit, calculer les effets marginaux des variables au point moyen de l'échantillon, puis pour une probabilité initiale égale à 0.9. Calculer si nécessaire les quasi-élasticités (semi-élasticités).
4. Proposer une mesure de la qualité d'ajustement du modèle.

Problème 19.3

Estimation du salaire et de la fonction d'offre de travail

1. Les déterminants du salaire sont supposés être les caractéristiques individuelles et les caractéristiques du marché du travail. Estimer avec la méthode adéquate une équation de salaire sur l'échantillon total. Les variables explicatives sont les variables : $AGEF$, $A2$ et f . Donner l'effet marginal de l'âge sur le salaire au point moyen. Interpréter les résultats.
2. On sélectionne le sous-échantillon des femmes qui ont une activité professionnelle. Proposer une spécification pour l'équation d'offre. Calculer l'élasticité par rapport au salaire et aux autres revenus.

5. ÉLÉMENTS DE CORRIGÉ

Problème 19.1

Construction des variables et analyse descriptive

1. Pour créer la variable indicatrice notée lfp , une sélection est réalisée par rapport à la variable mesurant les heures de travail (variable L). Ainsi, si $L > 0$ alors la variable vaut 1, sinon elle vaut 0. Il y a, dans notre échantillon, 38.58 % de femmes qui ne travaillent pas.

On dispose du revenu total du ménage (noté $REVENU$). Pour calculer le revenu non salarial (revenu considéré de manière exogène), on soustrait du revenu total, le revenu salarial c'est-à-dire le produit entre le nombre d'heures et le salaire horaire.

Pour créer la variable indicatrice f , une sélection est appliquée par rapport à la variable mesurant le nombre d'années de formation. Ainsi, suivant la

définition imposée, la variable f vaut 1 quand ($FORM > 12$), sinon elle vaut 0.

- Quand on compare les moyennes des différentes variables des deux sous-échantillons, on remarque que pour les deux variables «enfants», les moyennes sont supérieures dans le sous-échantillon des femmes qui ne travaillent pas aux moyennes dans le sous-échantillon des femmes qui travaillent. On s'attend donc à ce que ces deux variables aient un impact négatif sur la probabilité de travailler. Par contre, le niveau de formation est, en moyenne, plus élevé dans l'échantillon des femmes qui travaillent. Cette variable devrait avoir un impact positif sur la probabilité de travailler.

Tableau 19.1a Statistiques pour l'ensemble de l'échantillon

Series	Obs	Mean	Std Error	Minimum	Maximum
L	311	780.877814	821.883822	0.000000	3087.000000
ENF6	311	0.231511	0.524654	0.000000	3.000000
ENF618	311	1.289389	1.233870	0.000000	7.000000
AGEF	311	42.054662	8.079059	30.000000	60.000000
FORM	311	12.385852	2.353924	5.000000	17.000000
W	311	34.891732	36.648466	0.000000	263.576700
REVENU	311	233751.971704	121386.906053	42804.000000	893472.000000
CHOMAGE	311	8.673633	2.996839	3.000000	14.000000
LFP	311	0.614148	0.487580	0.000000	1.000000
NLFP	311	0.385852	0.487580	0.000000	1.000000
V	311	190011.310687	116442.524029	15192.876360	758172.118080
F	311	0.315113	0.465309	0.000000	1.000000
A2	311	1833.655949	694.923301	900.000000	3600.000000

Tableau 19.1b Statistiques pour l'échantillon des femmes qui travaillent

Series	Obs	Mean	Std Error	Minimum	Maximum
L	191	1271.481675	689.233807	15.000000	3087.000000
ENF6	191	0.151832	0.414201	0.000000	2.000000
ENF618	191	1.272251	1.217942	0.000000	5.000000
AGEF	191	41.350785	7.531512	30.000000	59.000000
FORM	191	12.910995	2.247205	5.000000	17.000000
W	191	56.813238	30.649689	36.980000	263.576700
REVENU	191	241116.454450	112305.316696	58695.600000	893472.000000
CHOMAGE	191	8.625654	3.034438	3.000000	14.000000
LFP	191	1.000000	0.000000	1.000000	1.000000
NLFP	191	0.000000	0.000000	0.000000	0.000000
V	191	169894.749862	98898.299754	15192.876360	758172.118080
F	191	0.382199	0.487202	0.000000	1.000000
A2	191	1766.314136	640.518241	900.000000	3481.000000

Tableau 19.1c Statistiques pour l'échantillon des femmes qui ne travaillent pas

Series	Obs	Mean	Std Error	Minimum	Maximum
L	120	0.000000	0.000000	0.000000	0.000000
ENF6	120	0.358333	0.645551	0.000000	3.000000
ENF618	120	1.316667	1.263471	0.000000	7.000000
AGEF	120	43.175000	8.798073	30.000000	60.000000
FORM	120	11.550000	2.285688	5.000000	16.000000
W	120	0.000000	0.000000	0.000000	0.000000
REVENU	120	222030.170000	134244.322318	42804.000000	757680.000000
CHOMAGE	120	8.750000	2.947011	3.000000	14.000000
LFP	120	0.000000	0.000000	0.000000	0.000000
NLFP	120	1.000000	0.000000	1.000000	1.000000
V	120	222030.170000	134244.322318	42804.000000	757680.000000
F	120	0.208333	0.407819	0.000000	1.000000
A2	120	1940.841667	764.274754	900.000000	3600.000000

Problème 19.2

Modèle de participation

1. Pour estimer la probabilité d'exercer une activité professionnelle, deux modèles alternatifs sont utilisés : le modèle logit et le modèle probit. Dans la partie économique, on a vu qu'il était nécessaire de spécifier une distribution statistique pour l'aléa (voir équation 18.43). Si l'on spécifie une loi logistique, on estime un modèle logit. Si l'on spécifie une distribution normale, on estime un modèle probit. Dans la pratique, les modèles logit et probit sont très similaires en termes d'ajustement statistique. Il n'y a pas d'argument pour choisir le modèle logit plutôt que le modèle probit. On commentera ici de façon détaillée les résultats du modèle probit⁴.

Les valeurs estimées des paramètres ne sont pas directement comparables⁵ et ne sont pas directement interprétables. La première information donnée par les estimations est transmise par le signe des paramètres estimés : un signe positif (respectivement négatif) indique qu'une augmentation de la variable considérée augmente (respectivement diminue) la probabilité d'exercer une activité professionnelle. Ainsi, le nombre d'enfants, quel que soit l'âge de ces derniers diminue la probabilité d'exercer une activité professionnelle.

⁴ Dans la suite de l'exercice, on applique la méthode d'Heckman qui utilise les résultats du modèle probit.

⁵ Pour comparer les valeurs numériques des paramètres estimés on multiplie les coefficients du probit par $\pi/\sqrt{3}$.

Tableau 19.2a *Modèle de participation avec modèle Logit*

Logit — Estimation by Newton-Raphson					
Dependent Variable LFP					
Convergence in	5 Iterations.	Final criterion was	0.0000023 <0.0000100		
Usable Observations	311	Degrees of Freedom	303		
Cases Correct	221				
Log Likelihood	-168.834954				
Average Likelihood	0.5810737				
Pseudo-R**2	0.2399207				
	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	-2.582890565	4.403585671	-0.58654	0.55751095
2.	ENF6	-1.507187103	0.363765985	-4.14329	0.00003424
3.	ENF618	-0.262960882	0.122370283	-2.14889	0.03164273
4.	AGEF	0.331076639	0.206455767	1.60362	0.10879780
5.	A2	-0.004866282	0.002365719	-2.05700	0.03968635
6.	F	1.770711155	0.366260899	4.83456	0.00000133
7.	V	-0.000006864	0.000001475	-4.65463	0.00000325
8.	CHOMAGE	-0.046772786	0.045144191	-1.03608	0.30016697

Tableau 19.2b *Modèle de participation avec modèle Probit*

Probit — Estimation by Newton-Raphson					
Dependent Variable LFP					
Convergence in	5 Iterations	Final criterion was	0.0000001 <0.0000100		
Usable Observations	311	Degrees of Freedom	303		
Cases Correct	223				
Log Likelihood	-168.738807				
Average Likelihood	0.5812534				
Pseudo-R**2	0.2404980				
	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	-1.557142251	2.578742929	-0.60384	0.54595154
2.	ENF6	-0.890561322	0.208082566	-4.27985	0.00001870
3.	ENF618	-0.156652961	0.071274387	-2.19789	0.02795725
4.	AGEF	0.199037864	0.120720001	1.64876	0.09919757
5.	A2	-0.002921528	0.001384877	-2.10959	0.03489339
6.	F	1.037065051	0.206555190	5.02076	0.00000051
7.	V	-0.000004081	0.000000848	-4.81510	0.00000147
8.	CHOMAGE	-0.028816949	0.027085454	-1.06393	0.28736172

2. Dans les deux modèles estimés, on a introduit la variable «âge» et la variable «âge au carré». L'ajout du terme quadratique autorise un effet non constant de l'âge sur la probabilité de participation. De manière générale, la variable «âge» a un coefficient positif tandis que la variable «âge au carré» a un impact négatif. Deux remarques concernant l'effet de l'âge

(i) L'effet global de l'âge sur la probabilité de participation n'est pas constant et dépend des valeurs de l'expression suivante

$$\widehat{\alpha}_{agef} \cdot AGEF + \widehat{\alpha}a2 \cdot A2 \quad (19.45)$$

avec $A2 = AGEF^2$. Cet effet est globalement positif quand

$$-\widehat{\alpha}_{agef} + \widehat{\alpha}a2 \cdot AGEF > 0$$

c'est-à-dire quand

$$AGEF < \frac{-\widehat{\alpha}_{agef}}{\widehat{\alpha}a2}$$

$$AGEF < \frac{0.1990}{0.0029} \simeq 68$$

c'est-à-dire quand l'individu a moins de 68 ans (noté *EAGE* dans le programme rats).

(ii) La probabilité de participation croît avec l'âge dans un premier temps puis décroît dans un second temps. Pour déterminer l'âge pour lequel la probabilité est maximale, on annule la dérivée de la probabilité par rapport à l'âge.

Dans le modèle probit, la probabilité estimée s'écrit

$$\widehat{p} = \Phi(X'\widehat{\beta}) \quad (19.46)$$

où Φ désigne la fonction de répartition de la loi normale centrée réduite, X le vecteur des variables explicatives (dont les variables «âge» et «âge au carré»), $\widehat{\beta}$ le vecteur des coefficients estimés. La dérivée par rapport à l'âge de la probabilité s'écrit

$$\frac{\partial \widehat{p}}{\partial AGEF} = (0.1990 + 2(-0.0029) \cdot AGEF) \phi(X'\widehat{\beta}) \quad (19.47)$$

où ϕ désigne la fonction de densité de la loi normale centrée réduite. L'âge à déterminer vérifie l'égalité suivante

$$(0.1990 + 2(-0.0029) \cdot AGEF) = 0 \quad (19.48)$$

L'effet est maximal pour un âge égal à

$$AGEF = -\frac{0.1990}{2(-0.0029)} \simeq 34 \quad (19.49)$$

On trouve que l'effet positif de l'âge est maximal à 34 ans (noté $AGEMAX$ dans le programme rats).

En résumé, l'effet de l'âge sur la probabilité est positif jusqu'à l'âge de 68 ans, l'impact positif maximal ayant lieu à 34 ans.

3. Pour comparer les effets des différentes variables explicatives, les effets marginaux sont calculés. Pour mesurer ces effets on distingue le cas des variables explicatives binaires et le cas des variables explicatives continues.

- (i) Seule la variable «formation» est binaire. Pour cette variable, on va raisonner en variation de probabilité. On choisit comme individu de référence un individu pour lequel les valeurs des autres variables sont égales aux valeurs moyennes de l'échantillon, $X = \bar{X}$.

Un individu n'ayant pas suivi des études supérieures, a une probabilité estimée d'avoir une activité professionnelle égale à 0.58.

$$\begin{aligned} P_m &= \Phi(-1.55 - 0.89\overline{ENF6} - 0.15\overline{ENF618} + 0.19\overline{AGEF} \\ &\quad - 0.003\overline{A2} - 4.08.10^{-6}\overline{V} - 0.03\overline{CHOMAGE}) \\ &= 0.58 \end{aligned}$$

La probabilité pour un individu ayant réalisé des études supérieures, les autres variables restant inchangées, est donnée par

$$\begin{aligned} P_{m1} &= \Phi(-1.55 - 0.89\overline{ENF6} - 0.15\overline{ENF618} + 0.19\overline{AGEF} \\ &\quad - 0.003\overline{A2} - 4.08.10^{-6}\overline{V} - 0.03\overline{CHOMAGE} + 1.03) \\ &= 0.89 \end{aligned}$$

Pour cet individu, la probabilité d'avoir une activité professionnelle est égale à 0.89.

La variation de la probabilité capte l'effet de la formation «études supérieures». Avoir suivi des «études supérieures» augmente la probabilité d'exercer une activité professionnelle de $0.31 = 0.89 - 0.58$.

- (ii) Pour les variables continues, on utilise les dérivées partielles pour déterminer l'effet marginal. On choisit comme individu de référence un individu pour lequel la valeur de la variable $f = 0$, les valeurs des autres variables sont égales aux valeurs moyennes de l'échantillon.

Pour le modèle probit, les dérivées partielles s'écrivent

$$\frac{\partial \hat{p}}{\partial X_k} = (\hat{\alpha}_k) \phi(X' \hat{\beta}) \quad (19.50)$$

où X_k représente la k ème variable, $\hat{\alpha}_k$ note le coefficient estimé pour cette variable, ϕ désigne la fonction de densité de la loi normale centrée réduite, X le vecteur des variables explicatives, $\hat{\beta}$ le vecteur des coefficients estimés.

Ces dérivées mesurent l'effet d'une augmentation d'une unité de la variable k sur la probabilité d'exercer une activité professionnelle. Ces calculs sont justes

pour toutes les variables apparaissant uniquement de manière linéaire dans la spécification. Pour la variable âge qui apparaît de manière quadratique, le calcul est donné par l'expression (19.47).

Par exemple, un enfant supplémentaire âgé de moins de 6 ans diminue la probabilité d'exercer une activité professionnelle de 0.35. On obtient cette valeur en calculant

$$\frac{\partial \hat{p}}{\partial ENF6} = -0.89\phi(-1.55 - 0.89\overline{ENF6} - 0.15\overline{ENF618} + 0.19\overline{AGEF} - 0.003\overline{A2} - 4.08.10^{-6}\overline{V} - 0.03\overline{CHOMAGE})$$

Pour deux variables (revenu exogène, taux de chômage) il est judicieux de calculer l'effet d'une variation de 1 % de la variable considérée (ce calcul permet de s'affranchir du problème de différence des unités et permet de comparer plus aisément l'effet des variations des variables explicatives). À titre d'exemple ici il est plus intéressant d'interpréter l'effet d'une variation de 1 % des revenus exogènes (plutôt qu'une augmentation d'un franc de cette variable). Ce calcul correspond à la quasi-élasticité (semi-élasticité) définie par le calcul suivant

$$QE = \frac{\partial \hat{p}}{\partial X_k} X_k \quad (19.51)$$

L'effet négatif d'une variation du revenu exogène est plus important en valeur absolue que l'effet négatif d'une variation du chômage. Par exemple, la quasi-élasticité du revenu est donnée par

$$\frac{\partial \hat{p}}{\partial V} V = -4.08.10^6\phi(-1.55 - 0.89\overline{ENF6} - 0.15\overline{ENF618} + 0.19\overline{AGEF} - 0.003\overline{A2} - 4.08.10^{-6}\overline{V} - 0.03\overline{CHOMAGE})\overline{V}$$

Tableau 19.3a Effets marginaux au point moyen de l'échantillon et quasi-élasticités

Effets marginaux	
enf 6	-0.34732
enf 618	-0.06110
agef	0.02971
revenu exterieur	-1.59173e-06
chomage	-0.01124
Calcul des quasi-élasticités	
enf 6	-0.08041
enf 618	-0.07878
revenu exterieur	-0.30245
chomage	-0.09748

Pour mesurer les effets marginaux on peut également retenir comme point de référence un niveau initial de probabilité de participation à une activité

professionnelle. Dans l'exemple, la probabilité est 0.9. Dans la table de la fonction de répartition de la loi normale centrée réduite on lit que pour $X\beta = 1.29$, $\Phi(X'\beta) = 0.9$ où Φ est la fonction de répartition. Puis on calcule la valeur de la fonction de densité en ce point et on peut calculer à nouveau les dérivées partielles. Par exemple, l'effet marginal de la variable $ENF6$ est obtenu en calculant

$$-0.89\phi(1.29) = -0.89 (0.1736) = -0.15$$

Ainsi, pour un individu ayant 90 % de chances d'exercer une activité professionnelle, avoir un enfant de moins de 6 ans diminue la probabilité de 0.15.

De la même manière on peut calculer les quasi-élasticités en multipliant l'effet marginal par la valeur moyenne de la variable. Par exemple, pour la variable de revenu on calcule

$$-4.08 \cdot 10^{-6} \phi(1.29) \bar{V} = -4.08 \cdot 10^{-6} (0.1736) \bar{V} = -7.08 \cdot 10^{-07} \bar{V} = -0.134$$

Tableau 19.3b Effets marginaux pour une probabilité initiale égale à 0.9 et quasi-élasticités

Effets marginaux pour une proba = 0.9	
enf 6	- 0.15460
enf 618	- 0.02719
agef	0.01322
revenu extérieur	- 7.08503e- 07
chomage	- 0.00500
Calcul des quasi-élasticités	
enf 6	- 0.03579
enf 618	- 0.03506
revenu extérieur	- 0.13462
chomage	- 0.04339

4. Un critère de validité du modèle repose sur le pourcentage de prévisions correct. Une méthode possible est le calcul pour chaque individu de la probabilité prédite de travailler. À partir de cette probabilité prédite, on classe chaque observation selon le critère suivant, si la valeur estimée de la probabilité est supérieure ou égale à 0.5 alors la réponse prédite est égale à 1 sinon la réponse prédite est égale à 0. Ainsi dans le sous-échantillon des 191 femmes qui travaillent, le modèle estimé a une bonne prévision pour 161 femmes, par contre le modèle a une prévision erronée pour 30 femmes. Dans le sous-échantillon des 120 femmes qui ne travaillent pas, le modèle estimé a une bonne prévision pour 62 femmes, par contre le modèle a une prévision erronée pour 58 femmes.

Au total, le pourcentage de bonnes prévisions est de 71.70 %.

$$\frac{161 + 62}{311} 100 = 71.70\%$$

Tableau 19.4 Test Prévisions en termes de proportion de prévisions correctes

individus dont LPF observé et estimé = 0	62.00000
individus dont LPF observé et estimé = 1	161.00000
individus dont LPF observé = 1 et estimé = 0	30.00000
individus dont LPF observé = 0 et estimé = 1	58.00000

Problème 19.3

Estimation du salaire et de la fonction d’offre de travail

1. L’estimation du salaire est réalisée à partir de l’échantillon total. On utilise un modèle Tobit pour estimer le salaire. La variable observée w_i prend la valeur 0 lorsque l’individu ne travaille pas, elle prend une valeur positive sinon. Pour estimer le modèle on utilise l’estimateur du maximum de vraisemblance. La fonction de vraisemblance est donnée par

$$\begin{aligned}
 \text{Log}L = & \sum_{w_i > 0} - \left(\frac{1}{2} \text{Log}2\pi + \frac{1}{2} \text{Log}\sigma^2 + \frac{1}{2\sigma^2} (w_i - Z'_i\Gamma)^2 \right) + \\
 & \sum_{w_i = 0} \text{Log} \left(1 - \Phi \left(\frac{Z'_i\Gamma}{\sigma} \right) \right)
 \end{aligned}
 \tag{19.52}$$

avec $Z'_i\Gamma = b_1 + b_2AGEF + b_3A2 + b_4f$

Tableau 19.5 Estimation de l’équation du salaire

Linear Regression — Estimation by Least Squares					
Dependent Variable W					
Usable Observations	311	Degrees of Freedom	307		
Centered R**2	0.138581	R Bar **2	0.130163		
Uncentered R**2	0.548842	T x R**2	170.690		
Mean of Dependent Variable	34.891731640				
Std Error of Dependent Variable	36.648466458				
Standard Error of Estimate	34.180206292				
Sum of Squared Residuals	358663.95617				
Regression F(3,307)	16.4629				
Significance Level of F	0.00000000				
Durbin-Watson Statistic	0.875330				

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	- 88.95315371	54.69314832	- 1.62640	0.10488972
2.	AGEF	5.68195539	2.60410882	2.18192	0.02987349
3.	A2	- 0.06772939	0.03027293	- 2.23729	0.02598467
4.	F	28.82962194	4.20313225	6.85908	0.00000000

MAXIMIZE — Estimation by BFGS					
Convergence in	37 Iterations.	Final criterion was 0.0000038 < 0.0000100			
Usable Observations	311				
Function Value	-941.09608286				
	Variable	Coeff	Std Error	T-Stat	Signif
1.	SIGMASQ	2535.118056	274.748698	9.22704	0.00000000
2.	B1	-200.461490	10.747492	-18.65193	0.00000000
3.	B2	10.596398	0.128340	82.56479	0.00000000
4.	B3	-0.128997	0.003796	-33.98244	0.00000000
5.	B4	39.528599	6.247771	6.32683	0.00000000

L'effet marginal de l'âge sur le salaire d'un individu dont l'âge est égal à \overline{AGEF} (42 ans dans l'échantillon total) et dont $f = 0$ est donné par

$$\begin{aligned} \frac{\partial \hat{w}}{\partial \overline{AGEF}} &= (10.60 + 2(-0.12)\overline{AGEF}) \\ &\quad \Phi \left(\frac{-200.56 + 10.60\overline{AGEF} - 0.12\overline{A2}}{\sqrt{2534.98}} \right) \\ &= -0.16 \end{aligned}$$

Le salaire d'un individu âgé de 42 ans n'ayant pas réalisé d'étude supérieur diminue lorsque son âge augmente. Si on considère un individu âgé de 20 ans, l'effet de l'âge sur le salaire est positif.

2. L'estimation des heures de travail est réalisée à partir du sous-échantillon des individus qui exercent une activité professionnelle. Ce sous-échantillon n'est pas déterminé de façon aléatoire car un critère de sélection est appliqué : seules les personnes pour lesquelles le salaire est supérieur au salaire de réserve sont retenues. Si l'équation est estimée par les Moindres Carrés Ordinaires on peut avoir un biais dans les estimations lié à cette sélection. Pour corriger le biais de sélectivité on utilise ici une méthode d'estimation de type ratio de Mills (méthode d'Heckman). Cette méthode donne des estimateurs sans biais des coefficients mais des estimateurs biaisés des variances. Aussi, on corrige la matrice de variance-covariance des paramètres estimés induite par cette méthode.

Les heures de travail sont positives lorsque l'individu travaille. La procédure d'Heckman consiste à estimer le modèle de participation (modèle probit du problème 18.2) et à en extraire le ratio de Mills. Dans un second temps, ce ratio est utilisé comme variable explicative des heures de travail pour les individus qui travaillent.

Au point moyen de l'échantillon, l'élasticité des heures de travail par rapport au salaire est égale à -0.07 %, le signe négatif indiquant que l'effet revenu est supérieur à l'effet substitution. La variable salaire apparaît seulement dans l'équation des heures de travail, l'élasticité est donnée par

$$-1.62 \frac{\bar{w}}{L} = -0.07$$

La variable revenu V est présente dans le modèle de participation (modèle probit) et dans le modèle d'offre de travail (cf. chapitre 18, Économétrie des variables qualitatives). L'élasticité des heures de travail par rapport au revenu est égale à -0.22 . Une augmentation du revenu V induit une baisse d'offre de travail de 0.22%

Tableau 19.6 Estimation des heures de travail

Linear Regression — Estimation by Least Squares					
Dependent Variable L					
Usable Observations	191	Degrees of Freedom	184		
Total Observations	311	Skipped/Missing	120		
Centered R**2	0.102567	R Bar **2	0.073302		
Uncentered R**2	0.797012	T x R**2	152.229		
Mean of Dependent Variable	1271.4816754				
Std Error of Dependent Variable	689.2338068				
Standard Error of Estimate	663.4918382				
Sum of Squared Residuals	81000741.165				
Durbin-Watson Statistic	1.983452				

	Variable	Coeff	Std Error	T-Stat	Signif
1.	Constant	7916.921594	4350.895442	1.81961	0.06881881
2.	ENF6	-330.510889	157.124377	-2.10350	0.03542222
3.	ENF618	-53.657457	54.023549	-0.99322	0.32060106
4.	AGEF	-2.687713	9.158926	-0.29345	0.76917604
5.	V	-0.000698	0.000592	-1.17901	0.23839386
6.	W	-1.619625	1.730329	-0.93602	0.34926219
7.	LAMBDA	-7893.026608	5830.334148	-1.35379	0.17580456

INDEX

A

Aléas 11, 33, 44
Aléas homoscédastiques 13, 44
Aléas non autocorrélés 13, 44
Anticipations adaptatives 177
Anticipations naïves 178
Arrelano Bond 324

B

Biais de sélection 422
Biais de simultanéité 202
Bruit blanc 13, 45

C

Coefficient d'ajustement partiel 179, 186, 190
Coefficient d'aplatissement 125
Coefficient d'asymétrie 125
Coefficient de détermination 15, 49, 83, 114, 305
Coefficient de détermination ajusté 49, 87, 167
Coefficient de détermination de Mc Fadden 410
Convergence en probabilité 47
Correction de White 145
Corrélogramme 335
Critère d'information AIC 336, 366, 383
Critère d'information BIC 336, 366, 383
Croissance potentielle 101

D

Délai moyen de retour à l'équilibre 361
Distribution d'une variable censurée 417
Distribution d'une variable tronquée 413
Distribution de Fisher-Snédecor 18
Distribution de l'estimateur des MCO 50

Distribution de Student 18
Distribution du Chi deux 18
Distribution normale 18
Données en coupe longitudinale 133
Données en coupe transversale 83, 133
Droite de régression 34
Droite de régression de l'échantillon 11, 86
Droite de régression de la population 12

E

Earning Yield 357
Effet fixe temporel 276
Effet marginal 181, 186, 190, 247, 411, 416, 423, 442
Effet revenu 432
Effet substitution 432
Effets aléatoires individuels 307
Effets fixes individuels 308
Elasticité 181, 189, 398
Elasticité partielle de la demande 223
Elasticité partielle de substitution d'Allen 223, 232, 241
Équation d'analyse de la variance 15, 49
Équation d'équilibre 245, 248
Équation d'Euler 313, 314, 316, 318, 319, 320, 325, 326, 329
Équation de comportement 245
Équilibre de long terme 343, 381
Erreur de mesure 150
Erreur de prévision 20
Erreur stochastique 11
Estimateur Between 310
Estimateur BLUE 16, 47, 136, 197
Estimateur convergent 12, 85
Estimateur Doubles Moindres Carrés 208, 249, 253

- Estimateur efficace 12
 Estimateur Maximum de Vraisemblance 126, 408, 416, 421
 Estimateur Moindres Carrés Généralisés 131, 198, 210
 Estimateur Moindres Carrés Ordinaires 13, 45, 305
 Estimateur Moindres Carrés Pondérés 136
 Estimateur Moindres Carrés Quasi Généralisés 132, 137, 143, 307
 Estimateur non biaisé 12
 Estimateur super convergent 344
 Estimateur Triples Moindres Carrés 209, 249, 257
 Estimateur Variables Instrumentales 147, 151, 208, 210, 306, 309
 Estimateur Within 273, 274, 275, 287, 308, 310
- F**
- Filtre de Hodrick et Prescott 105, 109
 Fonction d'autocorrélation 334
 Fonction d'autocorrélation partielle 334
 Fonction de vraisemblance 126
 Fonction homothétique 222
 Forme fonctionnelle flexible 224
 Forme réduite du modèle 181, 201
 Forme structurelle du modèle 181, 201, 249, 253
- H**
- Hypothèse de normalité 125
- I**
- Identification 379
 Identité comptable 249
 Intervalle de confiance 19, 29, 35, 37, 92
- L**
- Lemme de Shepard 220
- M**
- Marchés efficients 353
 Mean Absolute Error 64
 Mean Error 64
 Mécanisme correcteur d'erreurs 397
 MEDAF 158
 Méthode de Cochrane-Orcutt 143, 169
 Méthode de Engel et Granger 343
 Méthode de Hildreth-Lu 144, 170
 Méthode de Johansen 347, 380
 Méthode de Prais-Winsten 143
 Méthode de Zellner 198, 238
 Méthode de Zellner itérative 200, 240
 Méthode des Moments Généralisés 152, 279, 320, 323
 Méthode SUR 198
 Modèle à correction d'erreur 345, 346, 349, 360, 381
 Modèle à effets fixes 270
 Modèle à effets individuels aléatoires 283
 Modèle à équations apparemment non reliées 197
 Modèle à équations simultanées 201
 Modèle à retards échelonnés de Koyck 178
 Modèle ARMA 333
 Modèle autorégressif (AR) 187, 253, 333
 Modèle d'ajustement partiel 179, 246
 Modèle de moyenne mobile (MA) 333
 Modèle de Nerlove 179, 186
 Modèle de régression censuré 419
 Modèle de régression tronqué 413, 415
 Modèle de sélection 422
 Modèle dynamique 151, 187, 277, 289
 Modèle linéaire Généralisé 131
 Modèle Logit 407, 435, 439
 Modèle Probit 406, 435, 439, 442
 Modèle quasi-différencié 143, 169
 Modèle statique 175
 Modèle Tobit 419, 445
 Modèle VAR 346, 380
 Modèles à Correction d'Erreur 345
- N**
- Niveau de significativité 35, 83, 90, 92
- O**
- Output gap 101

P

PIB potentiel 101, 105
 Points de retournement 259
 Prédiction 12, 20, 62, 91
 Prévision 144
 Price Earning Ratio 356
 Prime de risque du marché 159, 357
 Probabilité limite 48
 Problème d'endogénéité 306, 309
 Procédure d'Heckman 423, 446
 Processus stationnaire 334

Q

Qualité prédictive du modèle 62

R

Ratio de Mills 416, 446
 Régression simple 35
 Relation de cointégration 347, 360, 368, 381, 389, 393
 Rendements d'échelle constants 222
 Résidu 12, 34, 48
 Résidus récursifs 117, 60
 Risque diversifiable 160
 Risque systématique 160
 Root Mean Square Error 64, 112

S

Salaire de réserve 430, 434
 Semi élasticité 398, 443
 Série intégrée 337, 386
 Séries cointégrées 343
 Séries stationnaires en différences (DS) 337
 Séries trend stationnaire (TS) 338
 Simulations Dynamiques 262
 Simulations Statiques 259
 Statistique de Box-Pierce 336
 Statistique de Fisher 20, 115, 168
 Statistique de Ljung-Box 336, 383, 389
 Statistique de Student 19, 115, 168
 Statistique de Theil 64, 264

T

Taux marginal de substitution 429
 Test d'exogénéité 188
 Test d'exogénéité faible 349, 391, 400
 Test d'Hausman 149, 149, 152, 188, 214, 281, 289
 Test d'homoscédasticité de Breush-Pagan 134
 Test d'homoscédasticité de Golfeld-Quandt 135
 Test d'homoscédasticité de White 133
 Test de Breusch-Pagan 288, 307
 Test de contraintes 38, 52, 53, 84
 Test de contraintes de sur-identification 212
 Test de contraintes linéaires sur les paramètres du modèle 52
 Test de Dickey Fuller 339
 Test de Dickey Fuller Augmenté 339, 366, 385
 Test de Fisher 88
 Test de l'hypothèse de normalité des aléas 125
 Test de la présence d'effets individuels fixes 275
 Test de la Trace 348, 390
 Test de la valeur propre maximale 349, 390
 Test de modèle emboîté 89, 387
 Test de non autocorrélation Durbin-Watson 169, 138, 169, 253, 369
 Test de non-autocorrélation statistique h de Durbin 139, 187, 253
 Test de normalité 126, 389
 Test de Sargan 153, 324, 325, 326
 Test de Sargan-Hansen 281
 Test de significativité des paramètres 19, 35, 38, 53, 83
 Test de significativité globale de la régression 38, 56, 83
 Test de Wald 127
 Test du CUSUM 60
 Test du CUSUM of Square 61
 Test du multiplicateur de Lagrange 129, 129, 134, 383, 389
 Test du ratio de vraisemblance 128, 128, 212, 387, 409

Test KPSS 341
Théorie de la dualité 220
Trend déterministe 338
Trend stochastique 338

V

Variable aléatoire 34
Variable binaire 405, 441
Variable endogène 11, 43, 249
Variable exogène 11, 44, 249

Variable explicative aléatoire 146
Variable indicatrice 65, 87, 92, 190, 381
Variable latente 405
Variable prédéterminée 213, 249
Variable qualitative 412
Vecteur de cointégration 343, 347
Vitesse d'ajustement 345, 347, 393, 397

W

Within 320, 322, 323

TABLE DES MATIÈRES

Avant-propos	5
1 Objectifs du livre	5
2 À qui s'adresse l'ouvrage	6
3 Contenu et plan de l'ouvrage	6
4 Utilisation du manuel et niveaux de difficulté	7
5 Remerciements	8
CHAPITRE 1	
Le modèle de régression linéaire simple	9
1 Présentation du modèle de régression linéaire simple	11
2 Estimation avec la méthode des Moindres Carrés Ordinaires	13
2.1 Hypothèses sur le modèle de régression linéaire simple	13
2.2 Estimateur des Moindres Carrés Ordinaires (MCO)	14
2.3 Propriétés de l'estimateur des MCO	16
3 Intervalles de Confiance et test de significativité des paramètres	18
4 Réalisation de prédictions et intervalle de confiance des prédictions ..	20
CHAPITRE 2	
Épargne, investissement et mobilité du capital	23
1 Introduction	25
2 Cadre d'analyse	26
2.1 Les relations macroéconomiques en économie ouverte	26
2.2 Spécification du modèle	27
2.3 Principaux résultats de Feldstein et Horioka	28
3 Évaluation de la relation épargne-investissement	30
3.1 Présentation de la base de données	30
3.2 Évaluation du degré de mobilité du capital à long terme	30
4 Éléments de corrigé	33

CHAPITRE 3

Le modèle de régression linéaire multiple	41
1 Présentation du modèle	43
1.1 <i>Estimateur des Moindres Carrés Ordinaires (MCO)</i>	46
1.1.1 Propriétés finies de l'estimateur des MCO	46
1.1.2 Propriétés asymptotiques de l'estimateur des MCO	47
1.2 <i>Équation d'analyse de la variance et qualité de l'ajustement</i>	48
2 Tests statistiques usuels	50
2.1 <i>Distribution de l'estimateur des MCO</i>	50
2.2 <i>Test de contraintes linéaires sur les paramètres du modèle</i>	52
2.2.1 Test d'une contrainte	52
2.2.2 Test de plusieurs contraintes	53
2.3 <i>Test de changement structurel test de Chow</i>	57
2.4 <i>Tests de stabilité temporelle basés sur les résidus récurrents</i>	58
2.4.1 Les résidus récurrents	59
2.4.2 Test du CUSUM	60
2.4.3 Test du CUSUM of Square	61
3 Prédications et qualité prédictive d'un modèle estimé	62
4 Les variables indicatrices dans les modèles économétriques	65

CHAPITRE 4

Convergence du PIB par tête des économies européennes	67
1 Introduction	69
2 Le modèle de Solow	69
2.1 <i>Structure du modèle</i>	70
2.2 <i>Analyse de l'équilibre de long terme du modèle</i>	71
2.3 <i>Dynamique d'ajustement vers l'équilibre</i>	73
2.3.1 L'ajustement vers l'équilibre	74
2.3.2 Définition de la convergence d'un groupe d'économies : convergence absolue versus convergence conditionnelle	75
3 Analyse empirique	76
3.1 <i>L'équation d'équilibre</i>	77
3.2 <i>L'équation de convergence</i>	78
3.3 <i>Analyse de la convergence</i>	80
4 Éléments de corrigé	83
5 Annexes	94
5.1 <i>Annexe 1 : Calcul du taux de croissance du stock de capital à l'équilibre</i>	94
5.2 <i>Annexe 2 : Détermination du taux de croissance du stock de capital par unité de travail effectif autour de l'équilibre</i>	94

5.3	<i>Annexe 3 : Solution de l'équation différentielle de taux de croissance de l'output par unité de travail effectif</i>	95
-----	-------------------------------------------------------------------------------------------------------------------------------	----

CHAPITRE 5

Les politiques monétaires européennes et la règle de Taylor		97
1	Introduction	99
2	Règles monétaires et règle de Taylor	100
2.1	<i>Règles monétaires actives</i>	100
2.2	<i>Règle de Taylor</i>	101
3	Adéquation des politiques monétaires européennes à la règle	104
3.1	<i>Données</i>	104
3.2	<i>Construction des variables d'intérêt : PIB potentiel, output gap et croissance potentielle</i>	105
3.2.1	La notion de PIB potentiel	105
3.2.2	Évaluation par la méthode de filtrage de Hodrick et Prescott	105
3.3	<i>Évaluation empirique</i>	106
4	Éléments de corrigé	109

CHAPITRE 6

Les hypothèses du modèle de régression linéaire multiple		123
1	La normalité des aléas	125
1.1	<i>Test de l'hypothèse de normalité des aléas</i>	125
1.2	<i>Le maximum de vraisemblance</i>	126
1.2.1	L'estimateur du maximum de vraisemblance	126
1.2.2	Tests d'hypothèses basés sur l'estimateur du maximum de vraisemblance	127
1.2.3	Test de Wald	127
1.2.4	Test du ratio de vraisemblance	128
1.2.5	Test du multiplicateur de Lagrange	129
2	L'homoscédasticité et la non-autocorrélation des aléas	131
2.1	<i>L'estimateur des MCG</i>	131
2.1.1	L'estimateur des MCO et le modèle linéaire généralisé	131
2.1.2	L'estimateur des MCG	131
2.2	<i>Cas de l'hétéroscédasticité des aléas</i>	133
2.2.1	Application des tests statistiques d'homoscédasticité des aléas	133
2.2.2	Estimation des paramètres en cas d'hétéroscédasticité des aléas	136
2.3	<i>Cas de l'autocorrélation d'ordre 1 des aléas</i>	137
2.3.1	Test de l'absence d'autocorrélation d'ordre 1 des aléas : le test de Durbin-Watson	138
2.3.2	Détermination de la matrice de variance-covariance des aléas	141
2.3.3	Estimation des paramètres en cas d'autocorrélation des aléas	142
2.3.4	Prévision en présence d'autocorrélation des aléas	144

2.4	<i>La correction de White</i>	144
3	Les variables explicatives du modèle	145
3.1	<i>Variables explicatives aléatoires et estimateur des MCO</i>	146
3.2	<i>L'estimateur des Variables Instrumentales</i>	147
3.3	<i>Test d'Hausman</i>	149
3.4	<i>Applications de l'estimateur des Variables Instrumentales</i>	150
3.4.1	Cas d'une erreur de mesure sur une variable explicative	150
3.4.2	Cas d'un modèle dynamique avec autocorrélation des aléas	151
3.5	<i>L'estimateur de la Méthode des Moments Généralisés</i>	152

CHAPITRE 7

	Le modèle d'évaluation des actifs financiers	155
1	Introduction	157
2	Le MEDAF : hypothèses et relations d'équilibre	158
2.1	<i>Hypothèses</i>	158
2.2	<i>Relations d'équilibre</i>	159
2.2.1	La droite des marchés de capitaux	159
2.2.2	La relation d'équilibre du MEDAF	160
3	Estimations et validation sur les sociétés du CAC 40	161
3.1	<i>Passage du modèle théorique au modèle économétrique</i>	161
3.2	<i>Données</i>	162
3.3	<i>Mise en œuvre de l'évaluation</i>	163
4	Éléments de corrigé	165

CHAPITRE 8

	La fonction d'offre de viande bovine	173
1	Introduction	175
2	Spécification de la fonction d'offre	175
2.1	<i>Dérivation de la fonction d'offre dans un cadre statique</i>	175
2.2	<i>Dynamisation de la fonction d'offre : l'exemple du modèle de Nerlove</i>	177
2.2.1	Spécification des anticipations	177
2.2.2	Existence de coûts d'ajustement : le concept d'ajustement partiel	179
2.2.3	Le modèle général de Nerlove	179
2.3	<i>Prise en compte des caractéristiques de la production de bœuf</i>	182
2.3.1	Existence d'une politique d'intervention	182
2.3.2	Caractère dynamique de la production	182
3	Analyse empirique	183
4	Éléments de corrigé	185

CHAPITRE 9

Les modèles à plusieurs équations	195
1 Les modèles à équations apparemment non reliées	197
2 Présentation des modèles à équations simultanées	200
3 Le problème de l'identification	203
3.1 <i>Présentation du problème de l'identification</i>	203
3.2 <i>Identification et FRM</i>	203
3.3 <i>Les conditions d'ordre et de rang</i>	206
4 Estimation d'un modèle à équations simultanées	207
4.1 <i>La méthode des DMC</i>	208
4.2 <i>La méthode des TMC</i>	209
5 Tests de spécification	212
5.1 <i>Test sur les contraintes de sur-identification</i>	212
5.2 <i>Test d'exogénéité des variables explicatives : le test d'Hausman</i>	213

CHAPITRE 10

Modélisation d'un système de demande de facteurs de production	217
1 Introduction	219
2 Approche par la fonction de coût	219
2.1 <i>Définition et propriétés fondamentales d'une fonction de coût</i>	220
2.1.1 Propriété issue de la théorie de la dualité	220
2.1.2 La contrainte de monotonie et la contrainte de concavité	221
2.1.3 La contrainte d'homogénéité de degré par rapport aux prix	221
2.2 <i>Propriétés supplémentaires</i>	222
2.2.1 Hypothèse d'homothétie de la fonction de coût	222
2.2.2 Hypothèse de rendements d'échelle constants	222
2.3 <i>Interprétation des résultats</i>	223
3 Illustration en spécifiant une fonction translog	224
3.1 <i>Propriétés de la fonction de coût</i>	225
3.1.1 Les contraintes de symétrie	225
3.1.2 La contrainte d'homogénéité de degré 1 par rapport aux prix	225
3.2 <i>Propriétés particulières</i>	226
3.2.1 Hypothèse d'homothétie de la fonction de coût	226
3.2.2 Hypothèse de rendements d'échelle constants	227
4 Estimation de la fonction translog	227
4.1 <i>Dérivation du système de parts de dépense</i>	227
4.2 <i>Illustration dans le cas de trois facteurs</i>	229

4.2.1	Contraintes de symétrie	230
4.2.2	Contrainte d'homogénéité de degré 1 par rapport aux prix	230
4.2.3	Contraintes d'additivité	232
4.2.4	Contrainte d'homothétie	232
4.2.5	Interprétation des coefficients : calcul des élasticités	232
4.2.6	Estimation des paramètres	233
5	Analyse empirique	233
6	Éléments de corrigé	234

CHAPITRE 11

	Modélisation d'un équilibre de marché	243
1	Introduction	245
2	Spécification du modèle	245
2.1	<i>Équations de comportement</i>	245
2.1.1	Équation d'offre	245
2.1.2	Équation de demande	246
2.1.3	Équation de transmission de prix	248
2.1.4	Équation du commerce extérieur	248
2.2	<i>La forme structurelle du modèle</i>	248
2.2.1	Équation d'équilibre	248
2.2.2	Écriture du modèle économétrique	248
2.3	<i>Estimation du modèle</i>	249
3	Analyse empirique	250
4	Éléments de corrigé	251

CHAPITRE 12

	Introduction à l'économétrie des données de panel	267
1	Présentation du modèle de base	269
2	Le modèle à effets fixes individuels	270
2.1	<i>Estimation du modèle</i>	270
2.2	<i>Test de la présence d'effets individuels fixes</i>	275
2.3	<i>Généralisation du modèle à effets fixes</i>	276
2.4	<i>Modèles à effets individuels fixes dynamiques</i>	277
2.5	<i>Tests sur les modèles à effets individuels fixes dynamiques</i>	281
3	Le modèle à effets individuels aléatoires	283
3.1	<i>Estimation du modèle</i>	283
3.2	<i>Test des effets individuels aléatoires : le test de Breush Pagan</i>	287
3.3	<i>Modèles à effets individuels aléatoires corrélés avec les régresseurs</i>	288
3.4	<i>Modèles à effets individuels aléatoires dynamiques</i>	289

CHAPITRE 13

Analyse du marché du travail non qualifié	293
1 Introduction	295
2 Les modèles de négociation	296
3 Le modèle du droit à gérer	298
3.1 <i>Spécification de l'équation de salaire</i>	298
3.1.1 Le comportement de l'entreprise	298
3.1.2 Le comportement du syndicat	298
3.1.3 Programme de négociation des salaires	300
3.2 <i>Évaluation du modèle du droit à gérer</i>	301
3.2.1 Les données issues des DADS	301
3.2.2 Les données issues de l'ESE	302
3.2.3 Les données issues des comptes nationaux annuels	302
4 Éléments de corrigés	305

CHAPITRE 14

Imperfection du marché du capital et décisions d'investissement	311
1 Introduction	313
2 Le modèle microéconomique	314
3 Spécification économétrique	318
4 Application empirique : Estimations sur données de panel françaises	319
5 Éléments de corrigé	320
6 Annexe	328
6.1 <i>Dérivation des équations (14.3), (14.4) et (14.6)</i>	328
6.2 <i>Équation (14.3)</i>	329
6.3 <i>Équation (14.4)</i>	329
6.4 <i>Équation (14.6)</i>	329

CHAPITRE 15

Econométrie des séries non stationnaires	331
1 Les modèles de série temporelle univariés	333
1.1 <i>Exemples</i>	333
1.2 <i>Définition de la stationnarité</i>	334
1.3 <i>Modélisation d'une série temporelle</i>	334
2 Processus non stationnaire et racine unitaire	337

2.1	<i>Les séries non stationnaires</i>	337
2.1.1	Séries DS	337
2.1.2	Séries TS	338
2.2	<i>Tests de racine unitaire</i>	339
2.2.1	Test de Dickey Fuller (test DF)	339
2.2.2	Test de Dickey Fuller Augmenté (test ADF)	339
2.2.3	Stratégie du test ADF	340
2.2.4	Test de Kwiatkowski, Phillips, Schmidt et Shin (KPSS)	341
3	<i>Cointégration et MCE : l'approche de Engel et Granger</i>	343
3.1	<i>Tests de Cointégration</i>	343
3.1.1	Cointégration : présentation	343
3.1.2	La procédure de Engel et Granger	343
3.2	<i>Modèles à Correction d'Erreur</i>	345
4	<i>Cointégration et MCE : l'approche de Johansen</i>	346
4.1	<i>Modèles à Correction d'Erreur à plusieurs équations</i>	346
4.2	<i>Tests de cointégration : test de la trace et test de la valeur propre maximale</i>	348
4.2.1	Test de la Trace	348
4.2.2	Test de la valeur propre maximale	349
4.3	<i>Test d'exogénéité faible</i>	349
4.4	<i>Test d'hypothèse sur les relations de cointégration</i>	350

CHAPITRE 16

La valorisation des marchés actions	351
1 Introduction	353
2 Théorie financière et prix d'équilibre des actions	354
2.1 <i>Recherche d'un prix d'équilibre à partir d'une relation d'arbitrage</i>	354
2.2 <i>Une première approche de la valorisation</i>	357
3 La construction d'un indicateur de valorisation	358
3.1 <i>Principes</i>	358
3.1.1 Identification d'une relation de cointégration entre Earning Yield et taux d'intérêt	359
3.1.2 Retour à l'équilibre	360
3.1.3 Indicateur de valorisation, niveaux théoriques de l'indice et degré de surévaluation	361
3.2 <i>Application sur les marchés boursiers américain et français</i>	362
3.2.1 Données et profil des séries	362
3.2.2 Tests, estimations et construction de l'indicateur	364
4 Éléments de corrigé	365

CHAPITRE 17

La relation salaires-prix à long terme	373
1 Introduction	375
2 Cadre analytique	376
2.1 <i>Le modèle WS-PS</i>	376
2.1.1 La courbe PS	376
2.1.2 La courbe WS	378
2.1.3 Le modèle WS-PS	379
2.2 <i>Identification du modèle théorique</i>	379
3 Stratégie d'estimation	380
4 Éléments de corrigé	383

CHAPITRE 18

Économétrie des variables qualitatives	403
1 Modèles à variables qualitatives binaires	405
1.1 <i>Présentation des modèles à variables qualitatives binaires</i>	405
1.2 <i>Les modèles probit et logit</i>	406
1.3 <i>Estimation des modèles probit et logit</i>	408
1.3.1 La méthode du maximum de vraisemblance	408
1.3.2 Tests statistiques et mesure de la qualité de l'ajustement	410
1.4 <i>Interprétation des résultats</i>	411
1.4.1 Les effets marginaux lorsque les variables explicatives sont continues	411
1.4.2 Les effets marginaux lorsque les variables explicatives sont qualitatives	412
1.5 <i>Comparaison des résultats des modèles probit et logit</i>	413
2 Modèles à variables dépendantes limitées	413
2.1 <i>Modèle de régression tronqué</i>	413
2.1.1 Distribution tronquée	413
2.1.2 Le modèle de régression tronqué	415
2.1.3 La fonction de vraisemblance	416
2.1.4 Interprétation des résultats	417
2.2 <i>Modèle de régression censuré --- Modèle Tobit</i>	417
2.2.1 Distribution d'une variable censurée	417
2.2.2 Le modèle de régression censuré : le modèle Tobit simple	419
2.2.3 Évaluation des paramètres du modèle	421
2.2.4 Interprétation des résultats	421
2.3 <i>Modèle de sélection</i>	422

CHAPITRE 19

Déterminants des décisions d'offre de travail	425
1 Introduction	427
2 Théorie microéconomique de la participation au marché du travail .	427
2.1 <i>Les déterminants de l'arbitrage entre travail et loisir</i>	427
2.1.1 Le programme de comportement	427
2.1.2 Interprétation des conditions du premier ordre	429
2.2 <i>Critère de participation au marché du travail : illustration graphique</i>	430
2.3 <i>Illustration paramétrique</i>	432
3 Approche économétrique de l'offre de travail	434
3.1 <i>Définition du salaire de réserve</i>	434
3.2 <i>Définition de l'offre de travail</i>	435
3.3 <i>Équation de salaire</i>	435
3.4 <i>Modèle de participation</i>	435
4 Analyse empirique	436
5 Éléments de corrigé	437
Index	449

OUVERTURES ◀▶ ÉCONOMIQUES

- ALLEGRET J.-P., LE MERRER P., *Économie de la mondialisation. Opportunités et fractures*
- ANDERSON R. D., SWEENEY J. D., WILLIAMS A. TH., *Statistiques pour l'économie et la gestion*, traduction de la 2^e édition anglaise par Cl. Borsenberger
- BÉNASSY-QUÉRÉ A., CŒURÉ B., JACQUET P., PISANI-FERRY J., *Politique économique*. 2^e édition
- BEREND IVAN T., *Histoire économique de l'Europe du XX^e siècle*
traduction de la 1^{re} édition anglaise par Amandine Nguyen
- BERGSTROM T., VARIAN H., *Exercices de microéconomie - 1. Premier cycle. Notions fondamentales*. 3^e édition
traduction de la 5^e édition américaine par A. Marciano
- BERGSTROM T., VARIAN H., *Exercices de microéconomie - 2. Premier cycle et spécialisation*. 2^e édition française
traduction de la 5^e édition américaine par J.-M. Baland, S. Labenne et Ph. Van Kerm
avec la collaboration scientifique d'A. Marciano.
- BISMANS F., *Mathématiques pour l'économie - Volume 1. Fonctions d'une variable réelle*
- BOUTILLIER S., PEAUCELLE I., UZUNIDIS D., *L'économie russe depuis 1990*
- BURDA M., WYPLOSZ C., *Macroéconomie. À l'échelle européenne*. 4^e édition
traduction de la 4^e édition anglaise par Stanislas Standaert
- CADORET I., BENJAMIN C., MARTIN F., HERRARD N., TANGUY S., *Économétrie appliquée*. 2^e édition
Méthodes, Applications, Corrigés
- CAHUC P., ZYLBERBERG A., *Le marché du travail*
- CAHUC P., ZYLBERBERG A., *Économie du travail. La formation des salaires et les déterminants du chômage*
- CARLTON D. W., PERLOFF J. M., *Économie industrielle*, traduction de la 2^e édition américaine par F. Mazerolle.
2^e édition
- CARTELIER J., *L'économie de Keynes*
- CAVES R.E., FRANKEL J. A., JONES R. W., *Commerce international et paiements*,
traduction de la 9^e édition américaine par M. Chiroleu-Assouline
- CAYATTE J.-L., *Introduction à l'économie de l'incertitude*
- COLLECTIF, *Économie sociale. Enjeux conceptuels, insertion par le travail et services de proximité*
- COMMISSARIAT GÉNÉRAL DU PLAN, *L'intégration régionale.*
Une nouvelle voie pour l'organisation de l'économie mondiale ?
- CORNET B. et TULKENS H. (Éds), *Modélisation et décisions économiques*
- CÔTÉ D., *Les holdings coopératifs. Évolution ou transformation définitive ?*
- CUTHBERTSON K., *Économie financière quantitative. Actions, obligations et taux de change*,
traduction de la 1^{re} édition anglaise par C. Puibasset
- DARREAU Ph., *Croissance et politique économique*
- DEFOURNY J., *Démocratie coopérative et efficacité économique. La performance comparée des SCOP françaises*
- DEFOURNY J., DEVELTERE P., FONTENEAU B. (Éds), *L'économie sociale au Nord et au Sud*
- DEFOURNY J., MONZON CAMPOS J.L. (Éds), *Économie sociale/The Third Sector. Entre économie capitaliste et économie publique/Cooperative Mutual and Non-profit Organizations*
- DE GRAUWE P., *Économie de l'intégration monétaire*, traduction de la 3^e édition anglaise par M. Donnay
- DE GRAUWE P., *La monnaie internationale. Théories et perspectives*,
traduction de la 2^e édition anglaise par M.-A. Sénégal

- DE KERCHOVE A.-M., GEELS TH., VAN STEENBERGHE V., *Questions à choix multiple d'économie politique*.
3^e édition
- DE MELO J., GREETHER J.-M., *Commerce international. Théories et applications*
- DEVELTERE P., *Économie sociale et développement*.
Les coopératives, mutuelles et associations dans les pays en voie de développement
- DRÈZE J., *Pour l'emploi, la croissance et l'Europe*
- DUPRIEZ P., OST C., HAMAIDE C., VAN DROOGENBROECK N., *L'économie en mouvement*.
Outils d'analyse de la conjoncture. 2^e édition
- ESCH L., *Mathématique pour économistes et gestionnaires*. 3^e édition
- ESSAMA-NSSAH B., *Inégalité, pauvreté et bien-être social. Fondements analytiques et normatifs*
- GAZON J., *Politique industrielle et industrie*
Volume I. Controverses théoriques. Aspects légaux et méthodologie
- GILLIS M. et al., *Économie du développement*, traduction de la 4^e édition américaine par B. Baron-Renault
- GOMEZ P.-Y., KORINE HARRY, *L'entreprise dans la démocratie, Une théorie politique du gouvernement des entreprises*
- GUJARATI D. N., *Économétrie*, traduction de la 4^e édition américaine par B. Bernier
- HARRISON A., DALKIRAN E., ELSEY E., *Business international et mondialisation. Vers une nouvelle Europe*
- HEERTJE A., PIERETTI P., BARTHÉLEMY PH., *Principes d'économie politique*. 4^e édition
- JACQUEMIN A., TULKENS H., MERCIER P., *Fondements d'économie politique*. 3^e édition
- JACQUEMIN A., PENCH L. R. (Éds), *Pour une compétitivité européenne*.
Rapports du Groupe Consultatif sur la Compétitivité
- JALLADEAU J., *Introduction à la macroéconomie. Modélisations de base et redéploiements théoriques contemporains*.
2^e édition
- JALLADEAU J., DORBAIRE P., *Initiation pratique à la macroéconomie. Études de cas, exercices et QCM*. 2^e édition
- JASKOLD GABSZEWICZ J., *Théorie microéconomique*. 2^e édition
- JONES Ch. I., *Théorie de la croissance endogène*, traduction de la 1^{re} édition américaine par F. Mazerolle
- JURION B., *Économie politique*. 3^e édition
- JURION B., LECLERCQ A., *Exercices d'économie politique*
- KOHLI U., *Analyse macroéconomique*
- KRUGMAN P. R. et OBSTFELD M., *Économie internationale*. 4^e édition
traduction de la 6^e édition américaine par A. Hannequart et F. Leloup
- KRUGMAN P., *L'économie auto-organisatrice*, traduction de la 1^{re} édition américaine par F. Leloup. 2^e édition
- KRUGMAN P., *Microéconomie*, traduction de la 2^e édition américaine par L. Baechler
- LANDAIS B., *Leçons de politique budgétaire*
- LANDAIS B., *Leçons de politique monétaire*
- LECAILLON J.-D., LE PAGE J.-M., OTTAVI CHR., *Économie contemporaine. Analyses et diagnostics*. 2^e édition
- LEMOINE M., MADIÈS P., MADIÈS T., *Les grandes questions d'économie et finance internationales*.
Décoder l'actualité
- LEROUX A., MARCIANO A., *Traité de philosophie économique*
- LESUEUR J.-Y., SABATIER M., *Microéconomie de l'emploi. Théories et applications*
- LÖWENTHAL P., *Une économie politique*

- MANKIW G. N., *Macroéconomie*, traduction de la 6^e édition américaine par Jihad C. El Naboulsi. 4^e édition
- MANSFIELD E., *Économie managériale. Théorie et applications*,
traduction et adaptation de la 4^e édition américaine par B. Jérôme
- MASSÉ G., THIBAUT FR., *Intelligence économique. Un guide pour une économie de l'intelligence*
- MARCIANO A., *Éthiques de l'économie. Introduction à l'étude des idées économiques*
- MILGROM P., ROBERTS J., *Économie, organisation et management*
- MONNIER L., THIRY B. (Éds), *Mutations structurelles et intérêt général
Vers quels nouveaux paradigmes pour l'économie publique, sociale et coopérative ?*
- NORRO M., *Économies africaines. Analyse économique de l'Afrique subsaharienne*. 2^e édition
- PERKINS D. H., RADELET S., LINDAUER D. L., *Économie du développement*. 3^e édition
- PROMEURO. *L'Euro pour l'Europe. Des monnaies nationales à la monnaie européenne*. 2^e édition
- RASMUSEN E., *Jeux et information. Introduction à la théorie des jeux*,
traduction de la 3^e édition anglaise par F. Bismans
- SALVATORE D. C., *Économie internationale*,
traduction de la de la 9^e édition américaine par Fabienne Leloup et Achille Hannequart
- SHAPIRO C., VARIAN H. R., *Économie de l'information. Guide stratégique de l'économie des réseaux*,
traduction de la 1^{re} édition américaine par F. Mazerolle
- SIMON C. P., BLUME L., *Mathématiques pour économistes*. traduction de la 1^{re} édition américaine par
G. Dufrenot, O. Ferrier, M. Paul, A. Pirotte, B. Planes et M. Seris
- SINN G., SINN H. W., *Démarrage à froid. Une analyse des aspects économiques de l'unification allemande*,
traduction de la 3^e édition allemande par C. Laurent
- STIGLITZ J. E., WALSH C. E., Lafay J.-D., *Principes d'économie moderne*. 3^e édition,
traduction de la 3^e édition américaine par F. Mayer
- VARIAN H., *Introduction à la microéconomie*. 6^e édition, traduction de la 7^e édition américaine par B. Thiry
- VARIAN H., *Analyse microéconomique*, traduction de la 3^e édition américaine par J.-M. Hommet. 2^e édition
- VAN DER LINDEN B. (Éd.), *Chômage. Réduire la fracture*
- ZÉVI A., MONZÓN CAMPOS J.-L., *Coopératives, marchés, principes coopératifs*

Une approche originale et très complète de la pratique des méthodes économétriques en prenant pour application des thèmes économiques diversifiés.

Économétrie appliquée

Cet ouvrage propose une **formation à la pratique de l'économétrie**. Il comprend à la fois un rappel de cours clair et pédagogique sur les **méthodes économétriques** (de base et avancées), et des **applications** sur des **thèmes économiques diversifiés** correspondant à des préoccupations actuelles. Il s'agit de montrer comment, de manière pratique, les différentes méthodes économétriques sont mises en œuvre. Cette seconde édition intègre de manière plus détaillée **l'économétrie des données de panel via un chapitre spécifique**, deux applications sur données de panel sont proposées : la première dans un cadre statique, la seconde dans un cadre dynamique

L'accent est mis sur la construction de modèles, à partir de fondements micro-économiques ou macro-économiques, et leur évaluation empirique. La démarche proposée est la suivante : (i) définition du modèle à estimer, (ii) création d'une base de données, (iii) application de la méthode d'estimation appropriée, (iv) analyse et interprétation des résultats.

L'originalité principale du manuel est de **travailler sur des données réelles** et de fournir un **corrigé complet pour chaque exercice d'application**.

Le lecteur peut réaliser lui-même les exercices en téléchargeant les données et les programmes développés par les auteurs avec notamment le logiciel économétrique *WinRats*. Des programmes développés sous SAS et *Eviews* sont également disponibles.

Isabelle Cadoret

Professeure de Sciences Économiques à l'Université de Rennes 1 et membre du Centre de Recherche en Économie et Management.

Catherine Benjamin

Directeur de recherche à l'INRA.

Franck Martin

Professeur de Sciences Économiques à l'université de Rennes 1, membre du centre de recherche en économie et management (Unité mixte de recherche du CNRS).

Nadine Herrard

Ingénieur d'études à l'INRA.

Steven Tanguy

Ingénieur de recherche au Rectorat de Limoges.

ECAPMO
ISBN 978-2-8041-5976-4
ISSN 2030-501X

www.deboeck.com



9 782804 159764