

UNIVERSITE DE MONTREAL

EXAMEN ET ESTIMATION D'UN MODELE
DYNAMIQUE DE DESEQUILIBRE
(ETUDE ECONOMIQUE)

PAR

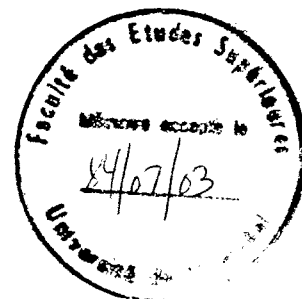
ALAIN FUGERE

DEPARTEMENT DE SCIENCES ECONOMIQUES

FACULTE DES ARTS ET DES SCIENCES

MEMOIRE DE MAITRISE PRESENTE A LA FACULTE DES ETUDES
SUPERIEURES EN VUE DE L'OBTENTION DU GRADE DE
MAITRE ES SCIENCES (M.Sc.)

MARS 1984



SOMMAIRE

Nous examinons, dans ce mémoire, une version dynamique du modèle de déséquilibre entre l'offre et la demande sur le marché d'un bien, le modèle de Dagenais (1980). Ce modèle permet de tenir compte des reports possibles du déséquilibre d'une période sur la suivante, soit par l'intermédiaire des stocks des producteurs, soit par la persistance (au moins partielle) de la demande excédentaire, quand elle survient.

La non-linéarité de la forme réduite de ce type de modèle en rend souvent l'estimation difficile. Notre principale contribution dans ce mémoire consiste ainsi dans la présentation et l'expérimentation d'un algorithme permettant d'estimer de façon itérative une équation non linéaire à l'aide de la méthode des moindres carrés linéaires, sous la condition cependant que l'équation soit de forme multi-linéaire.

Cette méthode est expérimentée en utilisant les données du marché du bois de sciage (Southern Pine) aux Etats-Unis pour les années 1956 à 1969.

De sérieux problèmes de multicollinéarité rendent cependant très instables les paramètres estimés. Notre mémoire se termine par un examen théorique de diverses voies permettant d'espérer atténuer ce type d'inconvénients.

TABLE DES MATIERES

Sommaire	iii
Introduction	1
Chapitre I - Eléments de théorie du déséquilibre	4
Chapitre II - Econométrie du déséquilibre	10
1. Modèles examinés par Maddala et Nelson (1974)	11
2. Modèle avec fonction d'ajustement déterministe des prix	16
2.1. Fonction de vraisemblance	16
2.2. Extensions : variables endogènes retardées autocorrélation des erreurs (AR(1))	19
3. Modèle avec fonction d'ajustement stochastique des prix ..	23
3.1. Fonction de vraisemblance	23
3.2. Extensions : autocorrélation des erreurs	25
variables endogènes retardées	25
4. Equation d'ajustement des prix et test de l'hypothèse de déséquilibre	30
Chapitre III - Modèle dynamique avec ajustement stochastique des prix	34
1. Modèle de Dagenais (1980)	35
2. Estimation du modèle	39
2.1. Modèle à estimer	39
2.2. Modèle en forme réduite	42
- Identification	44
- Structure du terme d'erreur	45
- Fonction de vraisemblance	47
2.3. Maximisation de la fonction de vraisemblance	48
- Maximisation non linéaire	48
- "Moindres carrés itérés"	51
2.4. Estimation des constantes	55
Chapitre IV - Application au marché du bois de sciage américain	57
1. Caractéristiques générales du marché	58
2. Spécification	59
2.1. Existante	59
2.2. Retenue	64
3. Equation estimée	65
4. Résultats	67

Chapitre V - Régressions ridge, bayésienne et mixte, non linéaires	71
1. Linéarisation du modèle	73
2. Régression ridge	74
3. Régression bayésienne	77
4. Régression mixte	80
Conclusion	84
Appendice A	88
Appendice B	91
Remerciements	97
Bibliographie	99

INTRODUCTION

Les études économiques dans un cadre de déséquilibre se font de plus en plus fréquentes depuis les premiers travaux de Patinkin (1948) et Clower (1966) en théorie économique, et de Fair et Jaffee (1972) et Maddala et Nelson (1974) en économétrie.

Le modèle étudié demeure cependant statique. Dans notre étude nous voulons examiner une version dynamique du modèle de déséquilibre, celle de Dagenais (1980). Nous voulons, en particulier, nous pencher sur les méthodes d'estimation qui peuvent lui convenir et en utiliser une pour estimer les paramètres de l'offre et de la demande sur un marché concret auquel nous appliquerons notre modèle.

Notre principale contribution dans ce mémoire consistera d'ailleurs dans la présentation et l'expérimentation d'un algorithme permettant d'estimer une équation non linéaire à l'aide de la méthode des moindres carrés linéaires, sous la condition cependant que l'équation soit de forme multilinéaire.

Dans une première section nous examinerons brièvement quelques éléments de la théorie économique du déséquilibre. Nous verrons ensuite les principaux modèles économétriques de déséquilibre étudiés dans la littérature, leur extension possible en modèle dynamique (ou à autocorrélation

des erreurs). La troisième section présentera le modèle dynamique que nous étudions et la méthode d'estimation que nous lui appliquerons. La quatrième section rapportera les détails de l'application : le marché étudié (celui du bois de scie - Southern Pine - aux Etats-Unis), la spécification retenue, l'équation estimée et les résultats. Une dernière section proposera des voies de recherche afin de contrer des problèmes de multicollinéarité qui pourraient survenir. En conclusion nous ferons un bilan provisoire sur le modèle étudié et sur son intérêt empirique.

CHAPITRE I

Eléments de théorie du déséquilibre

Une des questions importantes qui s'est posée aux premiers économistes à l'époque où le libéralisme et la liberté individuelle s'affirmaient, a été celle de la coordination, de l'harmonie ultime des actions individuelles et égoïstes des différents agents économiques. A cette époque, comme maintenant, tous étaient confrontés au double phénomène de l'existence continue du système économique, jour après jour, et de la récurrence cyclique des crises économiques. On pouvait alors souligner, comme le faisait Adam Smith, le fonctionnement relativement harmonieux du système et l'exprimer par la métaphore de la "main invisible", ce qu'on appellerait aujourd'hui le mécanisme d'ajustement automatique d'un marché compétitif. On pouvait aussi, d'autre part, insister avec Marx sur les crises et anticiper un éventuel effondrement total du système économique et social. Les prédictions apocalyptiques de ce dernier ne doivent pas cependant nous faire perdre de vue la pertinence de la difficulté initiale. C'est essentiellement toute la question des ajustements du marché et du cheminement vers l'équilibre qui se pose.

Devant ce problème, la théorie économique nous présente le paradigme du marché de concurrence parfaite. Dans un tel marché tous les agents ne prennent leurs décisions qu'à partir de leurs préférences et des prix. Ils déterminent ainsi leurs fonctions de demande et d'offre

ex ante. Les quantités ne sont pas une contrainte pour eux puisque, au contraire, ils décident de ces quantités. Par ailleurs, les plans fixés doivent se réaliser : ils sont déterminés aux prix d'équilibre et donc l'égalité de l'offre et de la demande survient nécessairement. Il n'y a ici aucune place pour le déséquilibre, les offres et les demandes réalisées (ex post) correspondant aux quantités prévues (ex ante).

Il faut expliquer cependant comment on arrive à ces prix d'équilibre. Plusieurs économistes ont proposé des modèles de détermination des prix d'équilibre. Le plus connu est celui de Walras. Il pose un commissaire-priseur fictif qui crie les prix et les rajuste jusqu'à ce qu'ils soient tels que les intentions d'achat et de vente des agents s'accordent. Il transmet ainsi gratuitement à tous les agents économiques l'information sur les prix. Tant qu'un prix d'équilibre n'est pas atteint, aucune transaction ne s'effectue. Cette condition rapproche la solution de Walras de celle de Edgeworth. Pour ce dernier, les agents transigent, mais de façon conditionnelle : tant qu'un des agents désire renégocier le contrat, parce que le prix n'est pas encore d'équilibre, le contrat n'est pas définitif. L'échange de biens n'a donc lieu qu'après qu'on soit arrivé à un contrat définitif, à un prix d'équilibre.

Ce modèle repose sur plusieurs présuppositions. Il y aurait d'abord une centralisation de l'activité économique et de l'information qu'on ne rencontre, en vérité, que dans les marchés boursiers. Le monde économique est au contraire caractérisé par un grand nombre de lieux de

décision et de transaction relativement séparés les uns des autres. L'information n'y est jamais complète, certaine, gratuite.

Les prix sont supposés parfaitement et instantannément flexibles à la hausse comme à la baisse. Il n'y aurait aucune rigidité, qu'elle soit d'ordre technique, psychologique et institutionnel, qui viendrait retarder ou bloquer le processus d'ajustement. L'instantanéité signifie que tous les prix qu'on peut observer, à quelque fréquence que ce soit, sont toujours des prix d'équilibre.

Aucune transaction ne se ferait hors des prix d'équilibre. Cette proposition implique que dans le moment qui "précède" l'ajustement instantané aucune transaction ne s'effectue. Dans le cadre d'une économie réelle, c'est dire qu'entre les moments où les prix sont observés, et où se déroulent les processus d'ajustement, on ne transige pas. Pour s'assurer de cette condition, Marshall pose le concept de courte période, et Hicks celui d'équilibre temporaire (et de 'semaine') qui désignent des laps de temps durant lesquels on suppose que les prix sont d'équilibre et qu'ils ne changent pas.

La non satisfaction d'un ou de plusieurs de ces points peut entraîner un déséquilibre observable par un des agents économiques, celui qui est rationné. Il y a alors des transactions qui s'effectuent à un prix qui n'est pas un prix d'équilibre et, par conséquent, il y a des agents qui sont rationnés. C'est là le champ d'étude de l'analyse du déséquilibre.

Le phénomène principal que cette analyse voulait expliquer, quand elle a commencé à se manifester, est celui du chômage involontaire keynésien. Clower va insister - après Patinkin - sur l'effet de débordement (ou de report, "spillover") qui en découle. Le chômage involontaire, par exemple, va entraîner une baisse sur un autre marché, celui de la consommation des biens et services. C'est donc dire qu'il faut faire entrer dans l'analyse plus que les prix; la dotation initiale et les prix attendus. Il faut aussi tenir compte des contraintes effectives qui font que ce qui est désiré - les courbes d'offre et de demande traditionnelles - ne peut se réaliser. Les transactions se produisent alors sur d'autres courbes, dites effectives.

Ces phénomènes de débordement caractérisent l'approche du déséquilibre. On peut distinguer deux types de débordement. Le premier, le "spillover", est un report contemporain, instantané. Il correspondrait, dans une perspective statique, à l'imposition de contraintes d'un marché sur un autre. C'est l'effet dont on a le plus parlé dans la littérature du déséquilibre.

Le deuxième type de débordement correspondrait à un report temporel. Il s'agirait alors de l'effet d'une période sur la suivante, dans un même marché. Cet effet se rencontre lorsque l'on tient compte des inventaires des demandeurs ou des offreurs, ou bien de la demande ou de l'offre excédentaire d'une période qui se reporte sur la suivante. On peut ainsi avoir une propagation du déséquilibre. Il est aussi possible de construire un type mixte où le report temporel s'accompagnerait de débordements sur un autre marché.

Pour l'essentiel, la théorie du déséquilibre a, jusqu'ici, été surtout développée pour répondre à des questions d'ordre macroéconomique. Elle peut, de plus, être considérée comme une sérieuse tentative de fondation microéconomique de la macroéconomie, dans la mesure où elle veut retrouver sous les agrégats et les relations macroéconomiques les comportements des agents au niveau des industries et des marchés microéconomiques. Il s'agit là, somme toute, d'une théorie du déséquilibre général.

On peut cependant penser aussi au déséquilibre au niveau d'un seul marché, et quitter ainsi l'approche générale. On laisse ainsi tomber les reports contemporains ("spillover") pour se concentrer sur la dynamique de l'ajustement des prix. L'économétrie du déséquilibre s'est surtout penchée sur de tels modèles.

CHAPITRE II

Econométrie du déséquilibre

1. Modèles examinés par Maddala et Nelson (1974)

L'économétrie du déséquilibre est encore plus jeune que la théorie économique du déséquilibre. L'impulsion a été donnée par l'article de Fair et Jaffee (1972), mais c'est celui de Maddala et Nelson (1974) qui doit être considéré comme le travail de base en économétrie du déséquilibre.

Le modèle type qu'ils étudient contient trois équations :

$$(2.1) \quad D_t = \underline{X}_{dt}' \underline{\alpha}_1 + u_{1t}$$

$$(2.2) \quad S_t = \underline{X}_{st}' \underline{\beta}_1 + u_{2t}$$

$$(2.3) \quad Q_t = \min(D_t, S_t)$$

D_t et S_t désignent les quantités demandées et offertes durant la période t , \underline{X}_{dt} et \underline{X}_{st} les vecteurs des variables qui influencent respectivement D_t et S_t , et u_{1t} et u_{2t} les termes d'erreur. Q_t est la quantité observée et transigée durant la période t ; elle correspond à la plus faible des deux valeurs D_t et S_t qui sont inobservables.

En ajoutant une équation d'ajustement des prix, ils obtiennent les quatre modèles suivants :

Modèle 1

- Comprend les équations (2.1), (2.2) et (2.3);
- le prix, P_t , peut se trouver de façon exogène dans \underline{X}_{dt} et/ou dans \underline{X}_{st} ;
- Q_t est observée.

Modèle 2

- Comprend les équations (2.1), (2.2) et (2.3);
- le prix, P_t , ne peut se trouver ni dans \underline{X}_{dt} ni dans \underline{X}_{st} ;
- on observe P_t et Q_t ;
- on pose que :

$$(2.4) \quad \begin{array}{lll} \Delta P_t > 0 & \text{si} & D_t > S_t \\ \Delta P_t < 0 & \text{si} & D_t < S_t \end{array}$$

où

$$\Delta P_t \equiv P_t - P_{t-1}$$

Modèle 3

- Les équations (2.1) et (2.2) sont modifiées ainsi :

$$(2.1') \quad D_t = \underline{X}'_{dt} \alpha_1 + \alpha_2 P_t + u_{1t}$$

$$(2.2') \quad S_t = \underline{X}'_{st} \beta_1 + \beta_2 P_t + u_{2t} \quad ;$$

- la condition (2.3) est inchangée;
- l'équation d'ajustement des prix est déterministe :

$$(2.4') \quad \Delta P_t = \gamma(D_t - S_t) ;$$

- on observe P_t et Q_t .

Modèle 4

- Comprend les équations (2.1'), (2.2') et (2.3);
- l'ajustement des prix est stochastique :

$$(2.4'') \quad \Delta P_t = \gamma(D_t - S_t) + u_{3t}$$

Quelques remarques au sujet de ces modèles sont nécessaires.

D'abord, ce sont tous des modèles de déséquilibre partiel, qui n'examinent qu'un seul marché à la fois. C'est donc dire qu'on ne peut inclure ici aucun effet de débordement, typique pourtant des modèles de la théorie du déséquilibre.

De plus ces modèles sont statiques. Il n'y a pas de liaison entre les périodes. Les reports temporels ne peuvent y être considérés. Seul est examiné le processus d'ajustement des prix et l'estimation des paramètres de l'offre et de la demande quand la quantité transigée et observée est la plus petite de la quantité offerte et de la quantité demandée.

Cette condition du minimum, exprimée par l'équation (2.3) nous indique que l'échange est volontaire, c'est-à-dire qu'aucun agent n'est forcé de vendre plus qu'il n'offre ou d'acheter moins qu'il ne demande, au prix donné (qui n'est pas le prix d'équilibre). Elle implique du même coup qu'on ne peut retrouver, à la fois, des vendeurs et des acheteurs rationnés sur un même marché.

Les quatre modèles diffèrent par ailleurs dans leur traitement de l'ajustement des prix - quand il est spécifié -, et par l'information qu'ils permettent (ou non) d'obtenir sur la répartition de chaque quantité transigée entre l'offre et la demande. Ces deux facteurs ont une influence déterminante sur la fonction de vraisemblance de Q_t , selon les modèles.

Dans le premier modèle, aucun processus d'ajustement des prix n'est présent : ils sont complètement exogènes et les déséquilibres ne les affectent pas. Il n'y a pas de connaissance préalable de l'appartenance de Q_t à D_t ou à S_t .

Dans le deuxième modèle, on peut associer chaque Q_t à la quantité offerte ou demandée grâce à l'information sur ΔP_t . Le prix est ici endogène, il est déterminé par le système. On ne peut par ailleurs l'inclure parmi les variables explicatives car il n'y a pas d'équation d'ajustement des prix. On ne peut donc déterminer la densité conjointe de Q_t et de P_t , car on n'a pas assez d'équations pour le faire.

Le troisième modèle va inclure une équation déterministe d'ajustement partiel et instantané des prix. On a dès lors deux variables endogènes conjointes, P_t et Q_t . Le signe de ΔP_t nous permet ici de distinguer les observations qui sont sur la courbe de demande de celles qui sont sur la courbe d'offre.

Le dernier modèle enfin est semblable au troisième à la différence que l'équation d'ajustement des prix est stochastique. On ne peut donc plus séparer nos observations Q_t . C'est le modèle qui s'approche le plus des intuitions habituelles d'ajustement des prix dans un marché. De plus, il est possible d'ajouter un vecteur de variables explicatives dans l'équation d'ajustement (2.4''). (On le pouvait aussi avec (2.4'), bien sûr). Ce vecteur permettrait de varier les hypothèses sur la structure concurrentielle du marché étudié. Nous en reparlerons plus loin.

Maddala et Nelson présentent des méthodes d'estimation par le maximum de vraisemblance pour les quatre modèles. Nous ne passerons pas en revue les fonctions de vraisemblance qu'ils développent pour les modèles 1 et 2. Leur traitement est très détaillé et par ailleurs ces deux modèles nous intéressent moins car ils ne comportent pas d'équation d'ajustement des prix. Nous examinerons plutôt les modèles 3 et 4 qui ont été l'objet d'extensions par Laffont et Monfort, Bowden, Ito et Ueda, Dagenais et enfin Orsi.

2. Modèle avec fonction d'ajustement des prix déterministe : Modèle 3

2.1. Fonction de vraisemblance

Supposons que les erreurs aléatoires u_{1t} et u_{2t} suivent une loi multinormale de moyenne nulle et de variance Σ . Elles ne sont pas auto-corrélées dans le temps :

$$(2.5) \quad \text{Cov}(u_{it}, u_{jt-m}) = 0 \quad \begin{array}{l} i, j = 1, 2 \\ m = 1, 2, \dots \end{array}$$

Cependant la corrélation instantannée n'est pas nulle :

$$(2.6) \quad \text{Cov}(u_{it}, u_{jt}) = \sigma_{ij} = \Sigma_{ij} \neq 0 \quad i, j = 1, 2$$

La fonction de densité conjointe de P_t et Q_t peut alors être déterminée par les étapes suivantes :

1. On sépare l'échantillon selon le signe de ΔP_t .

Si $\Delta P_t > 0$, alors $D_t > S_t$ et $Q_t = S_t$;

Si $\Delta P_t < 0$, alors $D_t < S_t$ et $Q_t = D_t$.

Les points où les prix seraient d'équilibre ($\Delta P_t = 0$) ne posent pas de problème car leur probabilité est nulle si les variables u_{1t} et u_{2t} sont continues. Il suffit en pratique de rattacher arbitrairement ces observations à l'un ou l'autre des deux groupes, l'offre ou la demande.

2. Pour un premier groupe, où $Q_t = S_t$, on isole les variables dépendantes observables. On obtient :

$$(2.7) \quad S_t = Q_t = \frac{X'_t}{-s_t} \underline{\beta}_1 + \beta_2 P_t + u_{2t}$$

et

$$(2.8) \quad \Delta P_t = \gamma(D_t - S_t) = \gamma\left(\frac{X'_t}{-dt} \underline{\alpha}_1 + \alpha_2 P_t + u_{1t} - Q_t\right);$$

de là, on peut tirer

$$(2.9) \quad P_t = \frac{X'_t}{-dt} \frac{\underline{\alpha}_1 t}{1 - \alpha_2 \gamma} + \frac{1}{1 - \alpha_2 \gamma} P_{t-1} - \frac{\gamma}{1 - \alpha_2 \gamma} Q_t + \frac{\gamma}{1 - \alpha_2 \gamma} u_{1t}$$

3. Des équations (2.7) et (2.9), connaissant la densité conjointe de u_{1t} et u_{2t} , on peut tirer la densité conjointe de P_t et de Q_t , quand $Q_t = S_t$. Désignons cette densité par $f_s(P_t, Q_t)$. On cherche alors :

$$(2.10) \quad f_s(P_t, Q_t) = |\text{dét } J| f(\underline{u}_t)$$

où $\underline{u}_t = (u_{1t}, u_{2t})$

et $f(\underline{u}_t) = (2\pi)^{-1} |\Sigma|^{-1/2} \exp[-1/2 \underline{u}_t \Sigma^{-1} \underline{u}_t']$

Il nous faut d'abord trouver $|\text{dét } J|$, la valeur absolue du déterminant de la matrice jacobienne des dérivées premières. On l'obtient ici par l'application en chaîne

$$(2.11) \quad \left| \det \begin{bmatrix} \frac{\partial \underline{u}_t}{\partial u_{1t}^*} \\ \frac{\partial \underline{u}_t}{\partial u_{2t}^*} \end{bmatrix} \cdot \det \begin{bmatrix} \frac{\partial u_t^*}{\partial (P_t, Q_t)} \end{bmatrix} \right| = \left| \det \begin{bmatrix} \frac{\partial \underline{u}_t}{\partial (P_t, Q_t)} \end{bmatrix} \right|$$

où $\underline{u}_t^* = (u_{1t}^*, u_{2t}^*)$ et $u_{1t}^* = \frac{\gamma}{1 - \alpha_2 \gamma} u_{1t}$

En utilisant les équations (2.7) et (2.9), on trouve

$$(2.12) \quad \left| \det \begin{bmatrix} \frac{\partial u_{1t}}{\partial u_{1t}^*} & \frac{\partial u_{1t}}{\partial u_{2t}^*} \\ \frac{\partial u_{2t}}{\partial u_{1t}^*} & \frac{\partial u_{2t}}{\partial u_{2t}^*} \end{bmatrix} \cdot \det \begin{bmatrix} \frac{\partial u_{1t}^*}{\partial P_t} & \frac{\partial u_{1t}^*}{\partial Q_t} \\ \frac{\partial u_{2t}^*}{\partial P_t} & \frac{\partial u_{2t}^*}{\partial Q_t} \end{bmatrix} \right| =$$

$$\left| \det \begin{bmatrix} \frac{1 - \alpha_2 \gamma}{\gamma} & 0 \\ 0 & 1 \end{bmatrix} \cdot \det \begin{bmatrix} 1 & \frac{\gamma}{1 - \alpha_2 \gamma} \\ -\beta_2 & 1 \end{bmatrix} \right| = |1/\gamma - \alpha_2 + \beta_2|$$

4. En répétant les étapes 2 et 3, pour le groupe des $Q_t = D_t$, on obtient la même valeur pour $|\det J|$.
5. On peut alors formuler la fonction de vraisemblance du modèle 3. Sa forme générale est

$$(2.13) \quad L = \prod_{\Delta P_t > 0} \pi f_{s,t}(P_t, Q_t) \cdot \prod_{\Delta P_t < 0} \pi f_{d,t}(P_t, Q_t)$$

puisque l'échantillon est séparable de façon certaine.

Si on développe $f_s(P_t, Q_t)$ sous l'hypothèse que $\underline{u}_t \sim N_2(0, \Sigma)$, on obtient :

$$(2.14) \quad \frac{|1/\gamma - \alpha_2 + \beta_2|}{2\pi |\Sigma|^{1/2}} \exp \{- 1/2 \underline{u}_t \Sigma^{-1} \underline{u}_t'\}$$

où

$$(2.15) \quad \underline{u}_t = \begin{bmatrix} Q_t + \Delta P_t / \gamma - \frac{X'_t}{dt} \underline{\alpha}_1 - \beta_1 P_t \\ Q_t - \frac{X'_t}{st} \underline{\beta}_1 - \beta_2 P_t \end{bmatrix}'$$

Dans le cas de $f_d(P_t, Q_t)$, on garde (2.14) mais alors

$$(2.16) \quad \underline{u}_t = \begin{bmatrix} Q_t - \frac{X'_t}{dt} \underline{\alpha}_1 - \alpha_2 P_t \\ Q_t + \Delta P_t / \gamma - \frac{X'_t}{st} \underline{\beta}_1 - \beta_2 P_t \end{bmatrix}'$$

2.2. Extensions du modèle 3

Laffont et Monfort (1976, 1979) et Orsi (1982) ont proposé d'intéressantes extensions au modèle 3. Laffont et Monfort, en particulier, ont formulé la vraisemblance du modèle quand les erreurs sont autocorrélées (AR(1)) et quand il y a des variables endogènes retardées inobservables parmi les variables explicatives.

Dans ce dernier cas, le modèle s'écrit :

$$(2.17) \quad D_t = \frac{X'_t}{dt} \underline{\alpha}_1 + \alpha_2 P_t + \alpha_3 D_{t-1} + \alpha_4 S_{t-1} + u_{1t}$$

$$(2.18) \quad S_t = \frac{X'_t}{st} \underline{\beta}_1 + \beta_2 P_t + \beta_3 D_{t-1} + \beta_4 S_{t-1} + u_{2t}$$

et l'ajustement (2.4') de même que la condition du minimum (2.3) demeurent.

La fonction de vraisemblance se calcule en passant par les mêmes étapes que plus haut.

La difficulté dans ces modèles dynamiques (ou lorsqu'il y a autocorrélation des erreurs) tient aux variables inobservables qui s'accumulent alors et qu'on ne peut éliminer qu'en prenant leur intégrale, en calculant leur probabilité. Les méthodes numériques dont nous disposons ne nous permettent pas de calculer de façon opérationnelle des intégrales ayant plus de cinq dimensions.

Cette difficulté à remplacer les variables inobservables peut être contournée dans le cas du modèle 3, car la relation déterministe d'ajustement des prix (2.4') nous permet de connaître avec certitude D_t (ou S_t) quand on sait que $Q_t = S_t$ (ou $Q_t = D_t$).

La vraisemblance sera donnée par la fonction

$$(2.19) \quad L = \prod_{t=2}^T h_t(Q_t, P_t, Q_{t-1}, P_{t-1}, \dots, Q_1, P_1)$$

$$(2.20) \quad h_t(Q_t, P_t) = |\text{dét } J| f(\underline{u}_t)$$

$$\text{où } |\text{dét } J| = |1/\gamma + \beta_2 - \alpha_2|$$

et, si $Q_t = S_t$,

$$\underline{u}_t = \begin{bmatrix} Q_t + \frac{\Delta P_t}{\gamma} - \frac{X'_{dt}}{\alpha_1} - \alpha_2 P_t - (\alpha_3 + \alpha_4)Q_{t-1} - \frac{\alpha_3}{\gamma} \Delta P_{t-1} \\ Q_t - \frac{X'_{st}}{\beta_1} - \beta_2 P_t - (\beta_3 + \beta_4)Q_{t-1} - \frac{\beta_3}{\gamma} \Delta P_{t-1} \end{bmatrix}'$$

et, si $Q_t = D_t$

$$\underline{u}_t = \begin{bmatrix} Q_t - \frac{X'_{dt}}{\alpha_1} - \alpha_2 P_t - (\alpha_3 + \alpha_4)Q_{t-1} + \frac{\alpha_4}{\gamma} \Delta P_{t-1} \\ Q_t - \frac{\Delta P_t}{\gamma} - \frac{X'_{st}}{\beta_1} - \beta_2 P_t - (\beta_3 + \beta_4)Q_{t-1} + \frac{\beta_4}{\gamma} \Delta P_{t-1} \end{bmatrix}'$$

Selon la valeur que prennent Q_t et Q_{t-1} on aura donc quatre fonctions h_{dd} , h_{ds} , h_{ss} , h_{sd} et la fonction de vraisemblance s'écrira :

$$(2.21) \quad L = \begin{matrix} \pi & h_{dd,t}(Q_t, P_t) & \cdot & \pi & h_{ds,t}(Q_t, P_t) \\ \Delta P_t < 0 & & & \Delta P_t < 0 \\ \Delta P_{t-1} < 0 & & & \Delta P_{t-1} > 0 \\ \cdot & & & & \\ \pi & h_{ss,t}(Q_t, P_t) & \cdot & \pi & h_{sd,t}(Q_t, P_t) \\ \Delta P_t > 0 & & & \Delta P_t > 0 \\ \Delta P_{t-1} > 0 & & & \Delta P_{t-1} < 0 \end{matrix}$$

Sous l'hypothèse que $\underline{u}_t \sim N_2(\underline{0}, \Sigma)$ on obtiendrait alors des équations analogues aux équations (2.14) et (2.15), (2.16).

Un modèle ayant des erreurs qui suivent un processus autorégressif du premier ordre (AR(1))

$$(2.22) \quad \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \underset{(2 \times 2)}{R} \begin{bmatrix} u_{1t-1} \\ u_{2t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \quad \text{où } \underline{\varepsilon}_t \sim N_2(\underline{0}, \Sigma)$$

pourra se ramener au cas précédent par simple réécriture. Si on écrit le modèle sous la forme

$$(2.23) \quad \begin{bmatrix} D_t \\ S_t \end{bmatrix} = A y_t + B P_t + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}$$

on obtient :

$$(2.24) \quad \begin{bmatrix} D_t \\ S_t \end{bmatrix} = R \begin{bmatrix} D_{t-1} \\ S_{t-1} \end{bmatrix} + A y_t - R A y_{t-1} + B P_t - R B P_{t-1} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

Ce modèle se ramène au modèle (2.17), (2.18), avec en plus, des contraintes sur les paramètres.

Le modèle dynamique de Orsi (1982) est de la même famille. Il s'écrit (dans notre notation) :

$$(2.25) \quad D_t = \frac{X'_t}{dt} \alpha_1 + \alpha_2 P_t + \alpha_3 (D_{t-1} - Q_{t-1}) + u_{1t}$$

$$(2.26) \quad S_t = \frac{X'_t}{st} \beta_1 + \beta_2 P_t + \beta_3 (S_{t-1} - Q_{t-1}) + u_{2t}$$

Il peut être vu comme un cas particulier de (2.17), (2.18). La fonction de vraisemblance se formule de la même façon que (2.21) en utilisant l'équation d'ajustement des prix pour exprimer la variable inobservable.

L'estimation de ces modèles peut se faire sans trop de difficultés. Dans le cas standard, (modèle (2.1'), (2.2'), (2.3) et (2.4')), on pourrait estimer par doubles moindres carrés non linéaires (voir Amemiya, 1974a, 1974b, 1975). On peut aussi utiliser la méthode du maximum de vraisemblance. Cette dernière méthode est aussi applicable pour les modèles à erreurs autocorrélées et/ou à variables endogènes retardées. Il est alors nécessaire de choisir un algorithme de maximisation approprié aux caractéristiques de la fonction. Pour le choix des valeurs initiales pour la maximisation, Laffont et Monfort (1979) recommandent une méthode à information limitée ou celle des doubles moindres carrés, qui donnent des estimations convergentes.

3. Modèle avec fonction d'ajustement des prix stochastique : modèle 4

3.1. Fonction de vraisemblance

Le modèle 4 se distingue du modèle 3 par la présence d'un terme d'erreur non nul dans l'équation d'ajustement des prix (2.4''). Cette différence est lourde de conséquences car elle élimine la possibilité de remplacer toutes les variables inobservables par des quantités observées, connaissant Q_t et ΔP_t . On ne peut plus séparer l'échantillon avec certitude.

On cherche la vraisemblance du modèle. On trouve facilement :

$$(2.27) \quad g_t(D_t, S_t, P_t | X_t) = |\text{dét } J| f_t(\underline{u}_t)$$

$$\text{où } \underline{u}_t = (u_{1t}, u_{2t}, u_{3t})$$

et donc

$$g_t(D_t, S_t, P_t | X_t) = |1 + \gamma(\beta_2 - \alpha_2)| f_t(\underline{u}_t)$$

$$\text{où } \underline{u}_t = \begin{bmatrix} D_t - \alpha_2 P_t - \frac{X'_t}{dt} \alpha_1 \\ S_t - \beta_2 P_t - \frac{X'_t}{st} \beta_1 \\ P_t - P_{t-1} - \gamma(D_t - S_t) \end{bmatrix},$$

Utilisant ensuite l'équation (2.3), on peut exprimer la vraisemblance par

$$(2.28) \quad L = \prod_{t=1}^n h_t(Q_t, P_t | X_t) = \prod_{t=1}^n \left[\int_{Q_t}^{\infty} g_t(Q_t, S_t, P_t | X_t) dS_t \right. \\ \left. + \int_{Q_t}^{\infty} g_t(D_t, Q_t, P_t | X_t) dD_t \right]$$

Selon l'hypothèse retenue concernant la loi de distribution des \underline{u}_t , on peut ensuite, sans trop de difficultés, expliciter la fonction de vraisemblance.

La méthode d'estimation qui se présente d'abord - si le modèle est identifié - est celle du maximum de vraisemblance à information complète (FIML). Cette méthode est cependant lourde à utiliser en pratique.

On trouve aussi dans Maddala et Nelson (1974) une méthode à information limitée qui est corrigée par Sen (1977) pour rendre ses estimateurs convergents.

3.2. Extensions du modèle 4

Les extensions examinées dans le cadre du modèle 3 ne sont plus aussi simples dans ce nouveau modèle.

L'autocorrélation des erreurs d'abord. Comme l'ont affirmé Maddala et Nelson, cette situation est très difficile dans le modèle 4. Dans le cas le plus simple où il n'y a qu'une seule variable endogène non observée, et seulement deux observations, la fonction de densité conjointe des deux observations, connaissant celle qui les précède, serait

$$(2.29) \quad f(X_1, X_2 | X_0) = f(X_2 | X_1, X_0) \quad f(X_1 | X_0)$$

La difficulté est toute entière dans le fait qu'on ne peut, par hypothèse, ramener $f(X_2 | X_1, X_0)$ à $f(X_2 | X_0)$.

Dans le cas général, supposons que nous avons n observations, que l'on connaît la valeur initiale $Q_0 = D_0 = S_0$ - que l'on suppose d'équilibre -, et qu'il y a deux possibilités pour chaque observation, selon la condition du minimum

$$(2.30) \quad \begin{array}{l} D_i = Q_i \quad \text{et} \quad S_i > Q_i \\ S_i = Q_i \quad \text{et} \quad D_i > Q_i \end{array}$$

La fonction de densité conjointe des variables D_t et S_t s'écrirait alors

$$(2.31) \quad f(D_n, S_n, D_{n-1}, S_{n-1}, \dots, D_1, S_1 | Q_0)$$

Avant d'écrire la fonction de vraisemblance il faut remarquer que chacune des n observations est une paire de variables dont l'une reste inconnue, inobservable. On trouve facilement qu'il y a 2^n combinaisons différentes de variables observées et non observées pour notre échantillon. La fonction de vraisemblance prendra donc la forme :

$$(2.32) \quad \begin{aligned} L = & \int_{Q_n}^{\infty} \dots \int_{Q_1}^{\infty} f(Q_n, S_n, Q_{n-1}, S_{n-1}, \dots, Q_1, S_1 | Q_0) dS_1 \dots dS_n \\ & + \int_{Q_n}^{\infty} \dots \int_{Q_1}^{\infty} f(Q_n, S_n, D_{n-1}, Q_{n-2}, S_{n-2}, \dots, Q_1, S_1 | Q_0) \\ & dS_1 \dots dS_{n-2} dD_{n-1} dS_n \\ & + \dots \\ & + \int_{Q_n}^{\infty} \dots \int_{Q_1}^{\infty} f(D_n, Q_n, D_{n-1}, Q_{n-1}, \dots, D_1, Q_1 | Q_0) dD_1 \dots dD_n \end{aligned}$$

Chacun des 2^n termes de cette fonction est une intégrale à n dimensions, pour tenir compte de la probabilité des variables non observées. La condition du minimum nous fait intégrer de Q_i ($i = 1, \dots, n$) à l'infini.

Chaque terme peut être réécrit comme le produit d'une conditionnelle et d'une marginale :

$$\begin{aligned}
 (2.33) \quad L &= \int_{Q_n}^{\infty} \dots \int_{Q_1}^{\infty} f(Q_n, S_n | Q_{n-1}, S_{n-1}, \dots, Q_1, S_1, Q_0) \\
 &\cdot f(Q_{n-1}, S_{n-1}, \dots, Q_1, S_1 | Q_0) dS_1 \dots dS_n \\
 &+ \dots
 \end{aligned}$$

Cette réécriture de chaque terme pourrait se poursuivre jusqu'à ce qu'on ait n-1 conditionnelles - une pour chaque observation conditionnelle à celles qui la précèdent -, et une marginale. On aurait alors pour le premier terme de L

$$\begin{aligned}
 (2.34) \quad &\int_{Q_n}^{\infty} \dots \int_{Q_1}^{\infty} f(Q_n, S_n | Q_{n-1}, S_{n-1}, \dots, Q_1, S_1, Q_0) \\
 &\cdot f(Q_{n-1}, S_{n-1} | Q_{n-2}, S_{n-2}, \dots, Q_1, S_1, Q_0) \\
 &\quad \vdots \\
 &\cdot f(Q_2, S_2 | Q_1, S_1, Q_0) \\
 &\cdot f(Q_1, S_1 | Q_0) dS_1 \dots dS_n
 \end{aligned}$$

La formulation (2.34) est la plus simple si on suppose que l'autocorrélation est d'ordre n, que l'observation présente dépend de toutes celles qui précèdent. Si l'autocorrélation est du premier ordre on peut factoriser l'expression (2.34).

$$(2.35) \quad \int_{Q_1}^{\infty} \left\{ \int_{Q_2}^{\infty} \left[\dots \left[\int_{Q_n}^{\infty} f(Q_n, S_n | Q_{n-1}, S_{n-1}) dS_n \right] \dots \right] f(Q_2, S_2 | Q_1, S_1) dS_2 \right\} f(Q_1, S_1 | Q_0) dS_1$$

On se retrouve donc à la fin avec, pour chaque terme, n intégrales simples emboîtées. Si on suppose, de plus, qu'à chaque observation il y a indépendance entre les quantités offertes et les quantités demandées, on peut finalement réécrire (2.35) sous la forme

$$(2.36) \quad f_d(Q_n | Q_{n-1}) \cdot \dots \cdot f_d(Q_1 | Q_0) \cdot \int_{Q_1}^{\infty} \left\{ \dots \left[\int_{Q_n}^{\infty} f_s(S_n | S_{n-1}) dS_n \right] \dots \right\} f_s(S_1 | Q_0) dS_1$$

Cependant, il faut remarquer que la simplification introduite n'est qu'un effet de la notation. Dans les conditionnelles :

$$\int_{Q_t}^{\infty} f_s(S_t | S_{t-1}) dS_t$$

on ne connaît pas la variable S_{t-1} . On ne peut donc pas calculer l'intégrale unidimensionnelle la plus profondément emboîtée et remonter jusqu'à

$$\int_{Q_1}^{\infty} \{ \dots \} f_s(S_1 | Q_0) dS_1$$

Le problème se réglerait si l'on pouvait ramener (2.36) à

$$(2.37) \quad f_d(Q_n | Q_0) \cdot \dots \cdot f_d(Q_1 | Q_0) \cdot \int_{Q_n}^{\infty} f_s(S_n | Q_0) \cdot \dots \cdot \int_{Q_1}^{\infty} f_s(S_1 | Q_0)$$

On n'aurait alors qu'un produit d'intégrales unidimensionnelles faciles à résoudre numériquement. C'est ce que Quandt (1981) a proposé. Cette suggestion ne peut malheureusement pas être retenue. Il faudrait pour que cela soit correct que la covariance entre Q_t et Q_{t-1} , étant donné Q_0 , soit nulle, ce qui n'est pas le cas. En fait si, par exemple :

$$Q_t = D_t \quad t = 0, 1, 2$$

$$u_{1t} = \rho_1 u_{1t-1} + \varepsilon_{1t} \quad \text{où } \varepsilon_t \sim N(\underline{0}, \Sigma), \text{ iid}$$

et D_t est spécifiée comme (2.1')

$$(2.1') \quad D_t = \underline{X}'_t \underline{\alpha}_1 + \alpha_2 P_t + u_{1t} \quad (\equiv \underline{X}'_t \underline{\alpha} + u_{1t}) ,$$

on obtient

$$(2.38) \quad D_t = \underline{X}'_t \underline{\alpha} + \rho_1 (D_{t-1} - \underline{X}'_{t-1} \underline{\alpha}) + \varepsilon_{1t}$$

et donc

$$D_1 = \rho_1 D_0 + (\underline{X}'_1 - \rho_1 \underline{X}'_0) \underline{\alpha} + \varepsilon_{11}$$

et

$$D_2 = \rho_1^2 D_0 + (\underline{X}'_2 - \rho_1 \underline{X}'_1) \underline{\alpha} + \rho_1 (\underline{X}'_1 - \rho_1 \underline{X}'_0) + \rho_1 \varepsilon_{11} + \varepsilon_{12}$$

La covariance entre D_1 et D_2 , étant donné D_0 , sera

$$(2.39) \quad \text{Cov}(D_2, D_1 | D_0) = \text{Cov}[(\rho_1 \varepsilon_{11} + \varepsilon_{12}), \varepsilon_{11}] = \rho_1 \sigma_1^2 \neq 0$$

L'autocorrélation des erreurs pose donc un problème pratique qui est pour l'instant insurmontable.

L'autre extension que nous avons examinée dans le cas du modèle précédent est l'inclusion de variables endogènes retardées parmi les variables explicatives. Il n'existe pas pour le modèle 4, à notre connaissance, de modèle dynamique, autre que celui suggéré par Dagenais (1980) que nous examinerons dans les sections suivantes. La quasi-absence de tels modèles se comprend quand on se rappelle le lien étroit entre les modèles dynamiques et les modèles à erreurs autocorrélées dont nous venons de montrer les difficultés.

Il reste un domaine de variation dont nous n'avons pas parlé, c'est celui des diverses formes que peut prendre l'équation d'ajustement des prix.

4. Equation d'ajustement des prix et test de l'hypothèse de déséquilibre

Les variations de cette équation sont nombreuses dans la littérature. Nous avons retenu plus haut (2.4), (2.4') et (2.4'') pour simplifier l'exposé.

On peut facilement concevoir un processus d'ajustement des prix qui soit

$$P_{t+1} - P_t = \gamma(D_t - S_t)$$

On pourrait aussi penser que la vitesse d'ajustement diffère selon que l'on se trouve dans une situation d'offre ou de demande excédentaire.

On aurait alors

$$\begin{aligned} P_t - P_{t-1} &= \gamma_1(D_t - S_t) & \gamma_1 > 0 & \text{ si } D_t > S_t \\ &= \gamma_2(D_t - S_t) & \gamma_2 > 0 & \text{ si } D_t < S_t \end{aligned}$$

Nous avons supposé plus haut que $\gamma_1 = \gamma_2 = \gamma$. On peut enfin penser que l'ajustement ne se fait pas de façon contemporaine, qu'il est retardé.

On aurait alors

$$P_t - P_{t-1} = \gamma(D_{t-1} - S_{t-1})$$

L'introduction de variables exogènes dans l'équation d'ajustement des prix permet d'examiner des hypothèses plus variées.

L'essentiel de la démarche examinée jusqu'ici peut être conservé lorsqu'on retient ces variations, bien que les résultats soient, bien sûr, légèrement différents. Quelques contraintes s'ajoutent cependant. Si l'on ne fait pas l'hypothèse de l'égalité des vitesses de réaction, il faut, pour que l'on puisse faire la transformation inverse, qui nous permet ensuite de trouver le déterminant du jacobien, supposer que $\alpha_2 > \alpha_1$. Sinon, la relation n'est plus biunivoque. On dit alors que le modèle n'est pas cohérent. Les conditions suffisantes et nécessaires de

cohérence (c'est-à-dire de l'existence d'une forme réduite unique) ont été énoncées par Gouriéroux, Laffont, Monfort (1980a, 1980b).

Un dernier développement dû à Bowden (1978) peut être mentionné. Il permet d'emboîter un modèle d'équilibre dans un modèle de déséquilibre de type 3 ou 4. En plus des équations (2.1') et (2.2'), Bowden en pose deux autres qui ne diffèrent que par le remplacement de P_t par P_t^* , le prix d'équilibre (inobservable), qui égalise l'offre et la demande. On a donc

$$(2.1'^*) \quad D_t^* = \frac{X}{dt} \alpha_1 + \alpha_2 P_t^* + u_{1t}$$

$$(2.2'^*) \quad S_t^* = \frac{X}{st} \beta_1 + \alpha_2 P_t^* + u_{2t}$$

Comme, par définition, $D_t^* = S_t^*$, on peut réécrire (2.4'') de la façon suivante :

$$\begin{aligned} (2.4''*) \quad \Delta P_t &= \gamma [(D_t - S_t) - (D_t^* - S_t^*)] + u_{3t} \\ &= \gamma [(D_t - D_t^*) - (S_t - S_t^*)] + u_{3t} \\ &= \gamma [(\alpha_2 (P_t - P_t^*) - \beta_2 (P_t - P_t^*))] + u_{3t} \\ &= \gamma [(\alpha_2 - \beta_2) (P_t - P_t^*)] + u_{3t} \end{aligned}$$

En isolant P_t , on obtient finalement

$$(2.4''*) \quad P_t = \mu P_{t-1} + (1-\mu) P_t^* + \mu u_{3t}$$

$$\text{où } \mu = \frac{1}{1 - \gamma(\alpha_2 - \beta_2)}$$

Sous cette forme on peut, après avoir estimé les paramètres structuraux, tester si μ est significativement différent de zéro. S'il ne l'est pas, on ne peut rejeter l'hypothèse que le modèle pertinent pour le marché étudié est un modèle d'équilibre. Ito et Ueda (1981) ont étendu ce test au cas où la vitesse d'ajustement à la hausse peut être différente de celle à la baisse.

CHAPITRE III

Modèle dynamique avec ajustement stochastique des prix

1. Modèle de Dagenais (1980)

Le modèle que nous voulons examiner est celui de Dagenais (1980). Il se distingue des modèles dont nous avons parlé dans la section précédente par sa prise en compte, simultanée, des effets de report de la demande et de l'offre - ce qui en fait un modèle dynamique -, et d'autre part de la nature stochastique du processus d'ajustement des prix. C'est pour le moment le seul modèle de déséquilibre à réunir ces deux caractéristiques, les autres modèles dynamiques posant une fonction déterministe d'ajustement des prix.

Les équations du modèle sont les suivantes :

$$(3.1) \quad D_t = \beta_0 + \frac{X'_t}{t} \beta_1 + \frac{X'_{dt}}{dt} \beta_2 + \beta_3 P_t + \beta_4 (D_{t-1} - Q_{t-1}) + u_{1t}$$

$$(3.2) \quad S_t = \lambda_0 + \frac{X'_t}{t} \lambda_1 + \frac{X'_{st}}{st} \lambda_2 + \lambda_3 P_t + \lambda_4 (I_{t-1} - I_t^d) + u_{2t}$$

$$(3.3) \quad Q_t = \min(D_t, S_t)$$

$$(3.4) \quad \Delta P_t = P_t - P_{t-1} = \gamma (D_{t-1} - S_{t-1}) + u_{3t}$$

$$(3.5) \quad I_t^d = \theta_0 + \theta_1 (\lambda_0 + \frac{X'_t}{t} \lambda_1 + \frac{X'_{st}}{st} \lambda_2 + \lambda_3 P_t + u_{2t}) + u_{4t}$$

où les symboles signifient :

- D, S : quantités demandées et offertes;
 Q : quantité effectivement transigée;
 P : prix observé;
 I^d : niveau d'inventaire désiré;
 $\underline{X}, \underline{X}_d, \underline{X}_s$: vecteurs de variables exogènes (vecteurs colonne);
 \underline{X} : variables communes à l'offre et à la demande;
 \underline{X}_d : variables propres à la demande;
 \underline{X}_s : variables propres à l'offre;
 I : niveau d'inventaire en fin de période;
 β_0, λ_0 : constantes;
 $\underline{\beta}_1, \underline{\beta}_2, \underline{\lambda}_1, \underline{\lambda}_2$: vecteurs de paramètres;
 $\beta_3, \beta_4, \lambda_3, \lambda_4$: paramètres (scalaires);
 $u_{1t}, u_{2t}, u_{3t}, u_{4t}$: erreurs résiduelles.

L'équation (3.1) linéarise la fonction de demande. La quantité demandée peut être fonction de facteurs communs à l'offre (\underline{X}), de facteurs spécifiques (\underline{X}_d), du prix du bien considéré (P), et de la demande excédentaire de la période précédente ($D_{t-1} - Q_{t-1}$). Cette dernière variable prédéterminée permet de dynamiser le système. La demande qui n'a pu être satisfaite à la période précédente se reporterait, en partie du moins, pour accroître la demande de la période suivante. On remarque que cette variable est inobservable, puisqu'on ne connaît pas D , et qu'on s'attend à ce que son coefficient soit positif, compris entre 0 et 1, puisque la demande excédentaire ne peut être que non négative. Le coefficient du prix devrait, bien entendu, être négatif.

L'équation d'offre (3.2) est elle aussi linéarisée. La quantité offerte peut être fonction de facteurs communs à l'offre et à la demande (X), de facteurs spécifiques (X_s) qui comprennent les coûts et les facteurs reliés à la production, du prix du bien (P), et d'une variable de variation des stocks. L'hypothèse implicite est ici que l'augmentation non désirée des stocks se transmet, en partie du moins, dans une augmentation de la quantité offerte. On s'attend à ce que $0 \leq \lambda_4 \leq 1$, puisqu'une diminution non désirée des stocks ferait aussi bien baisser la quantité offerte.

Le niveau d'inventaire désiré est spécifié dans l'équation (3.5). Il suit un modèle d'accélérateur et est fonction de la quantité offerte abstraction faite de tout ajustement des stocks. On pourrait ajouter des variables explicatives ou en changer sans modifier sensiblement les caractéristiques du modèle. On pourrait par exemple vouloir tenir compte des différents coûts associés à l'inventaire : son maintien, son ajustement.

L'équation d'ajustement des prix est très importante dans les modèles de marché puisqu'elle spécifie les hypothèses sur les caractères du marché étudié. Elle est déterminante aussi pour les méthodes d'estimation, comme on l'a vu plus haut. Ici, l'équation (3.4) pose que l'ajustement des prix ($P_t - P_{t-1}$) est proportionnel à la demande excédentaire de la période antérieure, avec en plus une composante stochastique additive. C'est dire que l'on fait d'une part l'hypothèse d'un marché compétitif - puisque seule la demande excédentaire influencerait sur le mouvement

des prix -, et, d'autre part, l'hypothèse de rigidités dans le processus d'ajustement, qui se manifestent par un délai d'ajustement.

Il faut remarquer que dans le cadre d'un modèle de déséquilibre, dire qu'un marché est compétitif n'implique que l'absence de profit anormal à long terme (ou la libre entrée dans l'industrie ou le grand nombre d'entreprises). La compétition n'implique pas ici que la courbe de demande associée à chaque entreprise est parfaitement élastique, comme c'est le cas dans un modèle d'équilibre. Le prix n'est pas celui auquel la firme peut tout vendre, puisqu'on peut se trouver en situation d'offre excédentaire. Il n'est pas non plus tel que si l'entreprise augmente son prix, les ventes tombent à zéro, puisqu'on peut se trouver en situation de demande excédentaire. La courbe de demande de la firme pourra donc avoir une pente négative, à court terme du moins, car à long terme la compétition (la libre entrée dans l'entreprise) peut agir sur les prix.

L'équation (3.4) n'est pas la plus générale que l'on puisse poser dans le cadre des autres équations du modèle. Il est possible d'ajouter d'autres variables et donc de spécifier d'autres structures de marché que celle de la concurrence parfaite. On peut ainsi avoir

$$(3.4') \quad \Delta P_t = \underline{Z}'_t \underline{\delta} + \gamma(D_{t-1} - S_{t-1}) + u_{3t}$$

où \underline{Z}_t est un vecteur de variables exogènes permettant de représenter la fixation de prix "administrés", comme dans un marché oligopolistique, par exemple.

2. Estimation du modèle

2.1. Modèle à estimer

Pour arriver à un modèle qu'on puisse estimer, il faut d'abord remplacer I_t^d dans l'équation (3.2) par l'équation (3.5). On obtient

$$(3.6) \quad S_t = \psi_0 + \frac{X_t}{-t} \underline{\psi}_1 + \frac{X_t}{-st} \underline{\psi}_2 + \psi_3 P_t + \psi_4 I_{t-1} + u_{5t}$$

$$\text{où } \psi_0 = \lambda_0 + \lambda_4(\theta_0 + \theta_1 \lambda_0)$$

$$\underline{\psi}_1 = \underline{\lambda}_1(1 - \lambda_4 \theta_1)$$

$$\underline{\psi}_2 = \underline{\lambda}_2(1 - \lambda_4 \theta_1)$$

$$\psi_3 = \lambda_3(1 - \lambda_4 \theta_1)$$

$$\psi_4 = \lambda_4$$

$$u_{5t} = u_{2t}(1 - \lambda_4 \theta_1) - \lambda_4 u_{4t}$$

On remarque ici qu'on ne peut à partir des ψ_i ($i = 0, \dots, 4$) retrouver l'ensemble des λ_i et des θ_i . Il faut connaître a priori un des coefficients λ_i ou θ_i ($i \geq 1$) pour identifier les autres paramètres. θ_0 et λ_0 sont bien sûr non identifiables. C'est ainsi, par exemple, que connaissant θ_1 et $\lambda_4 = \lambda_4$, on pourrait poser

$$\underline{\lambda}_3 = \underline{\psi}_3 / (1 - \theta_1 \lambda_4)$$

$$\underline{\lambda}_2 = \underline{\psi}_2 / (1 - \theta_1 \lambda_4)$$

$$\underline{\lambda}_1 = \underline{\psi}_1 / (1 - \theta_1 \lambda_4)$$

Si, par contre, on connaît λ_3 et $\lambda_4 = \psi_4$, on obtient alors

$$\theta_1 = (1 - \psi_3/\lambda_3)/\lambda_4$$

$$\lambda_2 = \psi_2/(1 - \theta_1 \lambda_4)$$

$$\lambda_1 = \psi_1/(1 - \theta_1 \lambda_4)$$

Le modèle à estimer va comprendre les équations (3.1), (3.6), (3.3) et (3.4). Cette dernière équation peut être réécrite à la façon de Bowden afin de pouvoir tester l'hypothèse du déséquilibre. Elle deviendrait

$$(3.4') \quad P_t = \mu P_{t-1} + (1-\mu) P_{t-1}^* + \mu u_{6t}$$

$$(3.7) \quad \text{où } \mu = 1 + \gamma(\beta_3 - \psi_3)$$

Quand μ tend vers zéro, le modèle tend vers un modèle d'équilibre dynamique. Après avoir estimé les paramètres γ , β_3 et ψ_3 , il est possible d'estimer μ et de tester s'il est significativement différent de zéro.

On peut trouver la variance de μ en appliquant la formule de Goldberger, Nagar, Odeh (1961). Essentiellement, elle pose que si $y = f(\underline{X})$ et que $\text{Cov}(\underline{X}) = S$, alors $\text{Cov}(\underline{y}) = DSD'$ où $D = (\partial f/\partial \underline{X}')_{\underline{X}=\underline{m}}$ et $\lim E(\underline{X}) = \underline{m}$. Ici, on obtient la matrice de covariance asymptotique de μ :

$$\begin{bmatrix} \frac{\partial \mu}{\partial \hat{\gamma}} \\ \frac{\partial \mu}{\partial \hat{\beta}_3} \\ \frac{\partial \mu}{\partial \hat{\psi}_3} \end{bmatrix}' S \begin{bmatrix} \frac{\partial \mu}{\partial \hat{\gamma}} \\ \frac{\partial \mu}{\partial \hat{\beta}_3} \\ \frac{\partial \mu}{\partial \hat{\psi}_3} \end{bmatrix} = \begin{bmatrix} (\hat{\beta}_3 - \hat{\beta}_3) \\ \hat{\gamma} \\ -\hat{\gamma} \end{bmatrix}' S \begin{bmatrix} (\hat{\beta}_3 - \hat{\psi}_3) \\ \hat{\gamma} \\ -\hat{\gamma} \end{bmatrix}$$

et S est la matrice de covariance de $\hat{\gamma}$, $\hat{\beta}_3$ et $\hat{\psi}_3$ obtenue suite à leur estimation.

Pour estimer les paramètres structuraux du modèle, il faut encore faire disparaître la variable inobservable D_{t-1} qui se trouve dans l'équation (3.1). En isolant D_{t-1} dans l'équation (3.4) on trouve

$$(3.9) \quad D_{t-1} = 1/\gamma P_t - 1/\gamma P_{t-1} + S_{t-1} - 1/\gamma u_{3t}$$

En substituant maintenant à S_{t-1} , dans l'équation (3.9) la valeur qu'elle prend dans l'équation (3.6) et en insérant le tout dans l'équation (3.1), on arrive finalement à l'équation suivante :

$$(3.10) \quad D_t = (\beta_0 + \beta_4 \psi_0) + \underline{X}'_t \underline{\beta}_1 + \underline{X}'_{dt} \underline{\beta}_2 + \underline{X}'_{t-1} \underline{\psi}_1 \beta_4 + \underline{X}'_{st-1} \underline{\psi}_3 \beta_4 \\ + (\beta_4/\gamma + \beta_3)P_t - \beta_4(1/\gamma - \psi_3)P_{t-1} + \psi_4 \beta_4 I_{t-2} - \beta_4 Q_{t-1} \\ + u_{1t} - (\beta_4/\gamma)u_{3t} + \beta_4 u_{5t-1}$$

Le système d'équations à estimer est donc maintenant (3.10), (3.6), (3.3) et (3.4).

Supposons que les termes d'erreur u_{1t} , u_{5t} et u_{3t} suivent une distribution multinormale et qu'ils ne sont corrélés que de façon contemporaine. Autrement dit :

$$(3.11) \quad (u_{1t}, u_{5t}, u_{3t}) \sim N_3(\underline{0}, \Sigma)$$

$$\text{Cov}(u_{it}, u_{jt}) \neq 0 \quad \forall i, j \quad i, j = 1, 3, 5$$

$$\text{Cov}(u_{it}, u_{jt-s}) = 0 \quad s \neq 0 \quad \forall i, j \quad i, j = 1, 3, 5$$

La méthode du maximum de vraisemblance à information complète (FIML) s'avère ici impraticable. Il y a de l'autocorrélation temporelle entre le terme d'erreur de l'équation (3.10) et celui de l'équation (3.6). Comme nous l'avons vu précédemment, une telle autocorrélation a pour conséquence que la fonction de vraisemblance présente des intégrales à n dimensions, où n est le nombre d'observations.

2.2. Modèle en forme réduite

Il faut donc se rabattre sur une méthode qui ne tiendra pas compte de toute l'information. Nous laisserons de côté la condition du minimum (3.3). La solution du système que nous avons peut alors être trouvée en passant par l'équation de P_t en forme réduite. On insère les équations (3.6) et (3.10), retardées d'une période, dans l'équation (3.4) :

$$\begin{aligned}
 (3.12) \quad P_t &= \xi_0 + \underline{X}'_{t-1} \underline{\xi}_1 + \underline{X}'_{t-2} \underline{\xi}_2 + \underline{X}'_{dt-1} \underline{\xi}_3 + \underline{X}'_{st-1} \underline{\xi}_4 + \underline{X}'_{st-2} \underline{\xi}_5 \\
 &+ \xi_6 P_{t-1} + \xi_7 P_{t-2} + \xi_8 I_{t-2} + \xi_9 I_{t-3} + \xi_{10} Q_{t-1} + \eta_t \\
 &= \underline{Y}_t \underline{\xi} + \eta_t
 \end{aligned}$$

où

$$\begin{aligned}
 \xi_0 &= \gamma(\beta_0 + \beta_4 \psi_0 - \psi_0) \\
 \underline{\xi}_1 &= \gamma(\underline{\beta}_1 - \underline{\psi}_1) \\
 \underline{\xi}_2 &= \gamma(\beta_4 \underline{\psi}_1) \\
 \underline{\xi}_3 &= \gamma \underline{\beta}_2 \\
 \underline{\xi}_4 &= -\gamma \underline{\psi}_2 \\
 \underline{\xi}_5 &= \gamma(\beta_4 \underline{\psi}_2) \\
 \xi_6 &= 1 + \beta_4 + \gamma(\beta_3 - \psi_3) \\
 \xi_7 &= \gamma(\psi_3 \beta_4) - \beta_4 \\
 \xi_8 &= -\gamma \beta_4 \\
 \xi_9 &= \gamma \beta_4 \psi_4 \\
 \xi_{10} &= -\gamma \beta_4 \\
 \eta_t &= -\gamma u_{5t-1} - \beta_4 u_{3t-1} + \gamma u_{1t-1} + \gamma \beta_4 u_{5t-2} + u_{3t} \\
 \underline{Y}_t &= (1, \underline{X}'_{t-1}, \underline{X}'_{t-2}, \underline{X}'_{dt-1}, \underline{X}'_{st-1}, \underline{X}'_{st-2}, P_{t-1}, P_{t-2}, I_{t-2}, I_{t-3}, Q_{t-2}) \\
 \underline{\xi} &= (\xi_0, \underline{\xi}_1, \underline{\xi}_2, \underline{\xi}_3, \underline{\xi}_4, \underline{\xi}_5, \xi_6, \xi_7, \xi_8, \xi_9, \xi_{10})'
 \end{aligned}$$

- Identification

Si l'on peut estimer de façon convergente le vecteur $\underline{\xi}$, il est possible d'identifier ensuite l'ensemble des paramètres structurels (à l'exception des deux constantes β_0 et ψ_0). On trouve (par ordre) :

$$\beta_4 = -\xi_9/\xi_8$$

$$\gamma = -\xi_{10}/\beta_4$$

$$\psi_4 = -\xi_8/\gamma$$

$$\psi_3 = (\xi_7/\beta_4 + 1)/\gamma$$

$$\beta_3 = (\xi_6 - 1 - \beta_4)/\gamma + \psi_3$$

$$\beta_2 = \xi_3/\gamma$$

$$\psi_1 = \xi_2/(\gamma \beta_4)$$

$$\beta_1 = \xi_1/\gamma + \psi_1$$

Il nous reste alors deux ξ_i , ξ_4 et ξ_5 , pour identifier le dernier paramètre ψ_2 :

$$\psi_2 = \begin{cases} -\xi_4/\gamma \\ \xi_5/(\gamma \beta_4) \end{cases}$$

Le système est donc suridentifié et il est possible de retrouver tous les paramètres.

Il est intéressant de remarquer que le système demeure identifiable sous diverses variations des équations. Ainsi l'introduction de

façon additive d'un vecteur de variables exogènes distinctes dans l'équation d'ajustement des prix (3.4') ne pose pas de problèmes d'identification. Des ajouts à l'équation des stocks désirés peuvent aussi être faits sans difficulté. Il est possible enfin que l'ajustement des prix se fasse de façon contemporaine (voir l'appendice A).

Cependant, il est important que le report de demande excédentaire d'une période à l'autre ne soit pas nul, (c'est-à-dire $\beta_4 = 0$) car alors on ne peut plus rien identifier, six des onze coefficients ξ_i s'annulant alors.

- Structure du terme d'erreur

Sous forme matricielle, l'équation (3.12) devient

$$(3.13) \quad \begin{array}{ccccccc} \underline{P} & = & Y & \underline{\xi} & + & \underline{\eta} & \\ (N-3 \times 1) & & (N-3 \times 11) & (11 \times 1) & & (N-3 \times 1) & \end{array}$$

On remarque que $\{\eta_t\}$ correspond à un processus gaussien de moyenne nulle et de variance σ_η^2 qui suivrait une moyenne mobile du deuxième ordre, MA(2). En effet, comme η_t est une somme de normales, il est normal; comme par ailleurs une variable aléatoire qui suit une distribution normale est complètement caractérisée par ses moments du premier ordre et du second ordre, on peut affirmer la correspondance avec une MA(2). Ce processus est nécessairement stationnaire si la moyenne et la matrice d'autocovariance existent.

La matrice de covariance de η peut s'écrire de la façon suivante :

$$(3.14) \quad E(\underline{\eta}\underline{\eta}') = \phi_0 R, \quad \text{où} \quad \phi_0 = \sigma_\eta^2$$

et

$$R = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & 0 \\ \rho_1 & 1 & \rho_1 & \dots & \dots \\ \rho_2 & \rho_1 & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots \end{bmatrix}_{(n-3 \times n-3)}$$

On pourrait l'exprimer aussi par

$$R = [r_{ij}]$$

$$\text{où } r_{ij} = \begin{cases} 1 & \text{si } i = j \\ \rho_1 & \text{si } |i - j| = 1 \\ \rho_2 & \text{si } |i - j| = 2 \\ 0 & \text{si } |i - j| > 2 \end{cases}$$

R est la matrice d'autocorrélation du processus et ρ_1 et ρ_2 en sont les coefficients. On peut les estimer, connaissant les valeurs calculées des éléments de η , par

$$(3.15) \quad \hat{\rho}_1 = \frac{\text{Cov}(\hat{\eta}_t, \hat{\eta}_{t-1})}{V(\hat{\eta}_t)}, \quad \hat{\rho}_2 = \frac{\text{Cov}(\hat{\eta}_t, \hat{\eta}_{t-2})}{V(\hat{\eta}_t)}$$

- Fonction de vraisemblance

Il nous faut maintenant exprimer la fonction de vraisemblance associée à notre modèle.

Le jacobien de la transportation de η à P est unitaire. La fonction de vraisemblance s'écrit donc :

$$(3.16) \quad \ell = (2\pi)^{-(N-3)/2} |\phi_0 R|^{-1/2} \exp \left\{ - \frac{[\underline{P} - Y \underline{\xi}]' R^{-1} [\underline{P} - Y \underline{\xi}]}{2\phi_0} \right\}$$

Quand on en prend le logarithme on obtient :

$$(3.17) \quad \ln \ell = L = - \frac{(N-3)}{2} \ln(2\pi) - \frac{(N-3)}{2} \ln \phi_0 - \frac{1}{2} \ln |R| \\ - \frac{1}{2} \left\{ \frac{[\underline{P} - Y \underline{\xi}]' R^{-1} [\underline{P} - Y \underline{\xi}]}{\phi_0} \right\}$$

On concentre ensuite la fonction L sur ϕ_0 , c'est-à-dire qu'on remplace ϕ_0 par sa valeur $\hat{\phi}_0$ qui satisferait la condition nécessaire de premier ordre d'un maximum par rapport à ϕ_0 , $\partial L / \partial \phi_0 = 0$.

$$\frac{\partial L}{\partial \phi_0} = - \frac{(N-3)}{2} \frac{1}{\phi_0} + \frac{[\underline{P} - Y \underline{\xi}]' R^{-1} [\underline{P} - Y \underline{\xi}]}{2} = 0$$

Donc

$$(3.18) \quad \hat{\phi}_0 = \frac{[\underline{P} - Y \underline{\xi}]' R^{-1} [\underline{P} - Y \underline{\xi}]}{N-3}$$

Remplaçant ϕ_0 par $\hat{\phi}_0$ dans L, nous obtenons la fonction concentrée L_c :

$$(3.19) \quad L_c = - \frac{(N-3)}{2} \ln(2\pi) - \frac{(N-3)}{2} \{ \ln[\underline{P} - Y \underline{\xi}]' R^{-1} [\underline{P} - Y \underline{\xi}] - \ln(N-3) \} \\ - \frac{1}{2} \ln |R| - \frac{(N-3)}{2}$$

En regroupant, on a :

$$(3.20) \quad L_c = C - \frac{1}{2} \ln |R| - \frac{(N-3)}{2} \ln \{ [\underline{P} - Y \underline{\xi}]' R^{-1} [\underline{P} - Y \underline{\xi}] \}$$

où C = constante = $-\frac{(N-3)}{2} [\ln(2\pi) + \ln(N-3) + 1]$.

Il importe de ne pas oublier les contraintes que nous impose la relation fonctionnelle non linéaire des éléments du vecteur $\underline{\xi}$ au vecteur \underline{K} des paramètres structurels :

$$\underline{\xi}(\underline{K}) \quad \text{et} \quad \underline{K} = (\xi_0, \beta_1', \beta_2', \beta_3, \beta_4, \psi_1', \psi_2', \psi_3, \psi_4, \gamma)'$$

2.3. Maximisation de la fonction de vraisemblance

- Maximisation non linéaire

Cette fonction de vraisemblance doit être maximisée par rapport à ses paramètres, le vecteur \underline{K} et les deux coefficients d'autocorrélation, ρ_1 et ρ_2 . La fonction à maximiser est non linéaire, on l'a vu. Dans un tel cas, l'ensemble des propriétés des estimateurs du maximum de vraisemblance demeurent valables. Les estimations sont convergentes et asymptotiquement normales, sous des conditions assez générales qui ne comprennent

pas la linéarité de la fonction. Comme nous ne tenons pas compte de la condition (3.3), elles ne sont pas asymptotiquement efficaces.

La méthode de maximisation de la fonction de vraisemblance proposée par Dagenais (1980) est celle de Fletcher-Powell (1963). Appelée aussi méthode de Davidon-Fletcher-Powell, c'est une méthode à métrique variable qui ne nécessite pas le calcul du hessien.

Pour cette méthode itérative, on peut choisir comme valeurs initiales les estimations dérivées d'un MCO sur l'équation (3.12), et, pour ψ_2 qui est surdéterminé, on peut prendre la moyenne des deux estimations obtenues. Quant à ρ_1 et ρ_2 , ils peuvent être estimés en calculant les coefficients de corrélation, selon les formules (3.15), à l'aide des résidus du MCO.

A la suite de la convergence de l'algorithme de Fletcher-Powell, on peut approximer la matrice de covariance de $(\underline{K}', \rho_1, \rho_2)'$ estimés, en utilisant la formule de Berndt, Hall et alii (1974) :

$$(3.21) \quad \sum_{t=4}^N \left(\frac{\partial \{-1/2 \ln |R| - (N-3)/2 \ln[(\underline{P} - Y \underline{\xi})' R^{-1} (\underline{P} - Y \underline{\xi})]\}_t}{\partial (\underline{K}', \rho_1, \rho_2)'} \right) \times (\text{id.})'$$

$$\underline{K}', \rho_1, \rho_2 = \hat{\underline{K}}, \hat{\rho}_1, \hat{\rho}_2$$

Il serait ensuite facile de calculer μ par la formule que nous avons déjà vue (3.7) et aussi de lui déterminer un intervalle de confiance par la méthode de Goldberger, Nagar et Odeh (3.8).

Cette méthode de calcul est appropriée à notre fonction de vraisemblance quand sa forme dépend de l'hypothèse (3.11) sur les termes d'erreur des équations (3.10), (3.6) et (3.4). Les résidus de l'estimation devraient alors suivre un processus de moyenne mobile du deuxième ordre. C'est là une conséquence vérifiable.

Une autre façon de procéder - que nous avons utilisée pour nos estimations, plus loin -, consiste à ne pas faire d'hypothèse initiale sur la forme des erreurs des équations structurelles (3.10), (3.6) et (3.4), et à supposer plutôt que le terme d'erreur de l'équation réduite en P_t - l'équation (3.12) - suit une loi normale de moyenne nulle et de variance σ^2 . La fonction de vraisemblance concentrée L_c deviendrait alors

$$(3.22) \quad L_c = C - \frac{(N-3)}{2} \ln \sigma^2 - \frac{(N-3)}{2} \ln \{ [\underline{P} - Y \underline{\xi}]' [\underline{P} - Y \underline{\xi}] \}$$

Nous ferons plutôt une hypothèse moins forte et choisirons de considérer η_t comme un bruit blanc sans spécifier sa loi de distribution. On abandonne alors la possibilité d'estimer par la méthode du maximum de vraisemblance. On peut utiliser par contre plusieurs autres méthodes d'estimation non linéaire. Nous avons choisi de retenir la méthode des moindres carrés non linéaires. Les algorithmes de minimisation utilisés sont ceux de la méthode de Gauss-Newton, de même que ceux d'une méthode

que nous appelons moindres carrés linéaires itérés et qui tire profit de la forme multilinéaire de la fonction à minimiser.

La méthode de Gauss-Newton n'a pas à être expliquée longuement, elle est bien connue. Elle consiste essentiellement à faire un développement linéaire de la fonction à optimiser en un point. Les moindres carrés sont ensuite appliqués à ce développement linéaire. La valeur obtenue à chaque estimation est utilisée pour le développement suivant, jusqu'à convergence.

- Moindres carrés itérés

La méthode des moindres carrés itérés consiste à pratiquer une séquence de moindres carrés sur des sous-ensembles des paramètres à estimer. A chaque étape, on fixe la valeur des paramètres exclus de la régression présente, à celle qu'on a obtenue à l'itération précédente. La série d'itérations se termine quand le critère de convergence est satisfait.

La convergence de la méthode repose sur un théorème énoncé par Sargan (1964) et repris de façon plus formelle par Oberhofer et Kmenta (1974). C'est une extension des méthodes que Maddala (1977, p. 279) appelle "two-step procedures" et dont on connaît bien celle de Cochrane-Orcutt et celle de Durbin pour les modèles où on pense que les termes d'erreur suivent un processus autorégressif du premier ordre.

Oberhofer et Kmenta prouvent d'abord le lemme suivant :

- soit la fonction $f(\theta)$ qu'on veut maximiser par rapport à $\underline{\theta}$; $f(\theta)$ est continue sur un compact non vide S ;
- soit la partition de $\underline{\theta}$ en deux sous-ensembles : $\underline{\theta}_1$ et $\underline{\theta}_2$, $\underline{\theta} = (\underline{\theta}_1, \underline{\theta}_2)$, $\underline{\theta}_1 \cap \underline{\theta}_2 = \emptyset$;
- et soit l'itération suivante :
 - a) de $\underline{\theta}_1^0$, vecteur de valeurs initiales de $\underline{\theta}_1$, on trouve $\underline{\theta}_2^0$ qui maximise $f(\underline{\theta})$, étant donné $\underline{\theta}_1^0$;
 - b) de $\underline{\theta}_2^0$, on trouve $\underline{\theta}_1^1$ qui maximise $f(\underline{\theta})$, étant donné $\underline{\theta}_2^0$;
 - c) de $(\underline{\theta}_1^j, \underline{\theta}_2^j)$, pour tout $0 \leq j$, on maximise $f(\underline{\theta}_1^j, \underline{\theta}_2^j)$ par rapport à $\underline{\theta}_2^j$. Le maximum sera atteint à $\underline{\theta}_2 = \underline{\theta}_2^{K+1}$. Puis on maximise $f(\underline{\theta}_1^j, \underline{\theta}_2^{j+1})$ par rapport à $\underline{\theta}_1^j$. On aura un maximum à $\underline{\theta}_1 = \underline{\theta}_1^{j+1}$.

Le lemme énonce alors que, pour $0 \leq j \leq K$,

- 1) la séquence $\{\underline{\theta}^k\}$ a au moins un point d'accumulation $\underline{\theta}^*$ sur l'ensemble S ;
- 2) si $\underline{\theta}^*$ et $\underline{\theta}^+$ sont deux points d'accumulation de la séquence, alors $f(\underline{\theta}^+) = f(\underline{\theta}^*)$;
- 3) pour tout point d'accumulation $\underline{\theta}^* = (\underline{\theta}_1^*, \underline{\theta}_2^*)$,

$$\max_{\underline{\theta}_1} f(\underline{\theta}_1, \underline{\theta}_2^*) = \max_{\underline{\theta}_2} f(\underline{\theta}_1^*, \underline{\theta}_2) = f(\underline{\theta}_1^*, \underline{\theta}_2^*) = f(\underline{\theta}^*).$$

On trouvera une preuve de ce lemme dans les articles déjà mentionnés.

Remarquons qu'il est facile de voir que l'algorithme d'itération pourrait comporter de nombreuses étapes supplémentaires. Le vecteur des variables pourrait se décomposer en

$$\underline{\theta} = (\underline{\theta}_1, \underline{\theta}_2) = (\underline{\theta}_3, \underline{\theta}_4) = \dots = (\underline{\theta}_{n-1}, \underline{\theta}_n)$$

où les $(\underline{\theta}_i, \underline{\theta}_{i+1})$ représentent des partitions différentes de l'ensemble $\underline{\theta}$. A la limite, on pourrait imaginer qu'une seule variable est maximisée à chaque étape de l'itération et donc qu'on doit se donner une valeur initiale pour toutes les variables de $\underline{\theta}$, moins une. Les valeurs fixées à chaque itération étant choisies parmi celles qui maximisaient la fonction à l'itération précédente, la séquence des itérations ne peut faire diminuer la valeur de la fonction. Au pire, elle peut demeurer stationnaire si un maximum a été atteint. Comme pour la plupart des autres algorithmes de maximisation, le maximum atteint peut être local ou global. Seule une preuve de la concavité de la fonction peut nous assurer d'un maximum global. Sinon, il faut tâtonner en changeant les valeurs initiales.

L'intérêt de ce lemme est d'abord pratique. Il permet de calculer plus facilement le maximum d'une fonction compliquée, quand on peut la simplifier en ne considérant qu'une partie à la fois de ses variables. Le prix à payer est par contre une augmentation du nombre d'étapes dans chaque itération. Il faut donc considérer, entre autres, le temps relatif des itérations, la précision des algorithmes qu'on peut utiliser, le temps nécessaire pour les mettre en place, le groupement optimal des variables.

Dans le cas qui nous occupe, on constate que l'équation (3.12) a une forme multilinéaire, plus précisément trilinéaire; la non-linéarité

de la fonction provient de simples multiplications de paramètres, qui sont groupés par trois, au plus. Il serait donc possible de ramener notre équation à trois équations linéaires faciles à calculer à l'aide des moindres carrés ordinaires.

L'équation (3.12) serait ainsi estimée par les étapes suivantes :

- 1) on chercherait par un MCO une valeur initiale pour β_4 et γ ;
- 2) tenant β_4 et γ fixés, on maximiserait par rapport à $\underline{\beta}_1, \underline{\psi}_1, \underline{\beta}_2, \underline{\psi}_2, \beta_3, \psi_3$ et ψ_4 ;
- 3) tenant $\underline{\beta}_1, \underline{\beta}_2, \beta_3, \beta_4, \underline{\psi}_1, \underline{\psi}_2, \psi_3$ et ψ_4 fixés, on maximise par rapport au seul γ ;
- 4) tenant enfin $\gamma, \underline{\psi}_1, \underline{\psi}_2, \psi_3$ et ψ_4 fixés, on peut maximiser par rapport à $\underline{\beta}_1, \underline{\beta}_2, \beta_3$ et β_4 ;
- 5) on retourne à l'étape 2).

Quand le critère de convergence numérique est satisfait, il est possible de vérifier l'hypothèse de bruit blanc imposée aux termes d'erreur en utilisant le test de Ljung-Box. S'il s'avérait qu'ils suivent plutôt un processus stochastique plus complexe, on pourrait identifier ce processus, si possible, et reprendre l'estimation sur cette base.

2.4. Estimation des constantes

Après avoir estimé les éléments de \underline{K} , on peut trouver des estimés pour β_0 et ψ_0 . De l'équation (3.12) on sait que

$$\xi_0 = \gamma[\beta_0 - (1 - \beta_4) \psi_0]$$

et donc que

$$(3.23) \quad (\beta_0 + \beta_4 \psi_0) = (\xi_0/\gamma + \psi_0)$$

On écrit ensuite l'équation Q_t (selon que $Q_t = D_t$ et que $Q_t = S_t$) de telle sorte qu'on isole à droite ψ_0 et le terme d'erreur. Si $Q_t = D_t$, utilisant (3.10) et (3.23), on obtient :

$$(3.24) \quad Q_t - \xi_0/\gamma - \frac{X'_t}{-t} \beta_1 - \frac{X'_{dt}}{-dt} \beta_2 - \frac{X'_{t-1}}{-t-1} \psi_1 \beta_4 - \frac{X'_{st-1}}{-st-1} \psi_2 \beta_4 - (\beta_4/\gamma + \beta_3) P_t \\ + \beta_4(1/\gamma - \psi_3) P_{t-1} - \psi_4 \beta_4 I_{t-2} + \beta_4 Q_{t-1} = Q_{1t} = \psi_0 + u_{7t}$$

$$\text{où } u_{7t} = -(\beta_4/\gamma) u_{3t} + u_{1t} + \beta_4 u_{5t-1} .$$

Si $Q_t = S_t$, en réarrangeant (3.6) on trouve :

$$(3.25) \quad Q_t - \frac{X'_t}{-t} \psi_1 - \frac{X'_{st}}{-st} \psi_2 - \psi_3 P_t - \psi_4 I_{t-1} = Q_{2t} = \psi_0 + u_{5t}$$

Comme on connaît Q_{1t} et Q_{2t} , on peut estimer ψ_0 par moindres carrés en trouvant $\hat{\psi}_0$ qui minimise

$$(3.26) \quad \phi = \sum_{t=3}^N [Q_{1t} Z_t + Q_{2t}(1 - Z_t) - \psi_0]^2$$

$$\begin{aligned} \text{où } Z_t &= 1 & \text{si } Q_t &= D_t ; \\ &= 0 & \text{si } Q_t &= S_t . \end{aligned}$$

Une solution naïve de ce problème demanderait 2^{N-2} calculs distincts. Dagenais et Dagenais (1982) ont cependant proposé un ingénieux algorithme qui permet de trouver la solution en un nombre de calculs distincts, de l'ordre de $N-1$.

CHAPITRE IV

Application au marché du bois de sciage américain

1. Caractéristiques générales du marché

Nous avons choisi d'examiner le marché du bois de sciage aux Etats-Unis à l'aide du modèle dynamique que nous avons développé dans les sections précédentes.

Ce marché, aux Etats-Unis comme au Canada d'ailleurs, peut être vu comme un marché concurrentiel. Il comprend de très nombreux participants et il est relativement facile d'entrer dans l'industrie. En 1967, il y avait 10 271 scieries possédées par 10 016 entreprises aux Etats-Unis. On peut donc penser que l'hypothèse d'un processus d'ajustement des prix qui dépend essentiellement de la demande ou de l'offre excédentaire peut être pertinent.

C'est ensuite un marché dont il est raisonnable de penser qu'il n'est pas souvent en équilibre. Une demande excédentaire peut se manifester qui entraînerait après quelque temps la mise en marché de nouvelles scieries. Cette demande est par ailleurs relativement ferme, elle peut difficilement se tourner vers des biens substitués et il est peu probable qu'à court terme on modifie ses projets et qu'on passe d'une charpente de bois à une charpente en blocs de béton. La demande qui n'est pas satisfaite en une période se reportera alors, en partie du moins, sur la demande de la période suivante. Cela pourrait signifier par exemple un délai

plus long entre la demande de permis (ou la mise en chantier) et le parachèvement des constructions domiciliaires.

On peut aussi concevoir qu'une offre excédentaire se manifeste. Elle tiendrait par exemple à une baisse marquée des mises en chantier d'habitations. Il y aurait alors une augmentation involontaire des stocks qui absorberaient le surplus d'offre. Eventuellement, des fermetures de scieries surviendraient ensuite afin de ramener l'offre dans l'industrie à un niveau compatible avec celui de la demande.

Nous pouvons donc penser a priori que l'utilistion de notre modèle de déséquilibre dynamique serait pertinent et que les paramètres des deux équations, celle de la demande et celle de l'offre, pourraient être estimés.

2. Spécifications

2.1. Existantes

A) L'étude de McKillop (1967) examine de façon approfondie la spécification d'un système d'équations simultanées pour l'offre et la demande de produits forestiers aux Etats-Unis.

On y relève que la consommation finale de bois de sciage, en 1952, se répartissait en 72% pour la construction, 16% pour l'emballage et la livraison et 12% pour la transformation manufacturière. A elle seule, la construction résidentielle représentait 36% de la consommation totale de produits forestiers.

Nous n'avons pas trouvé d'étude plus récente sur les utilisations finales du bois de sciage aux Etats-Unis. Il est cependant raisonnable de penser, croyons-nous, que la part de la construction résidentielle n'a pas diminuée. La construction non résidentielle par contre nous semble utiliser de façon beaucoup plus intensive l'acier et le béton. Les codes de sécurité limitent grandement l'emploi du bois dans ces constructions.

Comme, de plus, les chiffres cités par McKillop concernent la totalité du bois de sciage, qu'il soit dur ou mou, et que l'on sait que l'utilisation commerciale du bois mou est très faible en dehors de l'industrie de la construction, on peut considérer la construction résidentielle comme l'utilisation prépondérante de ce matériau.

McKillop pose que la quantité de bois demandée sera fonction, de son prix, bien sûr, des prix des matériaux substitués, des taux de salaire dans l'industrie de la construction, de la valeur totale de la construction, de la population et du revenu disponible par habitant.

Les variables de prix et de revenu s'expliquent simplement par la théorie économique. Les matériaux substitués peuvent être l'acier de structure, le béton, la brique et le contre-plaqué. Il nous semble cependant que les deux premiers ont plus de pertinence pour une étude où les données se répartissent sur une longue période et où l'on peut croire que des changements dans la façon de construire sont survenus. Le prix

de ces matériaux ne nous apparaît pas pouvoir être une variable de choix dans le court terme à cause des changements importants au niveau des plans et des ouvriers qu'ils impliquent. (McKillop construit une moyenne mobile des prix sur 10 ans pour tenir compte probablement d'une remarque semblable à la nôtre). De plus, ils n'auraient de pertinence que pour la construction non résidentielle, dont McKillop veut aussi rendre compte. L'inclusion de la brique enfin ne se justifie que si la période inclut l'époque où une proportion significative des maisons avaient un revêtement extérieur en bois.

Le contre-plaqué est le matériau substitut le plus probable. McKillop remarque cependant, avec raison, que ce n'est pas seulement son prix qui peut le faire préférer au bois de sciage, mais son plus faible coût installé. C'est pourquoi il veut inclure parmi les variables explicatives le taux de salaire des différents corps de métier, maçons et briqueteurs, menuisiers et charpentiers, par exemple. Il ne pourrait cependant trouver des taux différents pour la pose du bois ou du contre-plaqué, puisque ce sont les mêmes ouvriers qui installent les deux. En fait, on ne trouve qu'un taux agrégé, pour l'ensemble des métiers de la construction. C'est ce taux de salaire qu'il inclut dans la spécification de la fonction de demande pour tenir compte du facteur d'installation. Une pondération estimée tenant compte de la différence de temps nécessaire pour accomplir un travail équivalent (couvrir une superficie de mur, par exemple), nous semblerait permettre de mieux tenir compte des conditions d'une substitution.

La valeur de la construction totale, mesurée en dollars courants - comme le fait McKillop -, ne nous semble pas être un déterminant des quantités demandées. Ce serait plutôt à titre de 'proxy' pour le revenu disponible qu'elle serait pertinente. McKillop remarque d'ailleurs leur forte collinéarité. Il aurait été préférable d'utiliser la valeur de la construction en dollars constants, d'utiliser à la fois le revenu disponible et le nombre de mises en chantier résidentielles.

La fonction d'offre de bois de sciage dépendra du prix du bois, du coût des facteurs de production autres que l'énergie et le travail (essentiellement, les billes), du prix de l'énergie - électricité ou pétrole -, du salaire courant dans les scieries et d'un indice de productivité.

En pratique, les coûts considérés sont ceux des billes, qui correspondraient, en 1958, à plus de la moitié des coûts de production, et les salaires qui compteraient pour 77% des autres coûts variables. Sur une longue période, il faut tenir compte de l'intensité capitaliste dans l'industrie, et du coût du capital. Pour cela, McKillop utilise la production annuelle par ouvrier dans les scieries. Il fait comme si cette variable était exogène et n'avait pas de relation avec la production. Du même coup, on peut penser tenir compte des changements technologiques. L'auteur ne tient pas compte cependant des investissements, à cause du manque de données, nous dit-il.

Comme l'importation de bois de sciage canadien est croissante, (11% en 1960), le tarif douanier américain de même que le taux de change

du dollar canadien en dollar américain sont ajoutés aux variables explicatives. Le niveau des stocks n'est pas considéré, car McKillop estime que son rôle n'est significatif que lorsqu'on examine des données mensuelles ou trimestrielles. Il utilise des données annuelles pour la période de 1929 à 1960.

B) McCallum (1974) propose des modèles opérationnels d'ajustement des prix dans un marché compétitif. Ces modèles ont la particularité de tenir compte des inventaires. Ce sont des modèles d'équilibre où les inventaires ne viennent pas absorber les chocs des variations de la demande, mais sont gardés pour faire des gains de capital. Les variables qu'il retient pour déterminer la demande de bois mou (Southern Pine) sont le prix, le produit national brut réel et le prix du contre-plaqué. La production dépend du prix, des salaires, du prix des billes et d'un indice de capital.

C) Sen (1977) propose quant à elle un modèle de déséquilibre où les stocks servent de tampon pour accommoder la demande ou l'offre excédentaire et où le prix et la production sont des variables de contrôle. Elle utilise une spécification simple pour la demande : prix du bois, nombre d'unités résidentielles mises en chantier et prix du contre-plaqué. Pour l'offre - plus précisément la production -, elle retient la production antérieure, le prix, la demande contemporaine et le niveau antérieur des stocks. Elle utilise aussi le taux de salaire dans les scieries. Elle estime un modèle simultané à quatre variables dépendantes : la production, le prix, la demande et les livraisons.

2.2. Spécification retenue

Nous avons décidé de nous tenir près des spécifications de McCallum et de Sen, dans un premier temps, d'une part à cause de leur simplicité et d'autre part afin de pouvoir comparer nos résultats avec les leurs. Nous avons pensé qu'il était préférable d'estimer initialement notre modèle avec un nombre réaliste, mais réduit, de variables explicatives, quitte ensuite à le compléter si c'était nécessaire.

Nous avons donc retenu pour l'équation de demande, les variables de prix du bois (P_t), de prix du contre-plaqué (PLY_t) et de nombre de mises en chantier résidentielles (HS_t). S'ajoutait aussi la demande excédentaire de la période précédente.

$$(4.1) \quad D_t = \beta_0 + \beta_{21} HS_t + \beta_{22} PLY_t + \beta_3 P_t + \beta_4 (D_{t-1} - Q_{t-1}) + \epsilon_{1t}$$

Pour l'offre, nous avons retenu les salaires (W_t) comme représentant l'ensemble des coûts. Cette décision s'est avérée nécessaire car nous n'avons pu trouver de données sur les prix des billes. Nous avons aussi tenu compte des stocks de la période précédente et du prix du bois.

$$(4.2) \quad S_t = \psi_0 + \psi_2 W_t + \psi_3 P_t + \psi_4 I_{t-1} + \epsilon_{2t}$$

L'équation d'ajustement des prix est celle que nous avons déjà présentée :

$$(4.3) \quad P_t - P_{t-1} = \gamma(D_{t-1} - S_{t-1}) + \epsilon_{3t}$$

3. Equation estimée

L'équation en forme réduite de P_t est alors :

$$(4.4) \quad P_t = \xi_0 + \beta_{21} \gamma HS_{t-1} + \beta_{22} \gamma PLY_{t-1} + \psi_2 \gamma W_{t-1} + \psi_2 \gamma \beta_4 W_{t-2} \\ + [\gamma(\beta_3 - \psi_3) + \beta_4 + 1] P_{t-1} + \beta_4(\gamma\psi_3 - 1) P_{t-2} - \psi_4 \gamma I_{t-2} \\ + \psi_4 \gamma \beta_4 I_{t-3} - \beta_4 \gamma Q_{t-2} + \eta_t$$

Pour les fins de l'estimation par la méthode que nous avons appelée des moindres carrés itérés, cette équation se ramène à trois équations à estimer successivement par itération :

1) Estimation de β_{21} , β_{22} , β_3 , ψ_2 , ψ_3 , ψ_4 étant donné γ , β_4 :

$$(4.5) \quad P_t - (\beta_4 + 1) P_{t-1} + \beta_4 P_{t-2} + \beta_4 \gamma Q_{t-2} = \\ \xi_{01} + \beta_{21}(\gamma HS_{t-1}) + \beta_{22}(\gamma PLY_{t-1}) + (\beta_3 - \psi_3) (\gamma P_{t-1}) \\ + \psi_3(\gamma \beta_4 P_{t-2}) + \psi_4(\gamma \beta_4 I_{t-3} - \gamma I_{t-2}) \\ + \psi_2(\gamma \beta_4 W_{t-2} - \gamma W_{t-1}) + \eta_{1t}$$

2) Estimation de γ étant donné β_{21} , β_{22} , β_3 , β_4 , ψ_2 , ψ_3 et ψ_4 :

$$(4.6) \quad P_t - (\beta_4 + 1) P_{t-1} + \beta_4 P_{t-2} =$$

$$\xi_{02} + \gamma[\beta_{21} HS_{t-1} + \beta_{22} PLY_{t-1} - \psi_2 W_{t-1}$$

$$+ \psi_2 \beta_4 W_{t-2} + (\beta_3 - \psi_3) P_{t-1}$$

$$+ \beta_4 \psi_3 P_{t-2} - \psi_4 I_{t-2}$$

$$+ \beta_4 \psi_4 I_{t-3} - \beta_4 Q_{t-2}] + \eta_{2t}$$

3) Estimation de β_{21} , β_{22} , β_3 et β_4 étant donné ψ_2 , ψ_3 , ψ_4 et γ :

$$(4.7) \quad P_t + \psi_2 \gamma W_{t-1} - (1 - \gamma \psi_3) P_{t-1} + \psi_4 \gamma I_{t-2} =$$

$$\xi_{03} + \beta_{21}(\gamma HS_{t-1}) + \beta_{22}(\gamma PLY_{t-1}) + \beta_3(\gamma P_{t-1})$$

$$+ \beta_4[\psi_2 \gamma W_{t-2} + P_{t-1} + (\gamma \psi_3 - 1) P_{t-2}$$

$$+ \psi_4 \gamma I_{t-3} - \gamma Q_{t-2}] + \eta_{3t}$$

Les signes attendus des coefficients sont :

$$\beta_{21}, \psi_3, \gamma > 0$$

$$\beta_{22}, \beta_3, \psi_2 < 0$$

$$0 \leq \beta_4, \psi_4 \leq 1.$$

4. Résultats

Nous présentons au tableau A les résultats de nos estimations. Nous les avons effectuées sur l'ensemble des données d'abord, puis en écartant les observations de la dernière année qui avaient un comportement très différent du reste de l'ensemble.

On constate que les valeurs initiales obtenues par MCO (en une itération) sont raisonnables et correspondent aux attentes, sauf pour ψ_4 dont la valeur devrait se trouver entre 0 et 1 et pour β_{21} dont le signe devrait être positif. La première régression par la méthode de Gauss ne corrige pas la grandeur de ψ_4 . Le signe de β_{21} demeure inacceptable et on constate surtout qu'aucun des coefficients n'est statistiquement différent de zéro. Ces très grandes variances sont le résultat d'une forte multicollinéarité entre les séries d'observations. L'équation en forme réduite de P_t a en effet comme variables explicatives P_{t-1} et P_{t-2} , de même que W_{t-1} et W_{t-2} ainsi que I_{t-2} et I_{t-3} .

Les valeurs propres de la matrice des dérivées premières évaluées au point des valeurs initiales (dans le cadre de l'estimation par la méthode de Gauss), nous indiquent d'ailleurs qu'elle est "mal conditionnée" (ill-conditioned). Le rapport de la plus grande à la plus petite valeur propre est bien supérieur à 1 000 pour 1. C'est dire qu'une légère modification du vecteur des observations sur les variables explicatives peut entraîner une forte variation des solutions. C'est une caractéristique d'un système multicollinéaire¹.

¹Cette indication de 'conditionnement' vaut pour les systèmes d'équations linéaires. La méthode gaussienne linéarise les équations non linéaires comme nous l'avons vu. Voir K.E. Atkinson, (1978), pp. 457-461.

Pour tenter de contourner cette difficulté, nous avons procédé à une mise à l'échelle (scaling) des variables explicatives. Comme il n'y a pas de théorie précise sur la meilleure façon de le faire, nous sommes contenté de suivre le conseil de Jacoby, Kowalik et Pizzo (1972) et de ramener les valeurs des variables à un même ordre de grandeur. L'amélioration n'a pas été significative. De même, le choix de solutions initiales différentes n'a pas été plus heureux.

Toutes les estimations à l'aide de la méthode de Gauss ont présenté ce problème. Une seule a donné deux coefficients significativement différents de zéro; mais le très petit nombre d'itérations (3) n'inspire pas confiance.

La deuxième méthode de maximisation de la fonction de vraisemblance que nous avons utilisée, les moindres carrés ordinaires itérés (ITER dans le tableau A), a permis dans tous les cas de diminuer encore plus la somme des erreurs au carré. Les résultats sont cependant décevants car on constate que les coefficients de prix, dans les équations d'offre et de demande, β_3 et ψ_3 , sont déraisonnables et semblent suivre un trajet explosif alors que les autres convergent, quand le nombre d'itérations augmente. Ces deux coefficients étant essentiels pour le test de μ , nous n'avons pas cru utile de le faire.

Tableau A

Méthode de minimisation ¹	Itérations ²	(100) ³ Xφ	(10) B21	(10) B22	(10) B3	(-) B4	(100) P2	(10) P3	(-) P4	(10) G	SRR ⁴	Critère de conv. ⁵
(4-156) ITER	1	0,1418	-0,257	1,405	-1,52	0,7336	-41,1	6,93	1,165	0,026	1,20539	
1. Gauss t	81	0,146 (1,01)	-0,153 (0,73)	1,088 (0,94)	-12,62 (0,57)	0,3593 (1,33)	-16,139 (0,78)	-15,43 (0,55)	0,278 (0,67)	0,03995 (0,92)	1,119969	Tol. SSR (10 ⁻⁴)
2. ITER	1458	0,1208	-0,1622	1,04269	-162,025	0,0644	-12,2638	-169,9	0,1271	0,04625	1,02517	Maxit
(4-144) ITER	1	-0,2151	0,0723	0,0667	1,949	0,9685	-3,478	19,5124	1,4225	0,02068	0,438923	
3. Gauss t	64	-0,2174 (1,62)	0,02157 (0,40)	0,3335 (0,76)	-8,6267 (0,32)	0,3390 (0,74)	-5,822 (0,53)	-12,4911 (0,38)	0,04839 (0,26)	0,04546 (0,74)	0,421147	Tol. SSR (10 ⁻⁴)
4. ITER	570	0,00925	0,2493	-24,232	0,1955	-3,805	-29,728	0,00836	0,0592	0,419733	Maxit	

¹ITER désigne la méthode que nous avons appelée des moindres carrés itérés. Gauss pratique un moindre carré non linéaire. Sous les estimés, la statistique t de Student, entre parenthèses.

(4-156) : l'échantillon complet. (4-144) : échantillon amputé des 12 dernières observations.

²Le nombre d'itérations.

ITER 1 désigne la première passe qui nous donnait les valeurs initiales pour Gauss et ITER.

³Le nombre entre parenthèses est le facteur par lequel il faut multiplier l'estimé obtenu pour enlever l'effet de la mise à l'échelle (scaling).

⁴Somme des erreurs au carré (SSR), à la fin des itérations.

⁵Critère de convergence : maxit : nombre prédéterminé d'itérations.

$$\text{tol. SSR}(10^{-4}) : \left| \frac{\text{SSR}_t - \text{SSR}_{t-1}}{\text{SSR}_t} \right| < 10^{-4}$$

Nous avons terminé sur cette note nos estimations puisqu'il apparaissait qu'une nouvelle spécification ne pourrait pas nous permettre d'éviter d'avoir P_{t-1} et P_{t-2} comme variables explicatives pour P_t . Enlever empiriquement une des deux aurait été rendre notre modèle non identifiable. Nous avons plutôt décidé d'examiner quelques voies théoriques qui pourraient nous aider à surmonter les sérieux problèmes de multicollinéarité que nous y avons rencontrés.

CHAPITRE V

Régressions ridge, bayésienne et
mixte, non linéaires

Les méthodes pour traiter les problèmes de multicollinéarité sont nombreuses. Toutes reconnaissent que le problème tient avant tout aux données utilisées pour estimer le modèle. Une des solutions consiste donc à chercher de nouvelles données, ou des données supplémentaires, qui viendront ajouter de l'information à celle que renferme déjà l'échantillon, rendant possible l'estimation précise du modèle. Une autre approche consiste à modifier la spécification du modèle de telle sorte qu'il corresponde à l'information de l'échantillon : enlever les variables dont les coefficients ne peuvent être estimés avec précision, utiliser des ratios ou des premières différences, régresser sur les composantes principales. Une dernière approche enfin cherche à ajouter de l'information sur les coefficients afin de rendre l'estimation précise. Trois méthodes se distinguent ici : la régression mixte de Theil et Goldberger (1961) et Theil (1963), la régression ridge de Hoerl et Kennard (1970) et l'estimation bayésienne. Nous voulons dans la suite montrer brièvement, d'une part comment il serait possible d'effectuer une régression ridge sur un modèle non linéaire avec erreurs additives et, d'autre part, comment on peut interpréter cette régression dans les termes de l'estimation bayésienne et de la régression mixte.

1. Linéarisation du modèle

Notre modèle est donné par l'équation (3.12) (ou 4.4). C'est un modèle non linéaire dans les paramètres avec erreurs additives. On peut l'exprimer de façon très générale par

$$(5.1) \quad \underline{Y} = f(\underline{X}; \underline{\beta}) + \underline{u}$$

où $\underline{Y}(n \times 1)$ est la variable dépendante; $\underline{X}(l \times p)$ est un vecteur de variables prédéterminées; $\underline{\beta}(p \times 1)$ un vecteur de paramètres inconnus; f une fonction non linéaire; et \underline{u} un terme d'erreur de moyenne nulle et de variance $\sigma^2 I_p$ et dont les éléments sont indépendants et identiquement distribués.

Il est très utile de le linéariser. Ce faisant, on peut lui appliquer la méthode d'estimation des moindres carrés généralisés non linéaires de Dagenais (1978), précisée pour le cas où le modèle ne comporte qu'une seule équation (Dagenais, 1983), et où les erreurs suivent de plus une loi de distribution normale.

Après avoir trouvé une solution initiale $\underline{\beta}_0$ au vecteur $\underline{\beta}$, on linéarise l'équation (5.1) autour de cette solution :

$$(5.2) \quad \underline{Y}_\ell \approx Z_\ell \underline{\beta} + \underline{u}$$

$$\text{où } \underline{Y}_\ell = \underline{Y} - f(\underline{X}; \underline{\beta}_\ell) + Z_\ell \underline{\beta}_\ell$$

$$\text{et } Z_\ell = \left(\frac{\partial f(\underline{X}; \underline{\beta})}{\partial \underline{\beta}'} \right)_{\underline{\beta}=\underline{\beta}_\ell}$$

Par un MCO sur l'équation (5.2) on trouve une solution $\underline{\beta}_{\ell+1}$, à partir de laquelle on trouve une nouvelle solution initiale. On trouve le détail du processus itératif dans Dagenais (1978). A la convergence, après s'être assuré que le minimum n'est pas que local, on peut considérer l'estimé $\underline{\beta}_L$ comme la solution MCO de l'équation

$$(5.3) \quad \underline{Y}_L = Z_L \underline{\beta} + \underline{u}$$

L'estimateur est équivalent à un FIML. On peut aussi estimer σ^2 de façon convergente par

$$(5.4) \quad \hat{\sigma}^2 = S^2 = \frac{\underline{u}'_L \underline{u}_L}{n-p}$$

où $\underline{u}_L = \underline{Y} - f(\underline{X}; \underline{\beta}_L)$.

2. Régression ridge

La régression ridge a été développée par Hoerl et Kennard (1970) dans un cadre linéaire. Elle consiste essentiellement à ajouter à chaque élément de la diagonale de $X'X$ un facteur k_i , où k_i est petit et peut être différent (ou non) pour chacun des éléments de la diagonale. Suivant l'idée de Dagenais (1983), on peut cependant l'appliquer à notre modèle non linéaire, si on utilise l'équation (5.3). L'estimateur ridge généralisé est alors

$$(5.5) \quad \hat{\beta}_r = (Z'_L Z_L + K)^{-1} Z'_L \underline{Y}_L$$

où $K = \text{diag}[k_1, \dots, k_p]$.

Si $k_i = k$, $i = 1, \dots, p$, on a alors l'estimateur ridge ordinaire avec $K = k I_p$.

Règle générale, dans la littérature sur le ridge, (par exemple, Ullah et alii, 1981), on calcule l'estimateur ridge en mettant d'abord la matrice $Z_L' Z_L$ sous forme d'une matrice de corrélation, puis en diagonalisant la matrice obtenue. L'équation (5.3), mise sous forme de corrélation, peut s'écrire, (la constante disparaissant de $\underline{\beta}$) :

$$(5.6) \quad \tilde{Y}_L = \tilde{Z}_L' \hat{\Sigma}^{1/2} \hat{\Sigma}^{-1/2} \underline{\beta} + \tilde{u}$$

où le tilde indique que les variables sont mises en écart à la moyenne et où $\hat{\Sigma} = \text{diag}[1/\sum_{t=1}^n \tilde{Z}_{it}^2]$, $i = 1, \dots, p$. On réécrit, en diagonalisant

$$(5.7) \quad \tilde{Y}_L = W \underline{\alpha} + \tilde{u}$$

$$\text{où } W = (\tilde{Z}_L' \hat{\Sigma}^{1/2})G, \quad G'G = GG' = I_p$$

$$\underline{\alpha} = G'(\hat{\Sigma}^{1/2} \underline{\beta}), \quad W'W = \Lambda,$$

$$\Lambda = \text{diag}[\lambda_1, \dots, \lambda_p], \quad \text{matrice des valeurs caractéristiques de la matrice.}$$

L'estimateur MCO de l'équation (5.7) est alors

$$(5.8) \quad \hat{\underline{\alpha}} = (W'W)^{-1} W' \tilde{Y}_L = \Lambda^{-1} W' \tilde{Y}_L$$

L'estimateur ridge est alors

$$(5.9) \quad \begin{aligned} \hat{\underline{\alpha}}_r &= [\Lambda + K]^{-1} W' \tilde{\underline{Y}}_L \\ &= \Delta \hat{\underline{\alpha}} \end{aligned}$$

où $\Delta = \text{diag}[\lambda_i / (\lambda_i + \lambda_i)]$, $i = 1, \dots, p$.

Le problème principal qui se pose ici concerne la façon de calculer K (ou k). Lin et Kmenta (1982) ont trouvé que deux formules de calcul donnaient de bons résultats par rapport à ceux obtenus avec un MCO dans une situation de forte multicollinéarité. La première, due à Hoerl, Kennard et Baldwin (1975)

$$(5.10) \quad k_{\text{HKB}} = p \hat{\sigma}^2 / \hat{\beta}'_L \hat{\beta}_L$$

se justifie du fait que c'est la valeur de k qui minimise la somme des erreurs quadratiques moyennes quand $Z'_L Z_L = I$. L'autre est basée sur une distribution a priori de $\underline{\alpha}$ qui suivrait une $N(\underline{0}, w^2 I_p)$, où w^2 est fixé d'avance. Dempster (1973) en tire que

$$(5.11) \quad \sum_{i=1}^p \hat{\alpha}_i^2 / \sigma^2 [(1/k) + (1/\lambda_i)] \sim \chi_p^2$$

où $k = \sigma^2 / w^2$.

En remplaçant σ^2 par $\hat{\sigma}^2$ et comme $E(\chi_p^2) = p$, la valeur de k est trouvée en résolvant par itération

$$(5.12) \quad \sum_{i=1}^p \hat{\alpha}_i^2 / \hat{\sigma}^2 [(1/k) + (1/\lambda_i)] = p$$

Il est utile de remarquer ici que la valeur de k est tirée de l'échantillon. Il est en ce sens stochastique et il ne correspond pas à de l'information a priori. C'est plutôt comme le disent Lin et Kmenta (1982), "a convenient way for trading bias for a reduction in variance".

Les deux interprétations que nous allons voir permettent d'arriver à l'estimateur ridge avec de l'information a priori. Elles lui donnent ainsi un sens autre que purement statistique.

3. Régression bayésienne

On peut interpréter l'estimateur ridge, dans une perspective bayésienne, comme procédant d'un $\underline{\beta}$ stochastique suivant une distribution spécifiée a priori. La valeur de K ne dépend donc aucunement de l'échantillon utilisé.

Dans le cas non linéaire, Dagenais (1983) a récemment proposé une interprétation bayésienne pour les grands échantillons.

Si les erreurs sont normales et indépendantes, alors la densité conjointe a posteriori de $(\underline{\beta}, \sigma)$ est donnée par

$$\begin{aligned}
 (5.13) \quad p(\underline{\beta}, \sigma | \underline{Y}) &= \frac{p(\underline{\beta}, \sigma) p(\underline{Y} | \underline{\beta}, \sigma)}{p(\underline{Y})} \\
 &\propto p(\underline{\beta}, \sigma) \ell(\underline{Y} | \underline{\beta}, \sigma) \\
 &\propto p(\underline{\beta}, \sigma) \sigma^{-N} \exp \{-\underline{u}'\underline{u}/2\sigma^2\}^1
 \end{aligned}$$

¹Dans les équations (11) et (12) de Dagenais (1983), on doit lire σ^{-N} et non $\sigma^{-N/2}$.

où $p(\underline{\beta}, \sigma | Y)$ est la densité conjointe a posteriori de $(\underline{\beta}, \sigma)$, $p(\underline{\beta}, \sigma)$ la densité conjointe a priori de $(\underline{\beta}, \sigma)$ et $\ell(Y|\underline{\beta}, \sigma)$ et $p(Y|\underline{\beta}, \sigma)$ la fonction de vraisemblance de l'échantillon.

Par un développement en séries de Taylor, et en ne gardant que les termes jusqu'au deuxième ordre, on obtient :

$$(5.14) \quad p(\underline{\beta}, \sigma | Y) \propto p(\underline{\beta}, \sigma) \sigma^{-N} \exp \left\{ -\frac{\underline{u}'_L \underline{u}_L}{2\sigma^2} \right\} \\ \exp \left\{ -(\underline{\beta} - \underline{\beta}_L)' H_L (\underline{\beta} - \underline{\beta}_L) / 2\sigma^2 \right\}$$

où $\underline{u}_L = Y - f(X; \underline{\beta}_L)$

$$H_L = \left(\frac{\partial^2 \underline{u}'_L \underline{u}_L}{\partial \underline{\beta} \partial \underline{\beta}'} \right)_{\underline{\beta} = \underline{\beta}_L}$$

et $\left(\frac{\partial \underline{u}'_L \underline{u}_L}{\partial \underline{\beta}} \right)_{\underline{\beta} = \underline{\beta}_L} = 0$

Si la probabilité a priori $p(\underline{\beta}, \sigma)$ est basée sur un échantillon hypothétique non négligeable (par rapport à celui dont on dispose pour l'estimation), et si la densité conditionnelle a priori de $\underline{\beta}$, étant donné σ , est posée comme une $N(0, K^{-1} \sigma^2)$

$$(5.15) \quad p(\underline{\beta} | \sigma) \propto \frac{1}{\sigma^p} \exp \left\{ -\frac{\underline{\beta}' K \underline{\beta}}{2\sigma^2} \right\}$$

et si on ne connaît rien de σ , c'est-à-dire qu'on pose avec Jeffrey et Zellner que

$$(5.16) \quad p(\ln \sigma) \propto \text{constante} \quad (\text{ou } p(\sigma) \propto 1/\sigma)$$

alors on peut écrire

$$(5.17) \quad p(\underline{\beta}, \sigma | \underline{Y}) \propto \frac{1}{\sigma^{N+p+1}} \exp \left\{ - \frac{\underline{u}'_L \underline{u}_L}{2\sigma^2} \right\} \\ \exp \left\{ - \frac{(\underline{\beta} - \underline{\beta}_L)' H_L (\underline{\beta} - \underline{\beta}_L)}{2\sigma^2} \right\} \\ \exp \left\{ - \frac{\underline{\beta}' K \underline{\beta}}{2\sigma^2} \right\}$$

Comme, lorsque l'échantillon est grand, $Z'_L Z_L$ tend vers H_L (Berndt et alii, 1974), on réécrit

$$(5.18) \quad p(\underline{\beta}, \sigma | \underline{Y}) \propto \frac{1}{\sigma^{N+p+1}} \exp \left\{ - \frac{1}{2\sigma^2} [SE + (\underline{\beta}' - \underline{\beta}'_L) Z'_L Z_L (\underline{\beta} - \underline{\beta}_L) + \underline{\beta}' K \underline{\beta}] \right\}$$

où $SE = \underline{u}'_L \underline{u}_L$.

En complétant le carré, on peut finalement transformer l'expression entre crochets dans (5.18) de telle sorte que

$$(5.19) \quad p(\underline{\beta}, \sigma | \underline{Y}) \propto \frac{1}{\sigma^{N+p+1}} \exp \left\{ - \frac{1}{2\sigma^2} [\phi + (\underline{\beta} - \hat{\underline{\beta}}_r)' (Z'_L Z_L + K) (\underline{\beta} - \hat{\underline{\beta}}_r)] \right\}$$

où $\phi = SE + \hat{\underline{\beta}}'_r (Z'_L Z_L + K) \hat{\underline{\beta}}_r + \underline{\beta}'_L Z'_L Z_L \underline{\beta}_L$

et $\hat{\underline{\beta}}_r = (Z'_L Z_L + K)^{-1} Z'_L Z_L \underline{\beta}_L$ est l'estimateur ridge (5.5).

En intégrant (5.19) par rapport à σ , on obtient la marginale a posteriori de $\underline{\beta}$ qui a la forme

$$(5.20) \quad p(\underline{\beta} | \underline{Y}) \propto [\phi + (\underline{\beta} - \hat{\underline{\beta}}_r)' (Z'_L Z_L + K) (\underline{\beta} - \hat{\underline{\beta}}_r)]^{-(N+p)/2}$$

C'est là une student qui, dans un grand échantillon (> 30) suit alors approximativement une normale de moyenne

$$(5.21) \quad E(\underline{\beta} | \underline{Y}) = \hat{\underline{\beta}}_r$$

4. Régression mixte

L'interprétation de la régression ridge non linéaire dans les termes de la régression mixte va suivre elle aussi des lignes asymptotiques. Il va s'agir ici aussi d'information a priori, mais $\underline{\beta}$ ne sera pas vu comme aléatoire. C'est une valeur fixe inconnue.

On prend pour point de départ l'équation (5.3) qui est la fonction linéarisée à son point de convergence. En lui ajoutant ensuite l'une des deux équations suivantes :

$$(5.22) \quad \underline{0} = \sqrt{k} I_p \underline{\beta} + \underline{v}$$

ou

$$(5.23) \quad \underline{0} = k^{1/2} \underline{\beta} + \underline{v}$$

on obtient - avec (5.23) -,

$$(5.24) \quad \begin{bmatrix} \underline{Y}_L \\ 0 \end{bmatrix} = \begin{bmatrix} \underline{Z}_L \\ k^{1/2} \end{bmatrix} \underline{\beta} + \begin{bmatrix} \underline{u} \\ \underline{v} \end{bmatrix}$$

On pose que $\underline{v} \sim N(0, \Omega)$ et que \underline{u} et \underline{v} sont indépendants, ou, à tout le moins, ont une covariance nulle. La variance de $[\underline{u} \ \underline{v}]'$ est alors

$$(5.25) \quad \Sigma_M = \begin{bmatrix} \sigma^2 I_P & 0 \\ 0 & \Omega \end{bmatrix}$$

En appliquant les moindres carrés généralisés, on obtient l'estimateur

$$(5.26) \quad \hat{\underline{\beta}}_M = (Z'_L Z_L / \sigma^2 + K^{1/2} \Omega^{-1} K^{1/2})^{-1} Z'_L \underline{Y}_L / \sigma^2$$

On peut le réécrire

$$(5.27) \quad \hat{\underline{\beta}}_M = (Z'_L Z_L + \sigma^2 K^{1/2} \Omega^{-1} K^{1/2})^{-1} Z'_L \underline{Y}_L$$

Pour que cet estimateur $\hat{\underline{\beta}}_M$ soit semblable à l'estimateur ridge $\hat{\underline{\beta}}_K$, il faut préciser la forme des éléments de $K^{1/2}$ et Ω . Si l'information dont on dispose a priori nous permet de dire que très probablement la valeur du paramètre β_j se trouve entre les valeurs a_j et b_j (où $a_j < b_j$)

$$(5.28) \quad a_j \leq \beta_j \leq b_j ,$$

alors on peut écrire

$$(5.29) \quad \beta_j = (a_j + b_j)/2 + u_j$$

où, suivant Theil et Goldberger (1961, p. 72), on se donne un intervalle de confiance de 'deux sigmas', en posant

$$(5.30) \quad u_j \sim N(0, (b_j - a_j)^2/16)$$

L'information a priori peut alors prendre la forme

$$(5.31) \quad \begin{matrix} [(a_j + b_j)/2] \\ (px1) \end{matrix} = I_p \underline{\beta} + [u_j] \quad j = 1, \dots, p$$

En multipliant les deux membres de l'équation (5.31) par la matrice $\text{diag} \left[\frac{4\sigma}{b_j - a_j} \right]$, on fait disparaître l'hétéroscédasticité

$$(5.32) \quad \text{diag} \left[\frac{4\sigma}{b_j - a_j} \right] \begin{bmatrix} a_j + b_j \\ 2 \end{bmatrix} = \text{diag} \left[\frac{4\sigma}{b_j - a_j} \right] \underline{\beta} + \text{diag} \left[\frac{4\sigma}{b_j - a_j} \right] [u_j]$$

La covariance entre les divers éléments d'information a priori est supposée nulle. Après cette opération, l'estimateur $\hat{\underline{\beta}}_M$, donné par l'équation (5.27) devient

$$(5.33) \quad \hat{\underline{\beta}}_M = (Z_L' Z_L + K)^{-1} Z_L' Y_L = \hat{\underline{\beta}}_T$$

Il nous reste ensuite à comparer terme à terme (5.23) et (5.32) et à résoudre pour a_j et b_j :

$$(5.34) \quad \underline{0} = \text{diag} \left[\frac{4\sigma}{b_j - a_j} \right] \begin{bmatrix} a_j + b_j \\ 2 \end{bmatrix}$$

et

$$K^{1/2} = \text{diag} \left[\frac{4\sigma}{b_j - a_j} \right]$$

Des équations (5.34), on obtient, si $b_j > a_j$,

$$(5.35) \quad a_j = -b_j$$

et

$$b_j = 2\sigma K_j^{1/2}$$

On peut donc considérer l'estimateur ridge comme un estimateur mixte incorporant les restrictions a priori suivantes :

$$(5.36) \quad -2\sigma K_j^{1/2} \leq \beta_j \leq 2\sigma K_j^{1/2}$$

Ces différentes méthodes mériteraient d'être examinées de plus près afin d'en mieux connaître les propriétés et d'en évaluer les effets sur des problèmes concrets de multicollinéarité comme celui que nous avons rencontré lors de l'estimation de notre modèle.

CONCLUSION

Les méthodes que nous avons utilisées ne nous ont pas permis de parvenir à des conclusions fermes sur le marché que nous avons étudié. Des problèmes de multicollinéarité, (que la forme du modèle rend probables) nous ont empêché de nous prononcer sur l'existence de déséquilibre sur le marché du bois de sciage américain.

Les résultats ne sont pas que négatifs cependant. La méthode d'estimation que nous avons utilisée, les moindres carrés itérés, s'est révélée très pratique et plus efficace que la méthode de Gauss. Elle est plus simple à mettre en oeuvre puisqu'elle ne demande pas les dérivées premières (elles sont implicites dans les équations normales des moindres carrés). Si la fonction a un maximum, sa convergence est assurée, même si le nombre d'itérations pour y parvenir peut être élevé.

Il faudrait maintenant poursuivre plus avant et refaire les estimations en incorporant les méthodes ridge, mixte et bayésienne que nous avons vues. Ces méthodes, en particulier la méthode ridge et les critères de choix de la valeur de k , demandent encore des recherches. C'est donc dire qu'il reste bien du travail à faire.

L'effort mérite d'être fait. Le grand intérêt suscité par la théorie du déséquilibre doit pouvoir s'accompagner d'un examen de la

validité empirique de l'hypothèse de déséquilibre. Dans ce contexte un modèle comme celui de Dagenais (1980), - modèle qui retient à la fois un caractère dynamique et une détermination stochastique des prix -, permet d'étendre l'éventail des situations concrètes que nous pouvons examiner en économétrie du déséquilibre. Son étude devrait donc être poursuivie.

APPENDICES

APPENDICE A

$$(3.1) \quad D_t = \beta_0 + \underline{X}'_t \underline{\beta}_1 + \underline{X}'_{dt} \underline{\beta}_2 + \beta_3 P_t + \beta_4 (D_{t-1} - Q_{t-1}) + u_{1t}$$

$$(3.6) \quad S_t = \psi_0 + \underline{X}'_t \underline{\psi}_1 + \underline{X}'_{st} \underline{\psi}_2 + \psi_3 P_t + \psi_4 I_{t-1} + u_{5t}$$

$$(A.1) \quad P_t = \gamma(D_t - S_t) + u_{3t}$$

$$(3.3) \quad Q_t = \min(D_t, S_t)$$

De l'équation (A.1) on obtient

$$(A.2) \quad D_t = 1/\gamma P_t - 1/\gamma P_{t-1} + S_t - 1/\gamma u_{3t}$$

En retardant (A.2) d'une période et en l'insérant dans (3.1) on trouve :

$$(A.3) \quad D_t = (\beta_0 + \beta_4 \psi_0) + \underline{X}'_t \underline{\beta}_1 + \underline{X}'_{dt} \underline{\beta}_2 + \beta_3 P_t + \beta_4 (1/\gamma + \psi_3) P_{t-1} \\ - \beta_4/\gamma P_{t-2} + \underline{X}'_{t-1} \beta_4 \underline{\psi}_1 + \underline{X}'_{st-1} \beta_4 \underline{\psi}_2 + \beta_4 \psi_4 I_{t-2} \\ - \beta_4 Q_{t-1} + (\beta_4 u_{5t-1} - \beta_4/\gamma u_{3t-1} + u_{1t})$$

Les équations (3.6) et (A.3) dans l'équation (A.1) nous donnent :

$$\begin{aligned}
(A.4) \quad P_t &= \gamma(\beta_0 + \beta_4 \psi_0 - \psi_0)/\delta + \underline{X}'_t(\underline{\beta}_1 - \underline{\psi}_1)\gamma/\delta + \underline{X}'_{t-1} \gamma \beta_4 \underline{\psi}_1/\delta \\
&+ \underline{X}'_{dt} \gamma \underline{\beta}_2/\delta - \underline{X}'_{st} \gamma \underline{\psi}_2/\delta + \underline{X}'_{st-1} \gamma \beta_4 \underline{\psi}_2/\delta \\
&- \gamma \psi_4/\delta I_{t-1} + \gamma \beta_4 \psi_4/\delta I_{t-2} - \gamma \beta_4/\delta Q_{t-1} \\
&+ [1 + \gamma \beta_4(1/\gamma + \psi_3)]/\delta P_{t-1} - \beta_4/\delta P_{t-2} + [\gamma Z_t - \gamma u_{5t} + u_{3t}]/\delta
\end{aligned}$$

$$\text{où } \delta = 1 - \gamma(\beta_3 - \psi_3)$$

$$Z_t = (\beta_4 u_{5t-1} - \beta_4 u_{3t-1} + u_{1t}).$$

On peut réécrire (A.4) brièvement :

$$\begin{aligned}
(A.5) \quad P_t &= \xi_0 + \underline{X}'_t \underline{\xi}_1 + \underline{X}'_{t-1} \underline{\xi}_2 + \underline{X}'_{dt} \underline{\xi}_3 + \underline{X}'_{st} \underline{\xi}_4 + \underline{X}'_{st-1} \underline{\xi}_5 + \xi_6 I_{t-1} \\
&+ \xi_7 I_{t-2} + \xi_8 Q_{t-1} + \xi_9 P_{t-1} + \xi_{10} P_{t-2} + \eta_t
\end{aligned}$$

De là, on trouve :

$$\beta_4 = -\frac{\xi_7}{\xi_6} = -\frac{(\gamma \beta_4 \psi_4)/\delta}{-\gamma \psi_4/\delta}$$

$$\psi_4 = -\frac{\xi_7}{\xi_8} = -\frac{(\gamma \beta_4 \psi_4)/\delta}{-\gamma \beta_4/\delta}$$

$$\underline{\beta}_2 = -\frac{\xi_3 \psi_4}{\xi_6} = -\frac{\gamma \underline{\beta}_2 \psi_4/\delta}{-\gamma \psi_4/\delta}$$

$$\underline{\psi}_2 = \frac{\xi_4 \psi_4}{\xi_6} = \frac{-\gamma \underline{\beta}_2 \psi_4/\delta}{-\gamma \psi_4/\delta}$$

$$\underline{\psi}_1 = -\frac{\underline{\xi}_2 \psi_4}{\beta_4 \xi_6} = -\frac{\gamma \beta_4 \psi_4 \underline{\psi}_1 / \delta}{-\gamma \beta_4 \psi_4 / \delta}$$

$$\underline{\beta}_1 = -\frac{\psi_4 \underline{\xi}_1}{\xi_6} + \underline{\psi}_1 = -\frac{\psi_4 \gamma (\underline{\beta}_1 - \underline{\psi}_1) / \delta}{-\gamma \psi_4 / \delta} + \underline{\psi}_1$$

$$\gamma = -\frac{\underline{\xi}_2}{\underline{\psi}_1 \xi_{10}} = -\frac{\gamma \beta_4 \underline{\psi}_1 / \delta}{-\underline{\psi}_1 \beta_4 / \delta}$$

$$\begin{aligned} \psi_3 &= \left[\frac{\xi_9 \gamma \beta_4}{-\xi_8} - (1 + \beta_4) \right] / \gamma \beta_4 = \left[\frac{\gamma \beta_4 (1 + \beta_4 + \gamma \beta_4 \psi_3) / \delta}{\gamma \beta_4 / \delta} - (1 + \beta_4) \right] / \gamma \beta_4 \\ &= \frac{1 + \beta_4 + \gamma \beta_4 \psi_3 - 1 - \beta_4}{\gamma \beta_4} = \psi_3 \end{aligned}$$

$$\begin{aligned} \beta_3 &= \left[\frac{\gamma \psi_4}{\xi_6} - (1 + \gamma \psi_3) \right] / \gamma = \left[\frac{\gamma \psi_4 \delta}{-\gamma \psi_4} - (1 + \gamma \psi_3) \right] / \gamma \\ &= -[1 - \gamma \beta_3 + \gamma \psi_3 - 1 - \gamma \psi_3] / \gamma = \frac{\gamma \beta_3}{\gamma} = \beta_3 \end{aligned}$$

APPENDICE B

Sources des données

Les données sur les variables I (stocks) et Q (livraisons, quantités transigées) proviennent du Business Statistics, 1961-1971, le supplément biannuel du Survey of Current Business; les prix du bois (P) et du contre-plaqué (PLY) proviennent de Wholesale Price Index, 1957-1970; le taux horaire moyen de salaire dans les scieries (W) est tiré de Employment and Earnings, 1957-1970; les mises en chantier (HS) proviennent elles aussi du Business Statistics. Ce sont, en milliers, les mises en chantier non agricoles permanentes.

Nous avons colligé ces données de la thèse de Gita Sen (1977). Dans le cas des stocks, il faut noter qu'une erreur s'y est glissée : les données pour 1957 et 1958 correspondent à une définition différente de celle des données subséquentes. Les données que nous présentons comprennent les corrections nécessaires pour rendre la série homogène. Ces corrections se retrouvent dans la livraison de janvier 1964, p. 28, du Survey of Current Business.

Données sur le marché du bois de sciage (Southern Pine)

Année	Mois	I	Q	P	PLY	W	HS
1957	Janvier	1 576	498	117,8	92,1	1,75	64,2
	Février	1 578	490	116,1	91,6	1,76	65,8
	Mars	1 542	579	115,5	91,1	1,77	87,0
	Avril	1 522	577	115,1	92,1	1,78	93,7
	Mai	1 484	623	114,7	92,4	1,80	103,0
	Juin	1 458	565	114,1	94,2	1,84	99,9
	Juillet	1 426	593	114,5	92,6	1,82	97,8
	Août	1 380	627	113,5	89,3	1,83	100,0
	Septembre	1 353	574	113,3	88,3	1,84	91,9
	Octobre	1 331	626	113,4	91,6	1,84	97,0
	Novembre	1 336	486	113,9	90,6	1,84	78,2
	Décembre	1 399	403	113,7	89,1	1,81	63,4
1958	Janvier	1 449	489	113,4	89,1	1,77	67,9
	Février	1 460	427	113,0	85,3	1,78	66,1
	Mars	1 458	498	112,2	83,9	1,79	81,4
	Avril	1 439	544	111,7	86,6	1,79	99,1
	Mai	1 392	570	112,1	82,5	1,84	108,5
	Juin	1 373	542	112,0	87,6	1,84	113,0
	Juillet	1 331	581	111,9	94,0	1,86	112,8
	Août	1 270	592	112,2	97,6	1,88	124,0
	Septembre	1 200	621	113,2	101,0	1,89	121,0
	Octobre	1 173	651	113,9	102,2	1,89	115,0
	Novembre	1 173	543	114,4	97,4	1,88	109,4
	Décembre	1 224	487	114,1	94,9	1,87	91,2
1959	Janvier	1 253	483	114,1	96,1	1,84	99,2
	Février	1 257	482	113,8	103,4	1,84	100,0
	Mars	1 216	599	114,3	104,3	1,85	130,7
	Avril	1 165	631	115,4	107,7	1,87	155,9
	Mai	1 131	602	116,0	107,5	1,90	156,0
	Juin	1 085	610	116,9	104,8	1,94	153,4
	Juillet	1 043	633	117,4	99,9	1,93	149,7
	Août	1 000	583	117,7	97,2	1,96	142,4
	Septembre	998	589	118,5	89,1	1,96	140,0
	Octobre	1 017	588	118,6	89,1	1,95	123,3
	Novembre	1 073	475	118,5	85,3	1,94	106,5
	Décembre	1 156	459	118,4	90,4	1,92	96,4

Données sur le marché du bois de sciage (Southern Pine) (suite)

Année	Mois	I	Q	P	PLY	W	HS
1960	Janvier	1 202	396	118,1	92,2	1,91	87,4
	Février	1 254	393	117,5	89,5	1,91	93,4
	Mars	1 312	429	117,2	86,5	1,93	93,9
	Avril	1 325	483	117,2	86,9	1,92	124,8
	Mai	1 358	483	116,8	85,9	1,93	133,8
	Juin	1 397	501	116,0	85,5	1,98	128,2
	Juillet	1 419	422	114,7	85,5	1,98	118,3
	Août	1 412	474	113,8	84,0	1,99	135,1
	Septembre	1 405	478	113,0	87,1	1,99	102,6
	Octobre	1 423	452	112,0	88,3	1,95	113,2
	Novembre	1 465	410	110,8	86,8	1,91	94,5
	Décembre	1 463	382	110,4	86,2	1,91	70,9
1961	Janvier	1 461	391	109,6	80,1	1,92	73,1
	Février	1 460	372	108,6	79,6	1,91	79,3
	Mars	1 409	525	109,0	81,6	1,92	109,3
	Avril	1 389	470	110,1	92,9	1,97	117,1
	Mai	1 373	545	110,4	89,8	1,98	131,6
	Juin	1 370	502	110,3	89,8	2,00	140,6
	Juillet	1 341	470	110,0	89,9	2,02	129,9
	Août	1 317	542	109,9	86,9	1,78	130,3
	Septembre	1 292	509	110,0	84,3	1,83	131,2
	Octobre	1 268	529	110,1	79,8	1,80	129,9
	Novembre	1 288	471	110,2	80,7	1,79	106,1
	Décembre	1 352	357	109,8	79,9	1,79	86,6
1962	Janvier	1 396	373	95,9	85,1	1,82	83,6
	Février	1 415	434	95,8	86,5	1,78	78,5
	Mars	1 401	507	96,2	90,6	1,79	118,1
	Avril	1 382	490	96,3	90,6	1,81	152,5
	Mai	1 361	565	96,3	87,3	1,81	157,6
	Juin	1 353	503	96,4	88,4	1,84	140,2
	Juillet	1 333	492	95,9	87,5	1,82	140,0
	Août	1 342	524	95,5	87,1	1,83	149,5
	Septembre	1 340	475	95,3	87,2	1,85	117,0
	Octobre	1 343	522	95,2	86,7	1,82	138,0
	Novembre	1 344	467	95,2	85,6	1,84	122,5
	Décembre	1 388	352	94,7	84,6	1,84	94,9

Données sur le marché du bois de sciage (Southern Pine) (suite)

Année	Mois	I	Q	P	PLY	W	HS
1963	Janvier	1 420	438	94,3	84,7	1,81	81,8
	Février	1 450	414	94,2	84,8	1,82	90,7
	Mars	1 453	477	94,4	85,9	1,82	128,3
	Avril	1 431	532	94,6	85,6	1,80	166,2
	Mai	1 410	575	95,0	85,4	1,83	176,0
	Juin	1 395	514	95,3	88,3	1,86	157,7
	Juillet	1 358	570	96,3	102,3	1,86	153,1
	Août	1 327	570	96,6	107,7	1,91	148,1
	Septembre	1 334	518	96,4	87,6	1,94	148,1
	Octobre	1 327	576	95,8	87,2	1,92	174,5
	Novembre	1 313	518	95,8	86,9	1,91	118,9
	Décembre	1 337	404	95,6	86,7	1,92	98,6
1964	Janvier	1 368	469	95,5	85,0	1,94	96,7
	Février	1 389	481	95,5	86,9	1,93	102,6
	Mars	1 400	537	95,8	90,4	1,92	131,5
	Avril	1 394	596	95,9	90,9	1,93	145,9
	Mai	1 363	576	96,0	89,8	1,94	156,3
	Juin	1 328	589	95,7	87,1	1,95	163,6
	Juillet	1 305	586	95,5	86,6	1,97	144,6
	Août	1 290	541	95,3	85,8	1,99	140,1
	Septembre	1 284	549	95,3	85,7	2,01	122,8
	Octobre	1 262	547	95,4	84,7	1,99	145,7
	Novembre	1 272	501	95,4	83,8	1,96	115,1
	Décembre	1 362	444	95,1	83,2	1,96	97,7
1965	Janvier	1 355	513	95,4	90,6	1,96	85,8
	Février	1 367	491	95,6	88,0	2,00	83,0
	Mars	1 360	579	95,5	86,6	2,00	124,1
	Avril	1 348	578	95,5	86,2	1,97	151,3
	Mai	1 278	595	95,6	85,1	2,01	157,7
	Juin	1 233	590	96,0	83,8	2,04	158,6
	Juillet	1 203	598	96,6	84,7	2,04	141,5
	Août	1 132	620	97,7	90,7	2,05	131,6
	Septembre	1 108	612	98,8	88,6	2,06	126,2
	Octobre	1 078	584	99,3	85,6	2,05	135,2
	Novembre	1 067	558	100,1	85,8	2,04	112,7
	Décembre	1 087	531	100,8	86,4	2,02	101,9

Données sur le marché du bois de sciage (Southern Pine) (suite)

Année	Mois	I	Q	P	PLY	W	HS
1966	Janvier	1 079	512	101,4	89,4	2,02	81,9
	Février	1 080	508	102,6	89,6	2,05	79,0
	Mars	1 027	673	103,6	95,9	2,04	122,4
	Avril	983	607	107,6	103,3	2,08	143,0
	Mai	974	582	108,5	99,6	2,10	133,9
	Juin	996	595	108,5	86,8	2,12	123,8
	Juillet	1 002	510	108,7	84,7	2,13	100,1
	Août	1 022	557	109,4	82,2	2,15	103,7
	Septembre	1 061	523	109,4	81,0	2,16	91,9
	Octobre	1 117	485	108,6	79,0	2,14	79,1
	Novembre	1 159	481	106,7	77,1	2,13	75,1
	Décembre	1 230	431	106,0	77,8	2,13	62,3
1967	Janvier	1 271	447	105,0	77,7	2,15	61,7
	Février	1 279	477	104,6	80,9	2,19	63,2
	Mars	1 286	568	105,1	79,9	2,19	92,9
	Avril	1 269	517	105,1	79,3	2,20	115,9
	Mai	1 292	536	105,3	78,9	2,21	134,2
	Juin	1 302	544	105,8	79,1	2,26	131,6
	Juillet	1 289	504	106,3	82,2	2,29	126,1
	Août	1 271	575	106,6	84,7	2,30	130,2
	Septembre	1 265	561	108,5	93,1	2,31	125,8
	Octobre	1 237	591	108,5	84,1	2,30	137,0
	Novembre	1 258	558	109,1	80,8	2,29	120,2
	Décembre	1 297	470	109,7	84,7	2,31	83,1
1968	Janvier	1 315	515	111,1	84,0	2,37	82,6
	Février	1 285	565	113,8	92,4	2,38	87,1
	Mars	1 269	600	116,1	95,9	2,41	128,7
	Avril	1 254	605	118,0	95,4	2,41	164,9
	Mai	1 215	646	119,7	95,4	2,44	144,5
	Juin	1 213	565	120,4	97,0	2,48	142,5
	Juillet	1 214	605	121,8	102,5	2,48	142,2
	Août	1 189	620	123,4	104,2	2,51	141,0
	Septembre	1 150	612	124,4	112,3	2,53	139,7
	Octobre	1 127	685	126,2	121,1	2,52	143,3
	Novembre	1 116	622	128,5	122,4	2,53	129,4
	Décembre	1 137	574	131,2	150,2	2,52	99,3

Données sur le marché du bois de sciage (Souther Pine) (suite)

Année	Mois	I	Q	P	PLY	W	HS
1969	Janvier	1 156	662	134,5	161,1	2,51	105,8
	Février	1 145	645	139,3	180,6	2,51	94,8
	Mars	1 111	704	146,6	180,0	2,54	135,6
	Avril	1 143	669	148,3	117,2	2,55	159,9
	Mai	1 190	619	147,8	104,1	2,58	157,7
	Juin	1 278	563	141,2	87,9	2,60	150,8
	Juillet	1 330	590	132,1	87,4	2,62	126,5
	Août	1 339	598	126,9	85,9	2,67	127,6
	Septembre	1 320	613	122,9	87,6	2,72	132,9
	Octobre	1 317	674	122,8	90,0	2,71	125,8
	Novembre	1 316	577	123,5	96,7	2,75	97,4
	Décembre	1 348	520	123,7	91,9	2,73	85,3

Note : Les indices de prix P et PLY avaient initialement comme années de référence 1947 - 49 = 100. Cette référence a été par la suite déplacée à 1957 - 59 = 100. Il faut multiplier P, pour la période janvier 1957 à décembre 1961 par le facteur 0,8718185. Les données PLY de la même période doivent être multipliées par 1,0676473.

REMERCIEMENTS

Je désire remercier d'abord mon directeur de thèse, M. Marcel Dagenais, pour son appui constant tout au long de l'élaboration de ce travail. Ses nombreux conseils et ses encouragements répétés ont facilité cette étude et n'ont pu que permettre d'en améliorer le contenu.

Je tiens à remercier M. Tran Cong Liem pour son aide patiente. Mme Suzanne Larouche-Sidoti est quant à elle responsable de l'excellente dactylographie de ce mémoire.

Je veux remercier finalement mes proches, et en particulier François, qui ont dû me subir durant ce travail et qui ne m'ont pas ménagé leur soutien.

BIBLIOGRAPHIE

- AMEMIYA, T. (1974a), "A note on a Fair and Jaffee Model", Econometrica, 42, 759-762.
- AMEMIYA, T. (1974b), "The Nonlinear Two-Stage Least Squares Estimator", Journal of Econometrics, Vol. 2, 105-110.
- AMEMIYA, T. (1975), "The Nonlinear Limited-Information Maximum Likelihood Estimator and the Modified Nonlinear Two-Stage Least Squares Estimator", Journal of Econometrics, Vol. 3, 375-386.
- APOSTOL, T.M. (1974), Mathematical Analysis, Second Edition, Addison-Wesley Publishing Company, Reading, Mass., XVII + 492 p.
- ATKINSON, K.E. (1978), An Introduction to Numerical Analysis, John Wiley and Sons, New York, XII + 587 p.
- BERNDT, E.K., B.H. HALL, R.E. HALL et J.A. HAUSMANN (1974), "Estimation and Inference in Nonlinear Structural Models", Annals of Economic and Social Measurement, 3/4, 653-665.
- BOWDEN, R. (1978), "Specification, Estimation and Inference for Models of Markets in Disequilibrium", International Economic Review, Vol. 19, No. 3, (octobre), 711-726.
- CHOW, G. (1983), Econometrics, McGraw-Hill Book Company, New York, XIII + 432 p.
- CLOWER, R. (1966), "The Keynesian Counterrevolution : A Theoretical Appraisal", in The Theory of Interest Rates, Ed., Hahn, S.H. et Brechling, S.P.R., Macmillan, New York.
- DAGENAIS, M.G. (1978), "The Computation of FIML Estimates as Iterative Generalized Least Squares Estimates in Linear and Nonlinear Simultaneous Equations Models", Econometrica, Vol. 46, No. 6, 1351-1362.
- DAGENAIS, M.G. (1980), "Specification and Estimation of a Dynamic Disequilibrium Model", Economic Letters, Vol. 5, No. 4, 323-328.
- DAGENAIS, M.G. (1983), "Extension of the Ridge Regression Technique to Non-Linear Models with Additive Errors", Economic Letters, Vol. 12, No. 2, 169-174.

- DAGENAIS, M.G. et M. DAGENAIS (1982), "Algorithme de sélection sous contraintes d'un sous-ensemble d'éléments homogènes", Ann. Sc. Math. Québec, Vol. II, no 1, 43-51.
- DEMPSTER, A.P. (1973), "Alternatives to Least Squares in Multiple Regression", in D. Kabe et R.P. Gupta (Eds.), Multivariate Statistical Inference, Amsterdam : North-Holland Pub. Co., 25-40.
- FAIR, R.C. et D.M. JAFFEE (1972), "Methods of Estimation for Markets in Disequilibrium", Econometrica, Vol. 40, 497-514.
- FISZ, M. (1963), Probability Theory and Mathematical Statistics, 3^e Edition, 1980 Printing, Robert E. Krieger Publishing Company, Huntington, New York, XVI + 677 p.
- FLETCHER, R. et M.J.D. POWELL (1963), "A Rapidly Convergent Descent Method for Minimization", Computer Journal, 163-168.
- GOLDBERGER, A., A.L. NAGAR et H.S. ODEH (1961), "The Covariance Matrices of Reduced-Form Coefficients and of Forecast for a Structural Econometric Model", Econometrica, Vol. 29, 556-573.
- GOURIEROUX, C., J.J. LAFFONT et A. MONFORT (1980a), "Coherency Conditions in Simultaneous Linear Equations Models with Endogeneous Switching Regimes", Econometrica, 48, 675-696.
- GOURIEROUX, C., J.J. LAFFONT et A. MONFORT (1980b), "Disequilibrium Econometrics in Simultaneous Equations Systems", Econometrica, 48, 75-96.
- HOERL, A.E. et R.W. KENNARD (1970), "Ridge Regression : Biased Estimation for Nonorthogonal Problems", Technometrics, Vol. 12, No. 1, (février), 55-67.
- HOERL, A.E., R.W. KENNARD et K.F. BALDWIN (1975), "Ridge Regression : Some Simulations", Communications in Statistics, 4(2), 105-123.
- ITO, T. et K. UEDA (1981), "Tests of the Equilibrium Hypothesis in Disequilibrium Econometrics : An International Comparison of Credit Rationing", International Economic Review, Vol. 22, No. 3, (octobre), 691-708.
- JACOBY, S. et J.T. PIZZO (1972), Iterative Methods for Nonlinear Optimization Problems, Prentice-Hall, Englewood Cliffs, 274 p.
- LAFFONT, J.J. et A. MONFORT (1976), "Econométrie des modèles d'équilibre avec rationnement", Annales de l'INSEE, No. 24, 3-39.

- LAFFONT, J.J. et A. MONFORT (1979), "Disequilibrium Econometrics in Dynamic Models", Journal of Econometrics, Vol. 11, 353-361.
- LIN, K. et J. KMENTA (1982), "Ridge Regression Under Alternative Loss Criteria", Review of Economics and Statistics, Vol. 64, 488-494.
- LOVELL, M. (1961), "Manufacturers' Inventories, Sales Expectations, and the Acceleration Principle", Econometrica, Vol. 29, No. 3, (July), 293-314.
- McCALLUM, B.T. (1974), "Competitive Price Adjustments : An Empirical Study", The American Economic Review, Vol. 64, No. 1, (March), 56-65.
- McKILLOP, W.L.M. (1967), "Supply and Demand for Forest Products - An Econometric Study", Hilgardia, Vol. 38, No. 1, (March), 1-132.
- MADDALA, G.S. (1977), Econometrics, McGraw-Hill Book Company, New York, XII + 516 p.
- MADDALA, G.S. et F.D. NELSON (1974), "Maximum Likelihood Methods for Models of Markets in Disequilibrium", Econometrica, Vol. 42, 1013-1030.
- OBERHOFER, W. et J. KMENTA (1974), "A General Procedure for Obtaining Maximum Likelihood Estimates in Generalized Regression Models", Econometrica, Vol. 42, No. 3, 579-590.
- ORSI, R. (1982), "On the Dynamic Specification of Disequilibrium Econometrics : An Analysis of Italian Male and Female Labor Markets", CORE, Discussion Paper No. 8228, Louvain-la-Neuve, 24 p.
- PATINKIN, D. (1948), "Price Flexibility and Full Employment", American Economic Review, Vol. 38, No. 4, (September), 543-564.
- QUANDT, R.E. (1982), "Econometric Disequilibrium Models", Econometric Review, Vol. 1, No. 1, 1-63.
- QUANDT, R.E. (1981), "Autocorrelated Errors in Simple Disequilibrium Models", Economic Letters, 7, 55-62.
- SARGAN, J.D. (1964), "Wages and Prices in the United Kingdom : A Study in Econometric Methodology", in Econometric Analysis for National Economic Planning, Ed. P.E. Hart, G. Mills, J.K. Whittaker, Butterworths, London, 25-54.
- SEN, G. (1977), The Theory of Estimation of Disequilibrium Markets, Ph.D. Thesis, Stanford University, 136 p.
- THEIL, H. (1963), "On the Use of Incomplete Prior Information in Regression Analysis", Journal of the American Statistical Association, Vol. 58, 401-414.

- THEIL, H. et A. GOLDBERGER (1961), "On Pure and Mixed Statistical Estimation in Economics", International Economic Review, Vol. 2, No. 1, (January), 65-78.
- ULLAH, A. et H.D. VINOD (1981), "Recent Advances in Regression Methods", Coll. Statistics : Textbooks and Monographs, Vol. 41, Marcel Dekker Inc., New York et Basil, 361 p.
- ULLAH, A., H.D. VINOD et R.K. KADIYALA (1981), "A Family of Improved Shrinkage Factors of the Ordinary Ridge Estimator", in Proceedings of the Econometric Society European Meeting, 1979, Ed. by E.G. Charatsis, North-Holland Publishing Co., 259-277.
- ZELLNER, A. (1971), An Introduction to Bayesian Inference in Econometrics, John Wiley and Sons, Inc., New York, XV + 431 p.

