

DETERMINATION D'UNE RELATION DE LONG TERME SUR DONNEES DE PANEL : ELEMENTS DE METHODE

Alain PIROTTE

Chercheur INRETS-DEST

Institut National de Recherche sur les Transports et leur Sécurité
Arcueil - France

INTRODUCTION

Une des finalités principales de la modélisation est son utilisation prédictive; on essaye ainsi de réduire l'incertitude pesant sur le futur en déterminant des évolutions possibles. Pour entreprendre ce type d'exercice, deux conditions essentielles sont requises : disposer de variables observées sur un laps de temps suffisamment long, et ne pas se tromper sur la spécification du modèle censé expliquer le phénomène analysé (variables explicatives, nombre de retards, forme de la relation).

Pour appréhender et comprendre un phénomène, les séries temporelles ou séries chronologiques constituaient jusqu'à un passé récent la principale source statistique disponible. Cependant, suivant les particularités du phénomène étudié la dimension temporelle des données peut être faible (LAMBERT et MADRE [1989]). Ce constat rend fragile les exercices de projections réalisés à partir d'une telle source d'information.

L'amélioration continue de l'appareil statistique Français a permis d'obtenir de nouvelles bases de données dénommées données individuelles temporelles ou données de panel. Ces nouvelles statistiques permettent de réaliser des analyses plus fines surtout en matière de comportements individuels et peuvent constituer une solution au problème posé par une dimension temporelle souvent insuffisante. La caractéristique essentielle réside dans leur double dimension temporelle (T) et individuelle (N), ainsi l'on ne dispose plus seulement de T observations mais de NT observations. Ceci permet d'augmenter le nombre de degrés de liberté et de réduire les problèmes de multicollinéarité [1] qui peuvent se poser parmi les variables explicatives, la richesse de l'information étant considérablement accrue par son volume et sa diversité. Les régressions porteront sur plusieurs dizaines de points, cela permet d'appliquer les propriétés asymptotiques des estimateurs (si il n'y a pas d'erreur de spécification, les estimations issues de méthodes convergentes doivent donner des résultats pratiquement identiques, on voit que cela peut être un moyen de tester la validité de la spécification).

Au niveau de la spécification, l'on adopte le plus souvent une approche dynamique car beaucoup plus traductrice des comportements des agents économiques, les ajustements ne se faisant pratiquement jamais instantanément. Ces modèles nous permettent alors de distinguer les effets de court et de long terme. Sur séries temporelles, la théorie de la cointégration assure la cohérence entre les trajectoires des variables à long terme et leur dynamique de court terme (Théorème d'ENGLE ET GRANGER [1987]). Sur données de panel, on raisonne le plus souvent en termes d'estimateur Between (ou inter) et d'estimateur Within (ou intra). Le premier correspond aux différences permanentes entre les individus, ce que l'on a couramment interprété comme étant la traduction des comportements de long terme. Quant au second, il retrace les fluctuations plus ou moins instantanées des comportements, on l'analyse comme traduisant les ajustements de court terme. BALTAGI et GRIFFIN [1984] ont montré que cette distinction est fragile, et que l'interprétation que l'on peut donner aux estimateurs dépend de la validité de la spécification (nombre de retards adéquats dans le modèle).

Lorsque l'on s'intéresse à l'analyse de l'évolution de la circulation globale sur le long terme (des véhicules légers et utilitaires) au niveau de la France entière (la finalité étant d'en réaliser des projections), on se heurte à la longueur insuffisante des séries (en raisonnant à un niveau annuel) des variables explicatives (parc automobile, revenu des ménages, prix des carburants). La solution adoptée a été de changer de cadre spatial en passant à un niveau régional (Les 22 régions Françaises), donc de raisonner à partir de données de panel.

Nous nous proposons d'exposer une méthode permettant d'appréhender la dynamique des ajustements dans le temps. Tout d'abord, nous présenterons deux spécifications utilisées sur données de panel, et plus particulièrement le modèle à effets fixes (nous montrerons qu'il peut répondre à des besoins précis en fonction d'une problématique donnée). Ensuite, nous décrirons les procédures d'estimations des modèles dynamiques à effets fixes. Enfin, nous aborderons le problème des fluctuations importantes que peuvent connaître certaines relations au cours du temps.

1 - DEUX SPECIFICATIONS POSSIBLES SUR DONNEES DE PANEL.

1.1 - Distinction modèles à erreurs composées-modèles à effets fixes : Pourquoi choisir un modèle à effets fixes ?

La méthodologie des panels propose essentiellement deux types de modèles :

(1) le modèle à effets fixes ou modèle de la covariance [2]

$$y_{it} = \theta + \sum_{j=1}^{\gamma} x_{j,it} \beta_j + \sum_{i=1}^N \alpha_i + \sum_{t=1}^T \phi_t + \varepsilon_{it}$$

(2) le modèle à effets aléatoires ou à erreurs composées [3] :

$$y_{it} = \sum_{j=1}^{\gamma} x_{j,it} \beta_j + \varepsilon_{it}$$

avec : $\varepsilon_{it} = \alpha_i + \lambda_t + u_{it}$

La différence fondamentale entre les deux modèles [4] est qu'un effet spécifique certain intervient dans l'explication de l'endogène pour (1). Dans (2), cet effet spécifique (individuel et/ou temporel) est aléatoire (chaque observation pour un individu i à l'année t des variables explicatives $x_{j,it}$ suffit à caractériser l'endogène y_{it}). Il ne se manifeste qu'au niveau de la perturbation. De façon plus précise, le modèle à erreurs composées revient à décomposer la perturbation en trois parties distinctes : une composante temporelle (qui constitue un résumé des variables omises dans la liste des variables explicatives et dont la valeur est identique pour tous les individus à un instant t), individuelle et résiduelle. Les méthodes d'estimations de ces deux modèles sont basées sur l'utilisation de la variabilité dans tout ou partie de ces dimensions. Ainsi, la

décomposition de la variance totale de chaque série en deux sous-variances orthogonales : temporelle (ou intra) et individuelle (ou inter), nous fournissons une indication sur la variabilité dominante.

Cette décomposition s'écrit : $Var_{tot} = Var_{ind} + Var_{temp}$

$$\text{Où : } Var_{tot} = \sum_{i=1}^N \sum_{t=1}^T (y_{it} - y_{..})^2 \quad \text{avec } y_{..} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T y_{it}$$

$$Var_{ind} = T \sum_{i=1}^N (y_{i.} - y_{..})^2 \quad \text{avec } y_{i.} = \frac{1}{T} \sum_{t=1}^T y_{it}$$

$$Var_{temp} = \sum_{i=1}^N \sum_{t=1}^T (y_{it} - y_{i.})^2$$

A des fins d'illustration [5], cette décomposition a été calculée pour les variables consommations de carburant (CARB), parc automobile (PARC) et revenu des ménages (REV). (cf. tableau n°1, ci-dessous).

Tableau n°1 : Caractéristiques de moyenne et de dispersion des variables exprimées en logarithmes pour les 22 régions Françaises (1973-1988).

Variable	Moyenne $y_{..}$	Ecart type		% Variance inter dans la variance totale
		Total	Intra	
CARB	13.879	10.655	2.4545	94,7
PARC	13.629	10.443	2.8663	92,5
REV	11.221	11.436	1.7428	97,6

On constate, pour chaque variable l'extrême prédominance de la variance inter dans la variance totale. On observe ainsi, que 94,7 % de la variance du logarithme des consommations de carburant est attribuable aux différences inter-régionales. Cette part s'élevant à 92,5 % pour le parc automobile et 97,6 % pour le revenu des ménages.

Suivant la spécification retenue et les méthodes d'estimations, on choisit de privilégier telle ou telle variabilité. Pour le modèle à erreurs composées, on ne décompose (le plus souvent) le terme résiduel qu'en incluant une perturbation aléatoire individuelle, sans considérer la perturbation temporelle car la dimension individuelle est la composante dominante sur données de panel (Cf. exemple ci-dessus). L'on cherche ainsi à neutraliser l'impact de ces "bruits" et non pas à les faire ressortir comme des "signaux" à estimer [6]. Concernant le modèle à effets fixes, il permet de capter les caractéristiques individuelles (et/ou temporelles) en éliminant par le biais des effets fixes toute la variabilité inter-individuelle (et/ou temporelle). L'information qui reste alors à analyser concerne la variabilité intra-individuelle-temporelle. Si l'on se situe dans une perspective prédictive, il faut obtenir des coefficients qui retracent les évolutions dans le temps, cette spécification réserve une place prépondérante à l'information temporelle indispensable pour un tel exercice.

Lorsque l'on cherche par exemple à analyser la circulation automobile au niveau régional, il est évident que des caractéristiques propres à chaque région vont venir influencer son niveau telles que la structure du réseau autoroutier, la population, etc... Si l'on reprend la distinction faite auparavant, le modèle à effets fixes correspond bien à la prise en compte de ces particularismes qu'ils importent d'évaluer.

Généralement dans un panel, l'hétérogénéité entre les individus est souvent forte, cependant certains d'entre eux peuvent avoir des comportements similaires. Une solution est alors de les regrouper en sous-ensembles homogènes. En ayant empilé les individus et mis des effets fixes individuels dans le modèle statique (sans variables retardées), on obtient une information sur les sous-ensembles qu'il convient de constituer (implicitement, on admet que cette spécification est la meilleure pour retracer l'hétérogénéité). Ils sont alors établis à partir des valeurs des coefficients affectés aux effets fixes et par rapport à la significativité de leurs students respectifs (MADRE et PIROTTE [1992]). La constitution de groupes homogènes doit également s'appuyer sur un recouplement avec d'autres critères (dans notre exemple, carte du réseau autoroutier, indices de concentration de population, etc...).

1.2 - Modèles à effets fixes et dynamique des comportements.

Une fois les groupes homogènes constitués [7], on peut passer à l'estimation de la relation retraçant la dynamique des comportements. En adoptant une spécification autorégressive à effets fixes [8], soit :

$$y_{it} = \delta y_{i,t-1} + \sum_{j=1}^{\gamma} x_{j,it} \beta_j + \sum_{i=1}^N \alpha_i + \varepsilon_{it}$$

Il n'est malheureusement pas possible d'appliquer l'estimateur des MCO (non convergent) du fait de la corrélation asymptotique entre l'endogène retardée $y_{i,t-1}$ et les perturbations ε_{it} .

La solution couramment envisagée pour résoudre ce problème est d'utiliser la méthode des variables instrumentales. La difficulté réside dans le choix des instruments, en effet il n'existe pas de méthodes permettant de sélectionner des instruments optimaux. Plusieurs approches ont été élaborées à partir du modèle réécrit en différences premières [9], soit :

$$y_{it} - y_{i,t-1} = \delta (y_{i,t-1} - y_{i,t-2}) + \beta (x_{it} - x_{i,t-1}) + \varepsilon_{it} - \varepsilon_{i,t-1}$$

ANDERSON-HSIAO [1982] ont proposé d'utiliser deux types de variables instrumentales :

$$Z_{it}^1 = (y_{i,t-2}, x_{it} - x_{i,t-1}) \text{ ou } Z_{it}^2 = (y_{i,t-2} - y_{i,t-3}, x_{it} - x_{i,t-1})$$

Ensuite ARELLANO [1989] a démontré qu'il valait mieux se servir de Z_{it}^1 , car la variance de l'estimateur qui utilise comme variable instrumentale $\Delta y_{i,t-2}$ peut prendre des valeurs très élevées.

ARELLANO et BOND [1987] ont élaboré une autre procédure permettant d'obtenir des estimateurs plus efficaces. Le principe a été d'utiliser les conditions d'orthogonalité des variables du modèle.

A titre d'exemple, si l'on raisonne en considérant un panel couvrant cinq périodes (t variant de 0 à 4), on a :

$$\text{pour } t = 2 : y_{i2} - y_{i1} = \delta (y_{i1} - y_{i0}) + (x_{i2} - x_{i1})\beta + \varepsilon_{i2} - \varepsilon_{i1}$$

$$\text{pour } t = 3 : y_{i3} - y_{i2} = \delta (y_{i2} - y_{i1}) + (x_{i3} - x_{i2})\beta + \varepsilon_{i3} - \varepsilon_{i2}$$

$$\text{pour } t = 4 : y_{i4} - y_{i3} = \delta (y_{i3} - y_{i2}) + (x_{i4} - x_{i3})\beta + \varepsilon_{i4} - \varepsilon_{i3}$$

L'idée a été de déterminer pour chaque t des instruments adéquats. Pour t=2, les instruments valides sont $(y_{i0}, x_{i2}-x_{i1})$, à t = 3 est associé $(y_{i0}, y_{i1}, x_{i3}-x_{i2})$ et à t = 4 $(y_{i0}, y_{i1}, y_{i2}, x_{i4}-x_{i3})$

La matrice des variables instrumentales s'écrit :

$$Z_i^o = \begin{pmatrix} y_{i0} & 0 & 0 & 0 & 0 & 0 & x_{i2}-x_{i1} & 0 & 0 \\ 0 & y_{i0} & y_{i1} & 0 & 0 & 0 & 0 & x_{i3}-x_{i2} & 0 \\ 0 & 0 & 0 & y_{i0} & y_{i1} & y_{i2} & 0 & 0 & x_{i4}-x_{i3} \end{pmatrix}$$

Enfin, on constate que l'écriture du modèle à effets fixes en différences premières induit un résidu suivant un processus de moyenne mobile d'ordre 1, ce qui ne permet pas d'obtenir des estimateurs pleinement efficaces. On peut se reporter à SEVESTRE et TROGNON [1992] pour tenir compte de cette structure particulière de l'aléa.

1.2.1 - Les effets de court et long terme

La validité de l'évaluation des effets de court et de long terme est conditionnelle à la bonne spécification du modèle. On définit, la réponse de court terme comme la valeur affecté à la variable exogène non retardée autrement dit exprimée à la date t, c'est l'ajustement qui se produit au cours de la période courante. Sous l'hypothèse de stationnarité, la réponse de long terme correspond au coefficient que nous nommerons R_L et qui est déterminée à partir du modèle (a). Soit :

$$(y_{it} - y_{i,t-1}) - \delta (y_{i,t-1} - y_{i,t-2}) = \beta (x_{it} - x_{i,t-1}) + \varepsilon_{it} - \varepsilon_{i,t-1}$$

$$\Rightarrow R_L = \frac{\beta}{1 - \delta}$$

Si sur la variable exogène portent plusieurs retards, le numérateur de R_L s'exprimera comme la somme des coefficients de cette variable explicative, son dénominateur restant inchangé. R_L traduit l'effet à long terme d'une modification d'une unité de $(x_{it} - x_{i,t-1})$ sur le niveau de l'endogène.

2 - MOINDRES CARRÉS ORDINAIRES RECURSIFS, MODELES A COEFFICIENTS VARIANT DANS LE TEMPS ET TESTS DE STABILITE.

La stabilité des paramètres joue un rôle particulièrement important lorsque l'on veut réaliser des projections. Leur instabilité peut refléter des phénomènes structurels (mauvaise spécification, variables omises), ou ponctuels dans le temps (choc pétrolier, mesures de politiques économiques, nouvelles réglementations, etc...). Une approche efficace pour tester la stabilité des coefficients est de les estimer à l'aide du filtre de Kalman (KALMAN [1960], HARVEY [1981a,1981b]) de façon récursive, c'est à dire en ajoutant les observations une par une. On peut alors estimer la performance d'adaptation du modèle aux observations. Nous allons exposer plus précisément ce procédé, en décrivant le principe des Moindres Carrés Ordinaires Récursifs (MCOR) et en présentant ensuite brièvement l'estimation des modèles à coefficients variant dans le temps à l'aide de l'algorithme du filtre de Kalman. Cette dernière approche revenant à généraliser la procédure MCOR en appréhendant les coefficients comme des variables aléatoires.

2.1 - Le principe des Moindres Carrés Récursifs (MCOR).

Soit le modèle : (b) $y_{it} = \sum_{j=1}^{\gamma} x_{j,it} \beta_j + \varepsilon_{it}$

Où : $t = 1, \dots, T$: nombre de périodes;

$i = 1, \dots, N$: nombre d'individus;

$j = 1, \dots, \gamma$: nombre de variables explicatives

$y_{it} : (y_{i1}, \dots, y_{iN1}, \dots, y_{i1T}, \dots, y_{iNT})$ variable endogène

$x_{j,it} : (x_{j,i1}, \dots, x_{j,iN1}, \dots, x_{j,i1T}, \dots, x_{j,iNT})$ variables explicatives

β_j : paramètre d'intérêt

$\varepsilon_{it} = (\varepsilon_{i1}, \dots, \varepsilon_{iN1}, \dots, \varepsilon_{i1T}, \dots, \varepsilon_{iNT})$ résidu aléatoire centré

L'application des MCO sur la totalité de l'échantillon (données empilées) donnent une estimation unique de chaque paramètre pour l'ensemble des individus. Maintenant, si l'on utilise la méthode des MCOR, cela revient à estimer le même modèle mais en ajoutant les observations une par une. On raisonne alors sur un échantillon de taille croissante et l'on obtient une suite d'estimateurs distincts qui sont le résultat de l'intégration successive des observations. Les estimations de la dernière régression coïncident évidemment avec celles des MCO. Les estimations récursives des coefficients sont ensuite représentées graphiquement, afin de donner un premier renseignement sur la constance ou non de chaque coefficient au cours de la période. Sur données de panel, il est hasardeux de commenter à partir de ces graphiques leurs évolutions réelles car l'information en t est conditionnelle à l'information passée et repose sur l'hypothèse d'homogénéité. On se contente dans un premier temps de raisonner à un niveau global en négligeant l'hétérogénéité, on doit donc prendre les estimations de ces coefficients comme donnant des ordres de grandeurs, l'objectif principal étant de repérer les ruptures marquantes.

2.2 - Algorithme du filtre de Kalman et coefficients variant dans le temps.

La relation (b) est supposée constante, or des ruptures brutales peuvent se produire. Pour les détecter précisément, le principe va être de spécifier les coefficients comme étant eux-mêmes des variables aléatoires. Ainsi, la valeur d'un coefficient à l'instant t (θ_{it}) sera fonction de son propre passé mais également des valeurs passées des autres coefficients et d'un terme aléatoire. Cette relation est appelée relation d'état. Il en découle que θ_{it} est une variable d'état d'état inobservable, qui possède une représentation dynamique et qui influence une relation dit de mesure qui elle est observable y_{it} . Un tel modèle est appelé Modèle état-mesure et s'exprime sous la forme [10] :

$$\begin{cases} y_{it} = \theta_{it} x_{it} + \varepsilon_{it} \\ \theta_{it} = \phi \theta_{i,t-1} + v_{it} \text{ avec } v_{it} \sim (0, \sigma^2) \end{cases}$$

L'objectif de la procédure d'estimation va consister à obtenir une prévision optimale de la variable d'état en utilisant les valeurs présentes et passées du phénomène observé y_{it} . Plus précisément, l'algorithme du filtre de Kalman permet le calcul récursif de l'approximation optimale d'une quantité non observable par l'intermédiaire d'une autre variable que l'on peut observer. Le minimum de la variance étant le critère d'optimalité, la prévision va être donnée par l'espérance conditionnelle de θ_{it} sachant les valeurs passées et présente des y_{i1}, \dots, y_{it} .

KALMAN [1960] a élaboré une procédure récursive optimale incluant des équations de mise à jour de l'espérance conditionnelle et de la variance de l'erreur d'estimation sur θ_{it} (une application détaillée est fournie par CHOW [1983]).

2.3 - Les tests de stabilité.

BROWN, DURBIN et EVANS [1975] ont développé des tests de stabilité basés sur les résidus récurrents (tests graphiques CUSUM et CUSUM Carré). Le premier [11] détecte les mouvements systématiques dans la valeur des coefficients, il teste leur constance au cours du temps. Il traduit en fait la validité de la spécification adoptée et les éventuelles modifications structurelles sur les paramètres.

Le second [12] est basé sur le même principe mais sa finalité est différente : il détecte les mouvements que l'on peut qualifier de plus "aléatoires". Il fournit une analyse temporelle détaillée de la totalité de la période, en nous informant sur les éventuels chocs et en nous donnant un possible découpage en sous-périodes. Cette information est particulièrement utile lorsque par exemple pour une année, il s'est produit des événements "atypiques" qui viennent perturber fortement le comportement moyen des coefficients affectés à la relation. Par la suite, une solution souvent adoptée consiste à introduire une variable indicatrice temporelle pour neutraliser l'impact de ce choc.

Le découpage en sous-périodes suggéré par le test CUSUM carré peut être confirmé en mettant en oeuvre le test de CHOW [1960]. Par exemple, supposons que le test CUSUM carré nous est permis de distinguer au sein de la période initiale A deux sous-périodes B et C, l'idée est de tester l'égalité des coefficients des deux sous-périodes. Pour cela, on effectue les régressions sur A, B et C et on récupère les sommes des carrés des résidus respectives pour calculer la statistique de CHOW [13] (sous l'hypothèse d'égalité cette statistique suit une loi de Fisher).

On s'aperçoit que les éléments d'information que nous apportent ces tests sont fondamentaux pour réaliser par la suite des projections. En effet, ils nous fournissent des renseignements pour élaborer une spécification temporellement correcte.

CONCLUSION

Les données de panel constituent une approche intéressante pour améliorer la perception de certains phénomènes difficilement analysables dans la seule dimension temporelle. Cependant différentes études (KMENTA [1978], STAPLETON [1981], BALTAGI ET GRIFFIN [1984]) montrent que malgré la double dimension des données de panel, si T est faible, le problème du à la mauvaise spécification des retards est toujours présent. On s'aperçoit ainsi que la mise en évidence des effets court et long terme "véritables" serait conditionnelle à la validité de la structure de la spécification dynamique retenue, qui est elle même dépendante des deux dimensions des données de panel (surtout de la dimension temporelle).

NOTES

- [1] - Pour une approche détaillée sur les problèmes qu'engendre la multicolinéarité, cf. ROUSSE [1990].
- [2] - Pour une présentation plus détaillée cf. KMENTA [1990].
- [3] - De nombreuses présentations de ce modèle existent cf. WALLACE et HUSSAIN [1969], NERLOVE [1971], MADDALA [1971].
- [4] - Cependant, MUNDLAK [1978] a montré que le modèle à effets fixes pouvait constituer un cas particulier d'un modèle à erreurs composées.
- [5] - On peut se reporter également à DORMONT [1989, p. 49-53].
- [6] - Cf. MAZODIER [1971, p. 47].
- [7] - Au sein de ces groupes, on fait toujours intervenir les effets fixes car aussi proches soient les comportements, une hétérogénéité "relative" est toujours présente. Si l'on omet ces variables, cela revient à imposer la nullité de leurs coefficients. On se heurte alors au dilemme biais-variance entre l'estimateur des MCO et des MCC si ces contraintes ne sont pas vérifiées.
- [8] - On peut également introduire des retards sur les exogènes, le raisonnement restant identique.
- [9] - Afin de simplifier les notations, on ne considère qu'une seule variable explicative non retardée. La présentation des différentes méthodes d'estimations s'inspire de SEVESTRE et TRÓGNON [1992].
- [10] - L'écriture du modèle état-mesure proposée est très simple car ne retenant qu'une seule variable explicative. Pour une approche plus détaillée et généralisatrice cf. MAUREL [1985].
- [11] - cf. HARVEY [1981b, p. 151-152].
- [12] - cf. HARVEY [1981b, p. 152].
- [13] - Pour un exemple d'application, cf. BLANCHARD et SEVESTRE [1987].

REFERENCES

- ANDERSON T.W. and C. HSIAO [1982] "Formulation and Estimation of Dynamics Models Using Panel Data", *Journal of Econometrics*, Vol. 18, pp. 578-606.
- ARELLANO M. [1988] "A Note on the Anderson-Hsiao Estimator for Panel Data", *mimeo Institute of Economics*, Oxford University.
- ARELLANO M. and S. BOND [1987] "Some Tests of Specification for Panel Data : Monte Carlo Evidence and an Application to Employment Equations" Working paper, *Institute for Fiscal Studies*, London.
- BALTAGI B.H. et J.M. GRIFFIN [1984] "Short and Long Run Effects in Pooled Models", *International Economic Review*, Vol. 25, pp. 631-645.
- BLANCHARD P. et P. SEVESTRE [1987] "Anticipations et Rattrapage : Une Etude de la Stabilité du Processus d'Indexation en France", dans "*L'indexation des Salaires*", *Economica*, pp. 134-160.
- BROWN R.L., J. DURBIN et J.M EVANS [1975] "Techniques for testing the Constancy of Regression Relationships Over Time", *Journal of the Royal Statistical Society*, Séries B, Vol n° 37, pp. 149-192.
- CHOW G.C. [1960] "Tests of Equality Between Sets of Coefficients in Two Linear Regressions", *Econometrica*, Vol. 28, pp. 591-605.
- CHOW G.C. [1983] "Random and Changing Coefficient Models", in *Handbook of Econometrics*, Vol. 2, Chap. 21, North Holland, pp. 1214-1243.
- ENGLE R.F. and C.W.J. GRANGER [1987] "Co-Integration and Error Correction : Representation, Estimation and Testing", *Econometrica*, Vol. 55, pp. 251-276.
- HARVEY A.C. [1981a] "*Time Series Models*", Philip Alan, pp. 101-120.
- HARVEY A.C. [1981b] "*The Econometric Analysis of Time Series*", Philip Alan, pp. 151-154.
- HSIAO C. [1986] "*Analysis of Panel Data*", Cambridge University Press.
- KALMAN R.E. [1960] "A new Approach to Linear Filtering and Predictions Problems", *Journal of Basic Engineering*, Series D, Vol. 82, pp. 35-45.
- KMENTA J. [1978] "Some problems of Inference from Economic Survey Data", *Survey Sampling and Measurement*, New York Academic Press, pp. 107-120.
- KMENTA J. [1990] "*Elements of Econometrics*", Second Edition, Macmillan, pp. 630-635.

- LAMBERT T. et J.L. MADRE [1989] "Prévisions à long terme du trafic automobile", *Collection des rapports du CREDOC*, n°60.
- MADDALA G.S. [1971] "The Use of Variance Components Models in Pooling Cross Section and Time Series Data", *Econometrica*, Vol. 39, pp. 341-358.
- MADRE J.L. et A. PIROTTE [1992] "Régionalisation des Projections à Long Terme du Trafic Automobile", *Rapport Convention SETRA*, à paraître.
- MAUREL F. [1985] "Filtre de Kalman et Econométrie", *Document de Travail n°319/930*, Département des Etudes Economiques d'Ensemble, INSEE.
- MAZODIER P. [1971] "L'estimation des Modèles à Erreurs Composées", *Annales de l'INSEE*, n°7, pp. 43- 72.
- MUNDLAK Y. [1978] "On the Pooling of Time Series and Cross Section Data", *Econometrica*, Vol. 46, pp. 69-85.
- NERLOVE M. [1971] "Experimental Evidence on the Estimation of Dynamic Economic Relations from a Time Series of Cross-Sections", *Economic Studies Quarterly*, Vol. 18, pp. 42-74.
- ROUSSE H. [1990] "Détection et Effets de la Multicolinéarité dans les Modèles Linéaires Ordinaires", *Document de Travail n°9002*, Département des Etudes Economiques d'Ensemble, INSEE.
- SEVESTRE P. et A. TROGNON [1992] "Linear Dynamic Models", Chapitre 5, in *The Econometrics of Panel Data : Theory and Applications*, L. Matyas and P. Sevestre Eds, Kluwer, à paraître.
- STAPLETON D.C. [1981] "Inferring Long-Term Substitution Possibilities from Cross-Section and Time-Series Data", *Modelling and Measuring Natural Resource Substitution*, Cambridge Mass. M.I.T. Press, pp. 93-118.
- WALLACE T.D. and A. HUSSAIN [1969] "The Use of Error Components Models in Combining Cross-Section with Time-Series Data", *Econometrica*, Vol. 37, pp. 55-72.