mars 1988

N° 8809

HETEROGENEITE

I. LE CAS LINEAIRE

GOURIEROUX C. (*) et PEAUCELLE I. (*)

(*) CEPREMAP.

RESUME

HETEROGENEITE : I. LE CAS LINEAIRE

Dans les modèles spatio-temporels, l'introduction de variables explicatives se révèle souvent insuffisante pour prendre en compte toutes les différences individuelles et celles oubliées devraient naturellement conduire l'économètre à retenir des coefficients des relations dépendantes de l'individu. Lorsque cette hétérogénéité est entièrement ou partiellement oubliée au moment de l'estimation, les paramètres estimés peuvent ou non selon le modèle utilisé s'interpréter en fonction des paramètres désagrégés correspondants. Dans cet article nous analysons l'existence et l'importance des biais d'hétérogénéité pour des modèles linéaires.

Mots-clefs : Données spatio-temporelles, hétérogénéité, biais d'estimation.

ABSTRACT

HETEROGENEITY : I. THE LINEAR CASE

In the case of panel data, the introduced explanatory variables are often not sufficient for describing all the individual caracteristics and it may be useful to consider relation with individual varying coefficients. When this heterogeneity is even partially omitted, the estimated parameters may or not have some interpretations in term of the associated desaggregated parameters. In this paper we analyse the existence and the importance of heterogeneity biases for linear models.

Keywords : Panel data, heterogeneity, estimation bias.

JEL : 211.

1. INTRODUCTION

Le fait que les comportements dépendent de l'individu considéré est habituellement pris en compte dans les modèles microéconomiques en introduisant dans les relations des variables explicatives auxiliaires décrivant certaines caractéristiques de cet individu. Ainsi une fonction de consommation individuelle, qui est à la base des évaluations d'élasticités revenu ou d'élasticités prix, ne comporte pas les seules variables explicatives revenus et prix, mais généralement elle en contient un certain nombre d'autres comme la taille du ménage ou la catégorie socio-professionnelle du chef de ménage.

Cette introduction de variables auxiliaires se révèle souvent insuffisante pour prendre en compte toutes les différences individuelles et celles oubliées devraient naturellement conduire l'économètre à retenir les coefficients des relations dépendantes de l'individu. Ainsi si on dispose de données spatio-temporelles doublement indexées par l'individu $i=1, \ldots, n$ et le temps $t=1, \ldots, T$, si la variable endogène est notée y_{it} et les k variables exogènes x_{it}, et si on retient une formulation linéaire, le modèle désagrégé devrait se présenter sous la forme :

$$(1.1) \quad y_{it} = a_i + x_{it}b_i + u_{it}, \quad Eu_{it} = 0, \qquad i=1, \ldots, n \; ; \; t=1 \ldots T$$

Dans tout le papier, on considère que ce modèle désagrégé est bien spécifié. On a implicitement admis pour simplifier que les coefficients a_i, b_i sont indépendants du temps, ce qui revient à privilégier l'hétérogénéité individuelle par rapport à l'hétérogénéité temporelle. Nous supposerons également pour simplifier les calculs que les erreurs u_{it} sont centrées, non corrélées, de même variance σ^2 et qu'elles sont non corrélées aux valeurs x_{it} des variables exogènes.

De telles formulations à coefficients variables (on dit aussi avec hétérogénéité) sont rarement retenues en pratique et on préfère généralement remplacer le modèle initial par des formulations approchées plus simples à manier. Parmi celles-ci les deux plus classiques sont l'approche microéconométrique par individu représentatif et l'approche macroéconomique du modèle agrégé (dite encore méthode inter).

... ... cas, on retient une spécification commune au niveau microéconométrique en supposant les coefficients constants pour les diverses unités. Le modèle s'écrit :

$$(1.2)\quad y_{it} = a + x_{it}\, b + v_{it} \qquad ; \; E\, v_{it}=0,\; i=1\ldots n,\; t=1\ldots T.$$

Ce modèle peut être vu comme cas particulier du modèle (1.1) correspondant à l'hypothèse d'homogénéité : $a_i=a$, $b_i=b$ $\forall$ i. L'optique macroéconométrique s'intéresse aux relations pouvant exister entre variables agrégées. Définissant pour simplifier ces dernières par sommation :

$Y_t = \sum_{i=1}^{n} y_{it}$, $X_t = \sum_{i=1}^{n} x_{it}$, $t=1, \ldots T$, le modèle agrégé est pris du type :

$$(1.3)\quad Y_t = \alpha + X_t\, \beta + \omega_t,\; E\, \omega_t = 0,\; t=1, \ldots T,$$

et n'est évidemment considéré que sur données temporelles.

Il est clair que, lorsque les coefficients individuels a_i, b_i sont constants (cas de l'homogénéité), ces deux formulations (1.2) et (1.3) sont compatibles avec le modèle désagrégé et les coefficients sont liés par les contraintes : $a_i = a = \frac{\alpha}{n}$, $b_i = b = \beta$.

Cette compatibilité existe aussi sous des conditions plus générales. Ainsi si nous supposons que les coefficients individuels a_i, b_i peuvent être considérés comme indépendants, extraits d'une même loi de moyenne A, B, et si de plus ils sont indépendants de valeurs des variables explicatives x_{it}, on peut écrire le modèle désagrégé sous la forme :

$$\begin{aligned} y_{it} &= a_i + x_{it} b_i + u_{it} \\ &= A + x_{it} B + u_{it} + a_i - A + x_{it}(b_i - B). \\ &= A + x_{it}\, B + v_{it}. \end{aligned}$$

Les hypothèses faites assurent que le terme résiduel :

$$v_{it} = u_{it} + a_i - A + x_{it}(b_i - B),$$

est de moyenne nulle et non corrélé avec les exogènes x_{it}. Le modèle désagrégé est alors compatible avec un modèle à individu représentatif du type : $y_{it} = A + x_{it}B + v_{it}$. Il l'est aussi avec un modèle agrégé, par sommation sur l'indice i des relations du modèle à individu représentatif.

En revanche lorsque l'hétérogénéité est présente et que la condition d'indépendance avec les variables explicatives n'est pas satisfaite, les deux modèles de l'individu représentatif et de la relation agrégée sont mal spécifiés. L'utilisation de ce modèle conduit alors souvent à des biais d'estimation.

De façon à éviter de tels biais, on essaye parfois de prendre en compte partiellement l'hétérogénéité, par exemple par l'intermédiaire du terme constant. Le <u>modèle à effet fixe</u> est ainsi défini par :

$$(1.4)\ y_{it} = a_i^* + x_{it}\, b^* + v_{it}^*,\ E\, v_{it}^* = 0,\ i=1, \dots n,\ t=1, \dots, T.$$

Le modèle agrégé associé est évidemment identique à (1.3). Ce modèle aussi apparaît mal spécifié si les coefficients b_i des variables explicatives diffèrent selon les individus.

Comme en pratique le paramètre d'intérêt est généralement le coefficient b^*, on remplace souvent pour l'estimation de b^* le modèle à effet fixe par un modèle transformé ne contenant plus de terme constant. Pour cela on peut considérer le <u>modèle en différence</u> :

$$(1.5)\ y_{i,t} - y_{i,t-1} = (x_{i,t} - x_{i,t-1})\, b^{**} + v_{it}^{**}.$$

On peut aussi introduire la <u>forme intra</u> :

$$(1.6)\ y_{it} - y_{i.} = (x_{it} - x_{i.})\, b^{***} + v_{it}^{***},$$

où : $y_{i.}$ par exemple désigne la moyenne des valeurs y_{it} sur les dates.

Dans cet article nous nous intéressons aux propriétés des estimateurs des moindres carrés ordinaires des coefficients b, β, b^{**}, b^{***} calculés à partir des modèles (1.2), (1.3), (1.5), (1.6), lorsque le vrai modèle contient de l'hétérogénéité.

Dans le paragraphe 2, on décrit des cas de mauvaises spécifications, en donnant des modèles où l'hétérogénéité est corrélée aux variables explicatives du modèle. Un premier exemple concerne les modèles comportant des variables endogènes retardées. La loi de l'endogène dépend des coefficients a_i, b_i. Comme parmi les variables explicatives figure l'endogène retardée, on en déduit que la loi des explicatives dépend également des coefficients.

Un autre exemple important est relatif aux fonctions de production. Considérons une fonction de productions homogènes de degré 1, par exemple du type COBB-DOUGLAS :

$$\mathrm{Log}\, \frac{y_{it}}{x_{1it}} = a_i + \sum_{k=2}^{K} b_{ki}\, \mathrm{Log}\, \frac{x_{kit}}{x_{1it}} + u_{it},$$

où y désigne l'output et x_k , k=1, ... k les quantités des divers inputs. Si l'entreprise optimise l'allocation de ses inputs en fonction des prix et de la fonction de production, les demandes de facteurs devraient être telles que :

$$\frac{x_{kit}}{x_{1it}} = D_k\, (P_{1t} \ldots P_{kt}, a_i, b_{2i}, \ldots, b_{ki}).$$

Il y a une dépendance entre les variables explicatives $\mathrm{Log}\, \frac{x_{kit}}{x_{1it}}$ et les coefficients a_i, b_{2i}, ..., b_{ki}.

Dans le paragraphe 3 nous explicitons les estimateurs des moindres carrés ordinaires $\hat{b}$, $\hat{\beta}$, $\hat{b}^{**}$, $\hat{b}^{***}$ calculé sur les divers modèles approchés et déterminons leurs limites b_∞ , β_∞ , b_∞^{**} , b_∞^{***} , lorsque le nombre d'observations tend vers l'infini. Toute la difficulté est alors dans l'interprétation des résultats obtenus, c'est-à-dire essentiellement de ces valeurs limites. Comme on souhaite souvent qu'elles aient une interprétation similaire aux coefficients désagrégés b_j, il est important de regarder si ces bruits peuvent ou non s'interpréter comme des moyennes des coefficients b_j correspondants. Lorsqu'une telle interprétation est possible, il lui correspond implicitement des pondérations et il faut voir si celles-ci sont ou non naturelles, si elles dépendent ou non du paramètre étudié ...

Lorsqu'une telle interprétation n'est pas évidente, il faudra analyser l'écart entre le coefficient estimé et une moyenne donnée des coefficients désagrégés b_j . Cet écart peut être vu comme un biais d'hétérogénéité.

Le paragraphe 4 contient une analyse des biais d'hétérogénéité pour diverses applications : estimations de coefficient de corrélation serielle, étude de fonctions de production de type COBB-DOUGLAS ou C.E.S.

2. *HETEROGENEITE ET EXOGENEITE*

Pour admettre l'interprétation usuelle, le modèle désagrégé (1.1) doit être tel que $E\,(u_{it}/x_{it}) = 0$. Ceci entraîne la non corrélation entre les erreurs et les variables explicatives, conduit à des estimateurs des m-c-o au niveau désagrégé qui sont sans biais et permet d'interpréter la partie déterministe du modèle : $a_i + x_{it}b_i$ comme la moyenne conditionnelle de y_{it} sachant x_{it}. Cette condition $E\,(u_{it}/x_{it}) = 0$ est l'une des définitions classiques de l'exogénéité des variables explicatives. Nous l'appellerons exogénéité simple. Une autre définition usuelle de l'exogénéité (dite exogénéité faible) consiste à ajouter à cette condition le fait que la loi marginale des explicatives ne dépend pas des paramètres d'intérêt a_i, b_i. Si cette contrainte est réalisée le modèle conditionnel donné par (1.1) apporte toute l'information sur les paramètres d'intérêt et il n'est pas nécessaire de compléter la relation par des modèles expliquant la détermination des variables x.

Il est alors clair que les cas où les versions approchées (1.2), (1.3), (1.4), (1.5), (1.6) sont mal spécifiées, c'est-à-dire où facteurs d'hétérogénéité et variables explicatives sont liés, correspondent à des modèles où les explicatives bien que simplement exogènes ne sont pas faiblement exogènes. En fait, l'exogénéité faible est satisfaite s'il y a exogénéité simple et indépendance entre explicatives et paramètres d'intérêt. Nous donnons ci-dessous deux exemples simples de tels modèles.

2.a Modèle avec variable endogène retardée

A titre d'exemple, considérons des modèles autorégressifs d'ordre 1 :

$$(2.1)\quad y_{it} = \rho_i \, y_{i,t-1} + u_{it} \qquad ; \; E\, u_{it} = 0, \; V\, u_{it} = \sigma^2,$$

où tous les coefficients ρ_i sont pris de module inférieur à 1 et où les erreurs associées à des indices individuels ou temporels différents sont non corrélés. Les variables y_{it} sont de moyennes nulles et de variance $V\, y_{it} = \frac{\sigma^2}{1-\rho_i^2}$.
On en déduit que la loi de la variable explicative $y_{i,t-1}$ dépend, au moins par l'intermédiaire du moment d'ordre deux, du coefficient ρ_i apparaissant dans la partie déterministe. La variable $y_{i,t-1}$ n'est donc pas faiblement exogène vis-à-vis du paramètre d'intérêt ρ_i.

2.b Fonctions de production dans le cas de comportement optimal de la firme

i) Importance de l'hypothèse de rendements constants

Supposons que l'on souhaite estimer une fonction de production. Pour cela on introduit un modèle donnant les valeurs de l'output y_{it} en fonction des valeurs des divers inputs $x_{1it} \ldots x_{Kit}$. Après transformation éventuelle des variables le modèle pourra se présenter sous la forme :

$$(2.2)\quad h_0(y_{it}) = a_i + \sum_{l=1}^{L} b_{li}\, h_l\,(x_{1it}, \ldots, x_{Kit}) + u_{it}.$$

Supposons alors que les firmes choisissent les quantités d'inputs nécessaires à la production de façon à minimiser leurs coûts. Si les prix des inputs sont $P_{1t}, \ldots, P_{Kt}$, les mêmes pour toutes les entreprises, les variables explicatives sont déduites des demandes de facteurs et s'écrivent :

$$x_{kit} = D_k\,(P_{1t}, \ldots, P_{Kt} \; ; \; y_{it} \; ; \; a_i, b_{1i}, \ldots b_{Li}).$$

Elles dépendent des prix, de l'endogène y_{it} et des coefficients caractérisant la forme de la fonction de production. Sous l'hypothèse de comportement optimal, on voit que les explicatives h_l $(x_{1it}, \ldots, x_{Kit})$ dépendent de l'endogène y_{it} et sont de ce fait corrélées avec l'erreur u_{it}. Elles ne sont pas simplement exogènes. Le modèle (2.2) ne peut alors être utilisé directement sans entraîner un biais d'endogénéité.

Ce problème disparaît cependant si les fonctions de production au niveau désagrégé peuvent être supposées à rendements constants. En effet le modèle désagrégé est alors écrit sous une forme :

$$(2.3)\quad h_0\left(\frac{y_{it}}{x_{1it}}\right) = a_i + \sum_{l=1}^{L} b_{li}\, h_l\left(\frac{x_{2it}}{x_{1it}}, \ldots, \frac{x_{kit}}{x_{1it}}\right) + u_{it},$$

c'est-à-dire en terme de rapport de quantités.

D'autre part l'écriture des conditions du premier ordre associées à la minimisation du coût montre que, du fait de l'homogénéité de la fonction de production, les rapports :

$$\frac{x_{kit}}{x_{1it}} = \frac{D_k\,[P_{1t},\ldots,P_{Kt};y_{it};a_i,b_{1i},\ldots,b_{li}]}{D_1\,[P_{1t},\ldots,P_{Kt};y_{it};a_i,b_{1i},\ldots,b_{Li}]} = \frac{D_k[P_{1t},\ldots P_{Kt};1;a_i,b_{1i},\ldots,b_{Li}]}{D_k[P_{1t},\ldots P_{Kt};1;a_i,b_{1i},\ldots,b_{Li}]},$$

sont indépendants du niveau y_{it} de la production.

Il y a alors non corrélation entre les variables explicatives $h_l\left(\frac{x_{2it}}{x_{1it}}, \ldots, \frac{x_{Kit}}{x_{1it}}\right)$ et le terme d'erreur. Cependant on voit que ces variables explicatives dépendent encore des coefficients a_i, b_{1i}, ..., b_{Li}. Bien qu'il y ait exogénéité simple, il n'y a pas exogénéité faible.

ii) <u>Cas d'une fonction de COBB-DOUGLAS</u>

Si la fonction de production est de type COBB-DOUGLAS à rendements constants :

$$\text{Log } y_{it} = a_i + \sum_{k=1}^{K} b_{ki} \text{ Log } x_{kit}, \qquad \text{avec } \sum_{k=1}^{K} b_{ki} = 1,$$

les conditions du premier ordre correspondant au problème d'optimisation donnent :

$$\frac{b_{ki}}{b_{1i}} = \frac{x_{1it}}{x_{kit}} = \frac{P_{kt}}{P_{1t}} \text{ ,}$$

$$\Leftrightarrow \frac{x_{1it}}{x_{kit}} = \frac{P_{kt}}{P_{1t}} \frac{b_{1i}}{b_{ki}} \text{ .}$$

Les variables explicatives introduits dans la "formulation linéaire" de la fonction de COBB-DOUGLAS sont : Log $\frac{x_{kit}}{x_{1it}}$. Elles s'écrivent :
Log $\frac{x_{kit}}{x_{1it}}$ = Log $\frac{P_{kt}}{P_{1t}}$ + Log $\frac{b_{1i}}{b_{ki}}$. Elles apparaissent comme somme d'un effet temporel dû aux prix et d'un effet individuel dû à l'hétérogénéité des fonctions de production. Bien que les variables Log $\frac{x_{kit}}{x_{1it}}$ ne soient pas faiblement exogènes, on notera que les variables différenciées Log $\frac{x_{kit}}{x_{1it}}$ - Log $\frac{x_{kit-1}}{x_{1it-1}}$ le sont.

iii) <u>Cas d'une fonction CES</u>

Cette fonction se présente sous une forme :

$$y_{it}^r = \sum_{k=1}^{K} b_{ki}\, x_{kit}^r$$

$$\Leftrightarrow \left(\frac{y_{it}}{x_{1it}}\right)^r = b_{1i} + \sum_{k=2}^{K} b_{ki} \left(\frac{x_{kit}}{x_{1it}}\right)^r .$$

Les conditions d'optimalité du premier ordre donnent :

$$\frac{b_{ki}\, r\, x_{kit}^{r-1}}{b_{1i} r\, x_{1it}^{r-1}} = \frac{P_{kt}}{P_{1t}}$$

$$\left(\frac{x_{kit}}{x_{1it}}\right)^r = \left(\frac{P_{kt}}{P_{1t}}\right)^{\frac{r}{r-1}} \left(\frac{b_{1i}}{b_{ki}}\right)^{\frac{r}{r-1}} .$$

Les variables explicatives du modèle apparaissent sous forme de produits d'un effet temporel dû aux prix et d'un effet individuel dû à l'hétérogénéité.

3. ETUDE DES ESTIMATEURS DES M.C.O.

A. Forme des estimateurs

De façon à mettre en évidence les éventuels biais d'hétérogénéité, nous commençons par donner les formes des estimateurs des m.c.o. des paramètres b, β, b^{**}, b^{***} correspondant aux divers modèles approchés.

a) Modèle de l'individu représentatif

La régression est menée sur toutes les observations aussi bien individuelles que temporelles. On a :

$$(3.1)\quad \hat{b} = \left[V_e\, X\right]^{-1} \mathrm{Cov}_e\,(X,\, Y),$$

où X, Y désignent les vecteur et matrice d'observations relatives à tous les individus et à toutes les dates.

b) Modèle agrégé

L'estimateur est donné par :

$$(3.2)\quad \hat{\beta} = \left[V_e\,(\sum_i X_i)\right]^{-1} \mathrm{Cov}_e(\sum_i X_i,\, \sum_i Y_i),$$

où Y_i, X_i, sont les vecteur et matrice d'observations correspondant au i^e individu.

c) Modèle à effets fixes en différence

On a :

$$(3.3)\quad \hat{b}^{**} = \left[\sum_i \; E_e\, (X_i-\bar{X}_i)' \; (X_i-\bar{X}_i)\right]^{-1} \sum_i \; E_e \left[\; (X_i-\bar{X}_i)' \; (Y_i-\bar{Y}_i)\right],$$

où le symbole $^{-}$ signifie que les données sont décalées d'une période dans le temps.

d) <u>Modèle à effets fixes intra</u>

L'estimateur est :

$$(3.4)\quad \hat{b}^{***} = \left[\sum_i V_e X_i \right] \sum_i \mathrm{Cov}_e \, (X_i, Y_i).$$

B. <u>Limites des estimateurs</u>

Sous des conditions de régularité usuelles sur les erreurs u_{it} et les variables explicatives x_{it}, les divers estimateurs tendent vers des limites, lorsque le nombre T de dates tend vers l'infini. Ces limites s'obtiennent en remplaçant dans les diverses expressions Y_i par $a_i + X_i b_i$, c'est-à-dire en mettant à zéro l'erreur, et en prenant les limites en probabilité des variances et covariances empiriques. On note :

$$V X_i = \operatorname*{plim}_T \underset{e}{V} X_i, \quad \mathrm{Cov}\,(X_i, Y_i) = \operatorname*{plim}_T \underset{e}{\mathrm{Cov}}\,(X_i, Y_i).$$

Les limites des divers estimateurs sont données ci-dessous.

a) <u>Modèle de l'individu représentatif</u>

Pour obtenir la limite, il nous faut commencer par utiliser les formules d'analyse de la variance.

$$\underset{e}{V} X = \frac{1}{N} \sum_i \underset{e}{V} X_i + \frac{1}{n} \sum_i \left[\underset{e}{E} X_i - \frac{1}{n} \sum_i \underset{e}{E} X_i \right]' \left[\underset{e}{E} X_i - \frac{1}{n} \sum_i \underset{e}{E} X_i \right],$$

$$\underset{e}{\mathrm{Cov}}\,(X,Y) = \frac{1}{n} \sum_i \underset{e}{\mathrm{Cov}}\,(X_i, Y_i) + \frac{1}{n} \sum_i \left[\underset{e}{E} X_i - \frac{1}{n} \sum_i \underset{e}{E} X_i \right]' \left[\underset{e}{E} Y_i - \frac{1}{n} \sum_i \underset{e}{E} Y_i \right].$$

Comme les limites sont obtenues en remplaçant Y_i par $a_i + X_i b_i$, on en déduit

$$\underset{T}{\text{plim}}\ \hat{b} = \underset{T}{\text{plim}} \left[\frac{1}{n} \sum_i \underset{e}{V} X_i + \frac{1}{n} \sum_i \left(\underset{e}{E} X_i - \frac{1}{n} \sum_i \underset{e}{E} X_i \right)' \left(\underset{e}{E} X_i - \frac{1}{n} \sum_i \underset{e}{E} X_i \right) \right]^{-1}$$

$$\left[\frac{1}{n} \sum_i \underset{e}{\text{Cov}} \left(X_i , a_i + X_i b_i \right) + \frac{1}{n} \sum_i \left(\underset{e}{E} X_i - \frac{1}{n} \sum_i \underset{e}{E} X_i \right)' \left[a_i + \underset{e}{E} X_i b_i - \frac{1}{n} \sum_i a_i - \frac{1}{n} \sum_i \underset{e}{E} X_i b_i \right] \right],$$

$$\underset{T}{\text{plim}}\ \hat{b} = \left[\sum_i \left(V X_i + \left(E X_i - \frac{1}{n} \sum_i E X_i \right)' \left(E X_i - \frac{1}{n} \sum_i E X_i \right) \right) \right]^{-1}$$

$$\left[\sum_i \left(V X_i + \left(E X_i - \frac{1}{n} \sum_i E X_i \right)' \left(E X_i - \frac{1}{n} \sum_i E X_i \right) \right) b_i \right.$$

$$\left. + \sum_i \left(E X_i - \frac{1}{n} \sum_i E X_i \right) \left(a_i - \frac{1}{n} \sum_i a_i \right) \right].$$

b) <u>Modèle agrégé</u>

La limite s'obtient immdiatement.

$$\underset{T}{\text{plim}}\ \hat{\beta} = \underset{T}{\text{plim}} \left[\underset{e}{V} \left(\sum_i X_i \right) \right]^{-1} \underset{e}{\text{Cov}} \left(\sum_i X_i , \sum_i Y_i \right).$$

$$\text{(3.6)} \quad \underset{T}{\text{plim}}\ \hat{\beta} = \left[\underset{e}{V} \left(\sum_i X_i \right) \right]^{-1} \sum_j \text{Cov} \left(\sum_i X_i , X_j \right) b_j$$

c) Modèle à effets fixes en différence

Comme asymptotiquement la moyenne des valeurs x_{it} sur le temps coïncide avec la moyenne de valeurs retardées, on peut directement travailler sur données centrées. On a :

$$\underset{T}{\text{plim}}\ \hat{b}^{**} = \underset{T}{\text{plim}} \left(\sum_i \underset{e}{V} \left[X_i - X_i^-\right] \right)^{-1} \sum_i \underset{e}{\text{Cov}} \left(X_i - X_i^-\ ,\ Y_i - Y_i^-\right)$$

$$= \left(\sum_i V\ \left[X_i - X_i^-\right] \right)^{-1} \sum_i \text{Cov} \left(X_i - X_i^-\ ,\ a_i + X_i b_i - a_i - X_i^-\ b_i\right),$$

$$(3.7)\quad \underset{T}{\text{plim}}\ \hat{b}^{**} = \left(\sum_i V \left[X_i - X_i^-\right] \right)^{-1} \sum_i V\ (X_i - X_i^-)\ b_i .$$

d) Modèle à effets fixes intra

On a immédiatement :

$$(3.8)\quad \underset{T}{\text{plim}}\ \hat{b}^{***} = \left[\Sigma\ V\ X_i\right]^{-1} \sum_i \ VX_i\ b_i .$$

C. Interprétation des limites en terme de moyennes

On peut noter que les limites :

$$\beta_\infty = \underset{T}{\text{plim}}\ \hat{\beta}\quad ,\ b_\infty^{**} = \underset{T}{\text{plim}}\ \hat{b}^{**}\quad ,\ b_\infty^{***} = \underset{T}{\text{plim}}\ \hat{b}_\infty^{***}$$

, sont de la forme :

$$(3.9)\quad \beta_\infty = \sum_{j=1}^{n} C_j b_j,\qquad \beta_\infty^{**} = \sum_{j=1}^{n} \Gamma_j^{**} b_j,\qquad \beta_\infty^{***} = \sum_{j=1}^{n} \Gamma_j^{***}\ b_j,$$

avec $C_j = \left(V\left[\sum_i X_i\right]\right)^{-1} \; Cov\left(\sum_i X_i, X_j\right)$,

$$\Gamma_j^{**} = \left(\sum_i V\left[X_i - X_i^-\right]\right)^{-1} V \; \left(X_j - X_j^-\right),$$

$$\Gamma_j^{***} = \left(\sum_i V X_i\right)^{-1} V X_j.$$

Ces trois systèmes de matrices sont constitués d'éléments tels que :

$$\sum_{j=1}^{n} C_j = \sum_{j=1}^{n} \Gamma_j^{**} = \sum_{j=1}^{n} \Gamma_j^{***} = Id.$$

On voit donc que les diverses démarches approchées conduisent à interpréter $\hat{\beta}$, $\hat{b}^{**}$, $\hat{b}^{***}$ comme des estimateurs de "moyennes" des coefficients désagrégés b_j. Il faut bien remarquer cependant que les coefficients C_j, Γ_j^{**}, Γ_j^{***} ne possèdent a priori aucune propriété de positivité.

De plus si on souhaite expliciter les résultats composante par composante, on obtient des relations du type :

$$\hat{\beta}^k = \sum_{l=1}^{K} \left[\sum_{j=1}^{n} C_j(k,l)\; b_j^l\right],$$

où : $C_j(k,l)$ désigne l'élément d'indice (k,l) de la matrice C_j et $\hat{\beta}^k$, b_j^l les k^e et l^e composantes des vecteurs $\hat{\beta}$ et b_j.

Comme : $\sum_{j=1}^{n} C_j = Id$, on a : $\sum_{j=1}^{n} C_j(k,k) = 1$, $\sum_{j=1}^{n} C_j\,(k,l) = 0$, $k \neq l$.

On voit ainsi sous forme développée que le terme $\hat{\beta}^k$ n'est pas simplement une "moyenne" des paramètres individuels $\hat{b}_j^k$ correspondants. Il dépend aussi des paramètres individuels associés aux autres variables explicatives (avec un effet moyen nul). Ainsi dans un modèle de consommation écrit en logarithme où les variables explicatives sont le revenu et la taille du ménage, l'élasticité revenu calculée directement à partir de l'un des modèles approchés dépend aussi des élasticités tailles désagrégées. La présence d'un tel effet pervers a été initialement mise en évidence par THEIL [1954] pour le cas du modèle agrégé. L'effet est encore plus important, lorsqu'on considère le modèle à individu représentatif. En effet la limite de l'estimateur dépend aussi du terme constant a_i du modèle de régression.

D. Comparaison avec la moyenne des coefficients désagrégés

En pratique il peut être important de se donner a priori le paramètre d'intérêt déduit des coefficients microéconomiques. Le plus naturel est alors de définir ce paramètre d'intérêt par :

$$(3.10)\quad B^k = \sum_{j=1}^{n} \pi_j \, b_j^k,$$

où les π_j sont des pondérations données. On peut ainsi s'intéresser à la moyenne arithmétique des coefficients $\pi_j = \frac{1}{n}$ ou à une moyenne pondérée par une variable de taille. Nous décrivons le calcul du biais pour le modèle agrégé.

Notons $C_j(k,.)$, la k^e ligne de la matrice C_j. On voit alors que :

$$\beta_\infty^k - B^k = \sum_{j=1}^{n} C_j\,(k,.)\; b_j - \sum_{j=1}^{n} \pi_k\, b_j^k$$

$$= \sum_{j=1}^{n} \left(C_j\,(k,.) - [0,\; \ldots\; 0,\; \pi_k,\; 0,\; \ldots\; 0]\right) b_j.$$

Comme : $\displaystyle\sum_{j=1}^{n} \left(C_j\,(k,.) - [0,\; \ldots\; 0,\; \pi_k,\; 0,\; \ldots\; 0]\right)$

$= [0, \ldots 0, 1, 0, \ldots, 0] - [0, \ldots, 0, 1, 0 \ldots, 0] = 0$, on a donc:

$$\beta_{\infty}^{k} - B^{k} = \underset{e}{\mathrm{Cov}} \left(C_j(k,.) - [0, \ldots 0, \pi_k, 0, \ldots 0], b_j\right),$$

$$(3.11) \quad \beta_{\infty}^{k} - B^{k} = \underset{\pi,e}{\mathrm{Covi}} \left(C_j(k,.), b_j\right),$$

la covariance empirique étant ici calculée sur les individus, d'où la notation Covi. Remarquons d'ailleurs qu'une formule analogue peut être obtenue en attribuant à chaque individu le poids π_j. Notant $\mathrm{Covi}_{\pi,e}$ la covariance calculée avec ces poids, on a :

$$(3.12) \quad \beta_{\infty}^{k} - B^{k} = \underset{\pi.e}{\mathrm{Covi}} \left(\frac{C_j(k,.)}{\pi_k}, b_j\right).$$

Les formules (3.11), (3.12) fournissent évidemment le biais d'estimation, lorsqu'on s'intéresse au paramètre B^k et que celui-ci a été estimé par $\hat{\beta}_k$. Ce biais est dit <u>biais d'hétérogénéité</u>.

Lorsque les coefficients individuels b_j peuvent être considérés comme indépendants des explicatives X_j, on a aussi l'indépendance entre les coefficients et les "poids" $C_j(k,.)$, puisque ces derniers ne dépendent que des variables explicatives. Dans ce cas et si le nombre n d'individus est grand, on voit que $b_{\infty}^{k} - B^{k} \neq 0$. On retrouve ainsi la propriété de convergence de l'estimateur sur modèle agrégé, lorsque l'hétérogénéité est indépendante des explicatives déjà introduites dans le modèle.

4. CAS PARTICULIERS

A. <u>Modèles avec une seule variable explicative : forme des estimateurs</u>

Les résultats des paragraphes précédents prennent une forme plus simple, lorsque le modèle initial comporte une seule variable explicative véritable. Les diverses limites sont :

Modèle de l'individu représentatif

$$\underset{T}{\text{plim}}\ \hat{b} = \frac{\sum_j \sum_t (x_{jt}-x_{..})^2\, b_j}{\sum_i \sum_t (x_{it}-x_{..})^2} + \frac{\sum_j (x_{j.}-x_{..})\,(a_j - a_.)}{\sum_i \sum_t (x_{it}-x_{..})^2} \quad \text{(voir (3.5))}$$

Modèle agrégé

$$\underset{T}{\text{plim}}\ \hat{\beta} = \frac{\sum_j \text{Cov}\,(\sum_i X_i,\ X_j)\, b_j}{V\,(\sum_i X_i)} \quad \text{(voir (3.6))}$$

Modèle à effets fixes en différence

$$\underset{T}{\text{plim}}\ \hat{b}^{**} = \frac{\sum_j VX_j\ (1-\varrho_j)\ b_j}{\sum_i VX_i\ (1-\varrho_i)}, \quad \text{(voir (3.7))}$$

où ϱ_j désigne la corrélation d'ordre 1 du processus (x_{jt}).

Modèle à effets fixes intra :

$$\underset{T}{\text{plim}}\ \hat{b}^{***} = \frac{\sum_j VX_j\ b_j}{\sum_i VX_i} \quad \text{(voir (3.8)}$$

On remarque que les effets pervers d'hétérogénéité des autres coefficients ont disparu, sauf pour l'estimateur fondé sur le modèle de l'individu représentatif. D'autre part les deux estimateurs fondés sur le modèle à effets fixes convergent vers de véritables moyennes des coefficients désagrégés, les poids étant bien dans ce cas positifs. En revanche la positivité n'est en général pas satisfaite pour l'estimateur construit à l'aide du modèle agrégé. Seuls les estimateurs $\hat{b}^{**}$ et $\hat{b}^{***}$ conservent une interprétation raisonnable en fonction des paramètres désagrégés.

B. Modèles avec une seule variable explicative : étude des biais

Il est clair d'après les propriétés précédentes que les divers estimateurs convergent vers des limites en général différentes. Les biais peuvent facilement être étudiés en utilisant les résultats du paragraphe 3.D. On suppose que le paramètre d'intérêt est la moyenne empirique des coefficients désagrégés : $B = \frac{1}{n} \sum_{j=1}^{n} b_j$, et, à titre d'exemple, on considère les deux estimateurs construits à partir du modèle à effets fixes.

On a :

$$\operatorname{plim}_T \hat{b}^{**} - B = \operatorname{Cov}_e^i \left(\frac{VX_j \, (1-\rho_j)}{\sum_i VX_i (1-\rho_i)}, b_j \right)$$

$$= \frac{1}{\sum_i VX_i \, (1-\rho_i)} \operatorname{Cov}_e^i \left[VX_j \, (1-\rho_j), \, b_j \right],$$

$$\operatorname{plim}_T \hat{b}^{***} - B = \frac{1}{\sum_i VX_i} \operatorname{Cov}_e^i \left[VX_j, \, b_j \right],$$

$$\operatorname{plim}_T \hat{b}^{***} - \operatorname{plim}_T \hat{b}^{**} = \operatorname{Cov}_e^i \left(\frac{VX_j}{\sum_i VX_i} - \frac{VX_j \, (1-\rho_j)}{\sum_i VX_i \, (1-\rho_i)}, b_j \right).$$

Introduisant les covariances empiriques pondérées par $\frac{VX_j}{\sum_i VX_i}$, notées $\operatorname{Cov} i_{V,e}$ on a :

$$\operatorname{plim}_T \hat{b}^{***} - \operatorname{plim}_T \hat{b}^{**} = \operatorname{Cov} i_{V,e} \left[1 - (1-\rho_j) \frac{\sum_i VX_i}{\sum_i' VX_i (1-\rho_i)} , \, b_j \right]$$

$$= \frac{\sum_i VX_i}{\sum_i VX_i\ (1-\varrho_i)} \operatorname{Cov}^{i}_{V,e} (\varrho_j, b_j).$$

Ainsi l'estimateur intra $\hat{b}^{***}$ tend à surestimer la moyenne B des coefficients désagrégés, lorsque les variabilités temporelles des explicatives sont corrélées positivement avec les coefficients b_j. L'estimateur par différence $\hat{b}^{**}$ fournit une surestimation de B, si les variabilités temporelles des accroissements de l'explicative sont corrélées positivement avec les coefficients b_j. Finalement l'estimateur intra tend à donner des estimations plus élevées que l'estimateur par différence si les corrélations sérielles des explicatives sont liées positivement aux coefficients b_j.

Ceci permet de mettre en évidence quelques conditions de convergence.

<u>Propriété 4.1</u>

> Si la variable explicative se décompose additivement en un effet temporel et un effet individuel : $X_{it} = Z_t + W_i$, alors :
>
> $\operatorname{plim}_T \hat{b}^{**} = \operatorname{plim}_T \hat{b}^{***} = B.$

<u>Preuve</u>

i) On a : $VX_j = V\ (Z_t+W_j) = VZ_t$, puisque la composante W_j est constante en t. On voit alors que la variance VX_j est constante en j et on en déduit que :

$$\operatorname{plim}_T \hat{b}^{***} - B = \frac{1}{\sum_i VX_i} \operatorname{Cov}^{i}_{e} (V\, X_j, b_j) = 0.$$

ii) La même propriété s'applique à la variable différenciée, dont la décomposition est :

$$X_{it} - X_{it-1} = Z_t - Z_{t-1}.$$

L'utilisation du modèle à effets fixes permet ici d'obtenir la convergence.

Les deux autres estimateurs tendent respectivement vers les limites ci-dessous.

$$\operatorname*{plim}_{T} \hat{b} = \left[\sum_i \left[VZ_t + (W_i - \frac{1}{n} \sum_i W_i)^2\right]\right]^{-1}$$

$$\left[\sum_i \left[VZ_t + (W_i - \frac{1}{n} \sum_i W_i)^2\right] b_i + \sum_i (W_i - \frac{1}{n} \sum_i W_i)\, a_i\right],$$

$$\operatorname*{plim}_{T} \hat{b} - B = \left[VZ_t + V_i \left[W_j\right]\right]^{-1} \left[\mathrm{Covi}\left[(W_j - \frac{1}{n} \sum W_i)^2, b_j\right] + \mathrm{Covi}\left[W_j, a_j\right]\right]^1,$$

où Vi désigne la variance calculée sur les individus. On voit alors que si la composante temporelle est importante, c'est-à-dire si VZ_t est grand, le biais est petit.

L'estimateur calculé sur le modèle agrégé est tel que :

$$\operatorname*{plim}_{T} \hat{\beta} = \frac{\sum_j \mathrm{Cov}\,(\sum_i (Z_t + W_i),\ Z_t + W_j)\, b_j}{V\left[\sum_i (Z_t + W_i)\right]}$$

$$= \frac{\sum_j n\, V\, Z_t\, b_j}{V\,(nZ_t)} = \frac{1}{n} \sum_j b_j = B.$$

Propriété 4.2

> Lorsque la variable explicative se décompose additivement en un effet temporel et un effet individuel, l'estimateur calculé sur le modèle agrégé est convergent de la moyenne B des coefficients désagrégés.

C. Application aux modèles autorégressifs hétérogènes

Reprenons l'exemple d'un modèle autorégressif hétérogène :

$$y_{it} = \rho_i\, y_{it-1} + u_{it},\ |\rho_i| < 1,$$

où les erreurs sont centrées, indépendantes, de même variance σ^2. La variable y_{it} est centrée et, lorsque les indices individuels diffèrent, il y a indépendance entre y_{it} et $y_{j\tau}$, $i \neq j$.

Les divers modèles approchés sont :

$$y_{it} = \varrho\, y_{it-1} + u_{it},$$

$$\sum_i y_{it} = \varrho \sum_i y_{it-1} + \omega_t,$$

$$y_{it} - y_{i.} = \varrho\,(y_{it-1} - y_{i.}) + v^{*}_{it},$$

$$y_{it} - y_{it-1} = \varrho\,(y_{i,t-1} - y_{i,t-2}) + v^{**}_{it}.$$

Utilisant les propriétés des processus $(y_{it},\ t \in Z)$, on voit immédiatement que les estimations des m.c.o. de ϱ calculés sur ces divers modèles convergent tous vers :

$$\frac{1}{n}\sum_{j=1}^{n} \varrho_j + \frac{1}{\sum_j V(y_{jt})} \operatorname{Cov}_e^j \left[V\, y_{jt},\ \varrho_j \right]$$

$$= \frac{1}{n}\sum_{j=1}^{n} \varrho_j + \frac{1}{\sum_j \frac{\sigma^2}{1-\varrho_j^2}} \operatorname{Cov}_e^j \left(\frac{\sigma^2}{1-\varrho_j^2},\ \varrho_j\right).$$

Remarquant que $\frac{1}{1-\varrho_j^2}$ est une fonction croissante (resp. décroissante) de ϱ_j, si ϱ_j est positif (resp. négatif), on obtient le résultat suivant.

<u>Propriété 4.3.</u>

i) Les estimations des m.c.o. calculées sur les divers modèles approchés ont tous même biais.

ii) Si toutes les corrélations ϱ_j sont positives (resp. négatives), il y a surestimation (resp. sous-estimation) de la corrélation moyenne.

Ce résultat est à rapprocher de ceux obtenus par GRANGER (1980) et GONÇALVES-GOURIEROUX (1987) montrant que l'agrégation de processus autorégressif d'ordre 1 conduit à un accroissement des délais temporels.

Une étude plus précise du biais peut être faite dans le cas de deux sous-populations. Le biais est donné par :

$$\underset{T}{\text{plim}}\ \hat{\varrho} - \frac{1}{2}(\varrho_1+\varrho_2) = \frac{1}{4}\ \frac{(\varrho_1-\varrho_2)\ \dfrac{(\varrho_1+\varrho_2)}{2}}{1-\dfrac{\varrho_1^2+\varrho_2^2}{2}} = \frac{\frac{1}{4}(\varrho_1-\varrho_2)^2\ \dfrac{\varrho_1+\varrho_2}{2}}{1-\dfrac{(\varrho_1+\varrho_2)^2}{4}-\dfrac{(\varrho_1-\varrho_2)^2}{4}}.$$

On voit qu'il ne dépend pas de la variance σ^2 du bruit. Il est fonction croissante de l'hétérogénéité $(\varrho_1-\varrho_2)^2$ [à $\frac{\varrho_1+\varrho_2}{2}$ donné]. D'autre part il est fonction croissante de $\frac{\varrho_1+\varrho_2}{2}$ pour ϱ_1, ϱ_2 positifs et $|\varrho_1-\varrho_2|$ donné. Ainsi plus on se rapproche de la non stationnarité, c'est-à-dire de la racine 1 correspondant à la marche aléatoire, plus le biais risque d'être important. Finalement on notera que le biais peut ête nul dans des cas où l'hétérogénéité et la variable explicative sont dépendantes. Dans notre exemple ceci se produit pour $\varrho_2 = -\varrho_1$.

D. Application aux fonctions de production de type Cobb-Douglas

i) Cas de comportement optimal de la forme

Le modèle désagrégé s'écrit :

$$\text{Log}\ \frac{y_{it}}{x_{1it}} = a_i + \sum_{k=2}^{K} b_{ki}\ \text{Log}\ \frac{x_{kit}}{x_{1it}} + u_{it},$$

où y désigne l'output et x_k les divers inputs. Nous avons vu dans le paragraphe 2 que la minimisation du coût conduit à des demandes de facteurs telles que :

$$\text{Log } \frac{x_{kit}}{x_{1it}} = \text{Log } \frac{P_{kt}}{P_{1t}} + \text{Log } \left[\frac{1 - \sum_{k=2}^{K} b_{ki}}{b_{ki}} \right].$$

Il y a ainsi une décomposition des explicatives en somme d'un effet temporel et d'un effet individuel. On peut alors appliquer les propriétés 4.1. et 4.2.

Propriété 4.4.

> Dans le cas de fonctions de production de COBB-DOUGLAS, homogènes de degré 1, et d'un comportement de minimisation du coût, les estimateurs des m.c.o. fondés sur le modèle agrégé ou sur le modèle à effets fixes sont convergents.

L'estimateur fondé sur le modèle à individu représentatif est lui non convergent. Cependant le biais asymptotique sera faible si la variabilité des prix est importante.

ii) Cas général

On peut se demander si les divers modèles approchés estimés sur données réelles conduisent bien à des estimations proches correspondant à cette idée de comportement optimal de la firme. On peut à titre d'illustration reprendre les résultats obtenus par GRILICHES-MAIRESSE (1987). Les fonctions de production désagrégées sont : $\text{Log } \frac{y_{it}}{l_{it}} = a_i + b_i \text{ Log } \frac{k_{it}}{l_{it}} + u_{it}$, où k, l désignent respectivement le capital et le travail. Nous reportons ci-dessous les valeurs estimées des diverses limites pour la France, le Japon et les Etats-Unis.

	FRANCE	JAPON	U.S.A.
$B = \frac{1}{n} \sum_i b_i$	0.194	0.273	0.219
b_∞ représentatif	0.303	0.452	0.221
b_∞^{**} effet fixe différence	0.260	0.183	0.289
b_∞^{***} intra	0.196	0.278	0.213

(Les variables étant en logarithmes le modèle agrégé n'a évidemment pas de sens ici).

On s'attend à ce que l'estimateur par individu représentatif, qui peut contenir un effet pervers dû à l'hétérogénéité des niveaux de production (c'est-à-dire des a_i), puisse être sensiblement différent des autres estimateurs. Cet effet est immédiatement visible pour la France et le Japon.

Par ailleurs, on s'aperçoit que les autres estimateurs prennent des valeurs assez différentes ce qui est non compatible avec l'idée d'optimalité des firmes. Les signes des biais apportent alors de l'information sur les corrélations entre hétérogénéité et variables explicatives.

Si l'élasticité du capital b_i est grande, on peut penser que la substitution travail-capital sera très rapide, entraînant une forte corrélation sérielle pour la variable Log $\frac{k_{it}}{l_{it}}$. Ainsi il peut paraître naturel d'avoir Covi $(\varrho_j, b_j) > 0$, ce qui implique que b_∞^{***} est plus grand que b_∞^{**}. On voit qu'un tel raisonnement qui, pour être valable, demande une possibilité d'adaptation rapide des firmes, ne pourrait ici s'appliquer qu'au Japon.

iii) Rendements

Le comportement étant non optimal, les variables k_{it}, l_{it} présentent peut-être peu de caractères endogènes. Ceci nous conduirait alors à considérer un modèle approché où l'hypothèse de rendement constant n'est pas forcément introduite. De façon plus précise, supposons qu'au niveau désagrégé le modèle soit :

$$\text{Log } y_{it} = a_i + b_i \text{ Log } k_{it} + c_i \text{ Log } l_{it} + u_{it},$$

avec $b_i + c_i = 1, \forall i$.

On peut a priori considérer plusieurs modèles à effets fixes selon que l'on introduit ou non la contrainte de rendement constant. Deux modèles sont ainsi :

$$\text{Log } y_{it} = a_i^* + b^* \text{ Log } k_{it} + (1-b)^* \text{ Log } l_{it} + v_{it}^*,$$

où la contrainte est prise en compte, et :

$$\text{Log } y_{it} = \tilde{\tilde{a}}_i^* + \tilde{\tilde{b}}^* \text{ Log } k_{it} + \tilde{\tilde{c}}^* \text{ Log } l_{it} + w_{it}^* \quad ,$$

où les paramètres $\tilde{\tilde{b}}^*$ et $\tilde{\tilde{c}}^*$ ne sont pas soumis à contrainte. Que se passera t-il si nous estimons ce dernier modèle par exemple par la méthode intra ? Retrouvera-t-on ou non la contrainte : $b_\infty^{***} + c_\infty^{***} = 1$ au niveau des limites, c'est-à-dire l'estimateur naturel du rendement conduira-t-il à retrouver celui valable au niveau microéconomique pour toutes les entreprises ?

D'après les résultats du paragraphe précédent les limites sont du type :

$$\begin{pmatrix} b_\infty^{***} \\ c_\infty^{***} \end{pmatrix} = \sum_{j=1}^{n} \begin{bmatrix} \gamma_j(1,1) & \gamma_j(1,2) \\ \gamma_j(2,1) & \gamma_j(2,2) \end{bmatrix} \begin{bmatrix} b_j \\ c_j \end{bmatrix},$$

avec : $\sum_j \gamma_j(1,1) = \sum_j \gamma_j(2,2) = 1,$

$\sum_j \gamma_j(2,1) = \sum_j \gamma_j(1,2) = 0,$

$b_j + c_j = 1, \forall j.$

On en déduit que :

$$b_{\infty}^{***} + c_{\infty}^{***} = \sum_j \left[\gamma_j(1,1) + \gamma_j(2,1)\right] b_j + \sum_j \left[\gamma_j(1,2) + \gamma_j(2,2)\right] c_j$$

$$= 1 + \sum_j \left[\gamma_j(1,1) + \gamma_j(2,1) - \gamma_j(1,2) - \gamma_j(2,2)\right] b_j .$$

Ainsi lorsque la contrainte de rendement constant n'est pas imposée a priori sur le modèle à effets fixes, l'estimateur du rendement ne tend pas en général vers 1. Il y a un biais d'hétérogénéité sur le rendement. Le sens du biais peut être aussi bien positif que négatif. On connait cependant divers cas, où il y a tendance à sous-estimation du rendement [voir HOUTTAKER, (1955)].

E. Application aux fonctions de production C.E.S.

L'étude des biais dans le cas de comportement optimal de la firme peut être menée pour d'autres fonctions de production. Considérons ainsi une fonction CES dans laquelle le paramètre puissance r est supposé connu, indépendant de l'individu :

$$\left(\frac{y_{it}}{l_{it}}\right)^r = a_i + b_i \left(\frac{k_{it}}{l_{it}}\right)^r + u_{it}.$$

Soit ω_t le prix relatif du travail par rapport au capital. Les demandes de facteur sont telles que :

$$\left(\frac{k_{it}}{l_{it}}\right)^r = \left(\omega_t\right)^{\frac{-r}{r-1}} \left(\frac{a_i}{b_i}\right)^{\frac{r}{r-1}} .$$

On peut alors déterminer par exemple le biais de l'estimateur intra. Il est donné par :

$$\underset{T}{\text{plim}}\ \hat{b}^{***} - \beta = \frac{1}{\sum_i V\left[\left(\frac{k_i}{l_i}\right)^r\right]} \text{Cov}_e^i \left(V\left[\left(\frac{k_j}{l_j}\right)^r\right], b_j \right)$$

$$= \frac{1}{V\left(\omega^{-\frac{r}{r-1}}\right) \sum_j \left(\frac{a_j}{b_j}\right)^{\frac{2r}{r-1}}} \text{Cov}_e^j \left[V\left(\omega^{-\frac{r}{r-1}}\right) \left(\frac{a_j}{b_j}\right)^{\frac{2r}{r-1}}, b_j \right]$$

$$= \frac{\text{Cov}_e^j \left[\left(\frac{a_j}{b_j}\right)^{\frac{2r}{r-1}}, b_j \right]}{\sum_j \left(\frac{a_j}{b_j}\right)^{\frac{2r}{r-1}}} .$$

Ce biais est indépendant de l'évolution des prix et uniquement fonction de la distribution jointe des coefficients.

REFERENCES

FISHER F. (1969)
"The existence of aggregate production functions"
Econometrica 37, 553-577.

FORTIN A. (1987)
"In search of aggregation biases"
DP Univ. of British Columbia.

GONÇALVES E., GOURIEROUX C. (1987)
"Agregation de processus autorégressifs d'ordre 1"
CEPREMAP-DP, à paraître dans Annales d'Economie et de Statistiques.

GRANGER C. (1980)
"Long memory relationships and the aggregation of dynamic models"
Journal of Econometrics 14, 227-238.

GREEN H. (1964)
"Aggregation in Economic Analysis : An introductory Survey"
Princeton Univ. Press.

GRILICHES Z., MAIRESSE J. (1987)
"Heterogeneity in panel data : are there stable production functions ?"
D.P.

GRUNFELD Y., GRILICHES Z. (1960)
"Is aggregation necessarily bad ?"
Review of Economics and Statistics, ,1-13.

HOUTHAKKER H. (1955)
"The Pareto distribution and the Cobb-Douglas production-function in activity analysis".
Review of Economic Studies, 23, 27-31.

KLEIN L. (1946)
"Remarks on the theory of aggregation".
Econometrica 14, 303-312.

MAIRESSE J. (1987)
"Les lois de la production ne sont plus ce qu'elles étaient : une introduction à l'économétrie des panels"
La Revue Economique, 39, 225-272.

NATAF A. (1948)
"Sur la possibilité de construction de certains macro-modèles",
Econometrica 16, 232-244.

THEIL H. (1954)
"Linear Aggregation of Economic Relations",
North Holland.

VAN DAAL J., MERKIES A. (1984)
"Aggregation in Economic Research : from individual to macro relations". Dordrecht, Reidel.

ZELLNER A. (1969)
"On the Aggregation problem : A new approach to a troublesome problem" dans Economic Models, Estimation and risk programming. Essays in Honor of Gerhard Tintner, ed. par Fox et al, Springer-Verlag.