ANNALES D'ÉCONOMIE ET DE STATISTIQUE. – N° 20/21 — 1991

Conditions initiales et estimation efficace dans les modèles dynamiques sur données de panel : Une application au comportement d'investissement des entreprises

Richard BLUNDELL, Richard J. SMITH *

RÉSUMÉ. – Cet article fournit une méthodologie pour estimer de manière efficace des modèles dynamiques sur données de panel sous différentes hypothèses concernant les conditions initiales et les spécifications de la variance. Dans l'analyse des modèles sur données de panel qui incluent des variables dépendantes retardées, l'intérêt apporté aux conditions initiales reflète le fait que la plupart des fichiers de données de panel ont une dimension temporelle relativement faible. Notre spécification du modèle est un modèle dynamique linéaire au premier ordre dans lequel un sous ensemble de régresseurs strictement exogènes est supposé exister. Pour toutes les hypothèses concernant les conditions initiales sauf pour les plus triviales les MCG standards ne sont pas convergents. Nous montrons que notre approche fournit une procédure naturelle pour atteindre l'efficacité asymptotique et un cadre naturel pour tester les hypothèses concernant le processus sous jacent aux conditions initiales. La méthodologie est appliquée à l'estimation d'un modèle dynamique décrivant le comportement d'investissement et utilise des données de panel d'entreprises.

Initial Conditions and Efficient Estimation in Dynamic Panel Data Models: An Application to Company Investment Behaviour

ABSTRACT. – This paper provides a methodology for efficient estimation of dynamic panel data models under different assumptions concerning initial conditions and variance assumptions. In the analysis of panel data models that may involve lagged dependent variables, concerns about the initial conditions reflect the fact that most panel data sets have a relatively small time series dimension. Our model specification is a linear first-order dynamic model in which there are assumed to exist an identifying subset of strictly exogenous regressors. Under all but the most trivial assumption concerning the initial conditions the standard GLS procedure is inconsistent. We argue that our approach provides a natural procedure for testing assumptions concerning the process underlying the initial conditions. The methodology is applied to the estimation of a dynamic model of investment behaviour using panel data.

* R. BLUNDELL: University College London and Institute for Fiscal Studies; R. J. SMITH: Department of Applied Economics and Gonville and Caius College, University of Cambridge. Ce travail a été commencé au cours d'une invitation des deux auteurs à l'ENSAE (Paris) avril 1988. Nous sommes reconnaissants envers Stephen Bond, Christian Gouriéroux, Thierry Magnac, Patrick Sevestre, et Alain Trognon pour leurs commentaires et suggestions. Nous remercions le ESRC qui a financé ce projet sous le numéro : B0023 2150. Cet article a été traduit en français par A. BOUSQUET.

1 Introduction

Le but de cet article est de fournir une méthodologie pour estimer de manière efficace des modèles dynamiques sur données de panel sous différentes hypothèses concernant les conditions initiales. Notre intérêt pour cet aspect particulier de l'économétrie sur données de panel provient essentiellement des applications sur des panels relativement courts dans lesquels la première période dépend de la date à laquelle la firme ou le ménage est entré dans un nouvel état. Ainsi les conditions initiales doivent refléter un effet d'entrée ou de «start up». Dans ce cas, il ne semble pas raisonnable de supposer que le processus déterminant les observations initiales soit stationnaire et par conséquent les observations initiales doivent faire l'objet d'un traitement spécifique.

Dans l'analyse des modèles sur données de panel, qui incluent des variables dépendantes retardées, il est nécessaire d'étudier les conditions initiales. Ceci reflète le fait que la dimension temporelle des observations individuelles d'entreprises ou de ménages est généralement faible et les arguments asymptotiques reposent usuellement sur l'effet temporel moyen, calculé sur la dimension individuelle. Dans ce cas l'importance des conditions initiales ne disparaît pas asymptotiquement. Ainsi dans ce type de modèles de données de panel, la convergence des estimateurs des MCG des composantes de la variance dépend crucialement des hypothèses sur les propriétés stochastiques des observations initiales (voir ANDERSON et HSIAO [1982] et BHARGAVA et SARGAN [1983]).

Nous spécifions un modèle dynamique linéaire au premier ordre dans lequel nous supposons qu'un sous ensemble de régresseurs strictement exogènes existe. La spécification des erreurs est supposée suivre des généralisations de la forme standard de décomposition de la variance, dans laquelle le terme d'erreur se décompose additivement en un effet individuel spécifique invariant dans le temps et un processus aléatoire variable dans le temps. Notre objectif est de décrire une méthodologie pour l'estimation et les tests pour toute une série d'hypothèses concernant les spécifications stochastiques associées aux conditions initiales. Pour toutes les hypothèses concernant les conditions initiales sauf pour les plus triviales les MCG standards sont inconsistants.

Ce travail est très proche de celui de SEVESTRE et TROGNON [1985] et a été, dans une certaine mesure, stimulé par leur analyse. Par ailleurs, notre méthodologie et les estimateurs qui en résultent sont tout à fait différents sur plusieurs points. En particulier, nous suivons l'approche du maximum de vraisemblance conditionnel. Nous montrons que cette approche conduit à une procédure naturelle qui permet d'obtenir l'efficacité asymptotique et fournit un cadre de travail simple pour tester les différentes hypothèses du processus sous jacent aux conditions initiales.

Le plan de cet article est le suivant. Dans la section 2 nous traçons les grandes lignes du modèle et décrivons les différentes spécifications stochastiques qui seront analysées. La section 3 décrit alors en détail les différents estimateurs proposés. Dans la section 4 nous présentons une application à un modèle dynamique d'investissement des entreprises. Enfin, la section 5 conclut et donne un résumé des principaux résultats.

2 Spécification des erreurs et conditions initiales

Le modèle auquel on s'intéresse dans cet article est un modèle linéaire dynamique de données de panel, de forme générale :

$$y_{ht} = \alpha y_{ht-1} + \gamma' z_h + \beta' x_{ht} + u_{ht} \tag{1}$$

où $h = 1, \ldots, N$ et $t = 1, \ldots, T$. Dans (1) y_{ht} décrit complètement le comportement observé de l'individu h à l'instant t en fonction de sa valeur retardée y_{ht-1}, d'un vecteur de caractéristiques invariantes dans le temps z_h strictement exogènes et d'un vecteur x_{ht} de variables explicatives variables dans le temps. La condition initiale y_{h0} est décomposée sous la forme :

$$y_{h0} = \mu_{h0} + u_{h0}, \tag{2}$$

où $E(y_{h0}) = \mu_{h0}$ et u_{h0} est un terme d'erreur aléatoire, $h = 1, \ldots, N$.

Dans la spécification la plus couramment utilisée pour ce modèle. le terme d'erreur u_{ht} est supposé avoir la décomposition suivante :

$$u_{ht} = \eta_h + e_{ht}, \tag{3}$$

où η_h est un effet spécifique aux individus et invariant dans le temps, $h = 1, \ldots, N$, $t = 1, \ldots, T$. En général, bien que T soit fini, N est supposé être très grand. En suivant BHARGAVA et SARGAN [1983], cette spécification est couramment complétée par l'hypothèse de stationnarité $|\alpha| < 1$ et en écrivant :

$$u_{h0} = u^*_{h0} + \varepsilon_{h0}, \tag{4}$$

avec

$$u^*_{h0} = \sum_{k=0}^{\infty} \alpha^k v_{h,-k} \tag{5}$$

où $v_{ht} = \eta_h + e_{ht}$, $t = \ldots, -2, -1, 0$, est défini de manière similaire à u_{ht} dans (3).

Sachant η_h, chaque e_{ht} pour tout t et ε_{h0} sont mutuellement indépendants. La covariance entre u_{ht} et u_{h0} peut être exprimée par :

$$(6) \qquad E(u_{ht}u_{h0}) = \sigma_\eta^2/(1-\alpha), \qquad h=1,\ldots,N, t=1,\ldots T,$$

où σ_η^2, la variance de η_h, est pour le propos de cette discussion, supposée constante pour tout h. Sous une hypothèse similaire d'homoscédasticité de e_{ht} et ε_{h0}, nous avons également :

$$(7) \qquad \operatorname{var}(u_{h0}) = \sigma_e^2\left\{\frac{\rho^2}{(1-\alpha^2)} + \frac{1}{(1-\alpha)} + \frac{\sigma_\varepsilon^2}{\sigma_e^2}\right\},$$

où $\rho^2 = \sigma_\eta^2/\sigma_e^2$ et $\sigma_e^2, \sigma_\varepsilon^2$ sont respectivement les variances de e_{ht} et ε_{h0}; voir BHARGAVA et SARGAN [1983].

Il n'y a pas « *a priori* » de raisons pour que la spécification des erreurs pour y_{ht} puisse être reliée directement à y_{h0} sous la forme décrite par (3)-(7). Par conséquent, nous devrons travailler avec un certain nombre de spécifications alternatives pour les erreurs. Par exemple, $t=0$ peut indiquer le commencement d'un nouveau comportement, par exemple la première année de cotation d'une entreprise en bourse. Dans ce cas les décisions à la date 0 (et avant) peuvent être définies séparément par un processus stochastique représentant des effets de « start-up ». Il est quand même possible que u_{h0} et u_{ht} soient indépendants, dans ce cas les MCG standards ou l'estimateur des composantes de la variance seront convergentes et efficaces pour N grand. Pour cette raison nous introduisons trois spécifications alternatives des erreurs.

De manière à fournir un classement cohérent pour chaque spécification des erreurs, nous devons définir le vecteur $u_h = [u_{h1}, \ldots, u_{hT}]$ et sa matrice de variance-covariance en fonction de u_{h0} :

$$(8) \qquad E\left\{\begin{pmatrix} u_h \\ u_{h0}\end{pmatrix}(u_h', u_{h0})\right\} = \sigma_e^2\begin{bmatrix}\Omega & \omega_{10} \\ \omega_{10}' & \omega_{00}\end{bmatrix},$$

La première spécification des erreurs que nous considérons est celle où la matrice de variance-covariance entre u_{ht} et u_{h0} est non contrainte, bien que nous conserverons l'hypothèse d'homoscédasticité au travers de $h=1,\ldots,N$. Dans ce cas Ω, $\omega_{10}' = (\omega_{10}, \ldots, \omega_{T0})$ et ω_{00} sont non contraintes. Par exemple, il n'y a pas de raison pour que ω_{t0} soit constant pour $t=1,\ldots,T$. Cette spécification des erreurs, notée ES (i), est clairement assez faible. Ainsi, dans ce cas l'estimation en différence première est convergente.

La spécification des erreurs est particulièrement importante puisque si η_h dans (3) est non nul alors les estimatons des Moindres carrés ordinaires (MCO), Intra (I) et des Moindres Carrés Généralisés Contraints (MCGC) (ou à erreurs composées) de (1) sont biaisées même si $N \to \infty$ quand t est fini; voir NICKELL [1981]. Par ailleurs, comme SEVESTRE et TROGNON [1985] l'ont fait remarquer, puisque les estimateurs MCGC rendent la matrice de variance covariance diagonale pour chaque individu, ils se rapprochent de la convergence et peuvent être ajustés pour N grand. Le biais asymptotique des estimateurs MCGC dépend de manière cruciale de la corrélation entre les effets individuels spécifiques η_h et la condition initiale y_{h0}. Notre solution au problème de l'estimation efficace de (1) avec la spécification des erreurs (3) suit l'approche du maximum de vraisemblance conditionnel. Dans cette approche un régresseur supplémentaire est inclu qui permet de « déplacer » la dépendance de u_{ht} sur y_{h0}. Nous justifierons que cette approche fournit un estimateur convergent plus commode que ceux suggérés dans la littérature et à partir duquel l'efficacité asymptotique peut être obtenue directement par une procédure en deux étapes.

Pour la spécification suivante des erreurs, nous permettons deux possibilités non emboîtées: ES (*ii*) (*a*) dans laquelle la covariance entre u_{ht} et u_{h0} est invariante dans le temps: $\omega_{t0} = \omega_{10}$, $t = 1, \ldots T$, avec Ω non contrainte et ES (*ii*) (*b*) où l'on permet à Ω d'être déterminée via la formulation avec erreurs composées:

$$\Omega = \sigma_e^2 (\mathrm{I_T} + \rho^2 ll'), \tag{9}$$

où l est un vecteur à T composantes formé de un et σ_e^2 et ρ^2 ont été définis plus haut dans (7). Ils correspondent à la spécification des erreurs de (3). Dans les deux cas ω_{00} est non contraint.

Dans la troisième spécification des erreurs, ES (iii), nous imposons à la fois les hypothèses de non variabilité temporelle et les hypothèses d'erreurs composées de ES (ii) (*a*) et ES (ii) (*b*) donnant ω_{00} non contraint. La dernière spécification des erreurs ES (iv) impose l'hypothèse de stationarité (4)-(7) sur ES (iii) ce qui donne le modèle dynamique standard à effet aléatoire discuté au début de cette section.

Pour générer notre modéle *conditionnel* nous décomposons le processus des erreurs u_{ht} comme:

$$u_{ht} = \tau_t u_{h0} + v_{ht}, \tag{10}$$

où

$$\tau_t = \omega_{t0}/\omega_{00}, \tag{11}$$

et donc $\mathrm{E}(v_{ht} \mid u_{h0}) = 0$. La relation entre la variance des erreurs dans le modèle conditionnel et (8) est alors donnée par:

$$\mathrm{E}(v_{hs} v_{ht}) = \omega_{st} - \omega_{s0}\, \omega_{t0}/\omega_{00}, \tag{12}$$

pour s, $t = 1, \ldots, \mathrm{T}$ où ω_{st} est le (s, t) élément de Ω. Ceci nous permet d'écrire (1) comme:

$$y_{ht} = \alpha y_{h\,t-1} + \gamma' z_h + \beta' x_{ht} + \tau_t u_{h\,0} + v_{ht}, \tag{13}$$

où maintenant chaque u_{h0}, pour tout z_h et x_{ht}, est orthogonal à v_{ht}, $h = 1, \ldots, \mathrm{N}$, $t = 1, \ldots, \mathrm{T}$.

Le modèle conditionnel (13) est complété en définissant une spécification de la moyenne de y_{h0}: (2) ci-dessus. Dans ce qui suit nous distinguerons deux possibles spécifications pour la moyenne MS (i):

$$\mu_{h0} = \psi' z_{h0}, \tag{14}$$

où z_{h0} représente un vecteur de variables instrumentales strictement exogènes, $h=1, \ldots, N$; voir HAUSMAN et TAYLOR [1981], AMEMIYA et MACURDY [1986] et BREUSCH, MIZON et SCHMIDT [1989] pour différentes hypothèses d'exogénéité, et MS(ii):

$$\mu_{h0}=\mu_0, \tag{15}$$

où μ_{h0} est simplement supposé constant pour tout $h=1, \ldots, N$.

Notre méthodologie pour l'estimation développée formellement dans la section suivante est définie de la manière suivante. Premièrement nous construisons $\dot{u}_{h0}$ en utilisant:

$$\dot{u}_{h0}=y_{h0}-\dot{\psi}' z_{h0}, \tag{16}$$

sous MS(i), où $\dot{\psi}$ est l'estimateur des MCO de ψ, et sous MS(ii) cette expression ce réduit simplement à:

$$\dot{u}_{h0}=y_{h0}-\bar{y}_0, \tag{17}$$

où $\bar{y}_0$ est la moyenne en coupe de y_{h0}, $h=1, \ldots, N$. Ensuite, sachant $\dot{u}_{h0}$, nous suivons BLUNDELL et SMITH [1989] et SMITH et BLUNDELL [1986] et estimons le modèle conditionnel (13) en remplaçant u_{h0} par $\dot{u}_{h0}$. Bien que $\dot{\psi}$ soit en général convergent mais non efficace, la distribution de l'estimateur du modèle conditionnel sur un grand échantillon est indépendante de $\dot{\psi}$. Ainsi, sous l'hypothèse de normalité, cette procédure est une estimation du maximum de vraisemblance conditionnel et est pleinement efficace. La méthode d'estimation optimale dépend de la spécification choisie pour les erreurs et pour cette discussion nous renvoyons à la section 3.

3 Estimation et test

Premièrement, dans le contexte du modèle conditionnel (13), nous considérons le modèle donné par ES(i) qui conduit à la formulation la moins restrictive pour la variance:

$$\Omega_* = \Omega - \omega_{10}\,\omega_{10}'/\omega_{00},$$
$$\tau = \omega_{10}/\omega_{00},$$

où $\Omega_* = E(v_h v_h')$ et $v_h = (v_{h1}, \ldots, v_{hT})'$, $h=1, \ldots, N$. [1]

1. Nous avons explicitement reparamétré le modèle par une transformation bijective: $\Omega \to \Omega_*$ et $\omega_{10} \to \tau$.

Le modèle *conditionnel* qui en résulte, dérivé des équations (1) et (10) peut être écrit sous forme vectorielle plus commode en empilant les observations :

$$\text{(18)} \qquad \mathrm{B}\, y_h - \alpha e_1 y_{h0} - (l \otimes \gamma') z_h - (\mathrm{I_T} \otimes \beta') x_h - \tau u_{h0} = v_h,$$

$$\text{(19)} \qquad y_{h0} - \mu_{h0} = u_{h0}, \qquad h = 1, \ldots, \mathrm{N},$$

où $y_h = (y_{h1}, \ldots, y_{hT})'$, $x_h = (x_{h1}, \ldots, x_{hT})'$, e_1 est le premier vecteur de dimension T de la base canonique, B une matrice $(\mathrm{T} \times \mathrm{T})$ avec des 1 sur la diagonale et $-\alpha$ sur la première diagonale principale inférieure. Sous MS (i) $\mu_{h0} = \psi' z_{h0}$ alors que sous MS (ii) $\mu_{h0} = \mu_0$. Par la suite nous supposons MS (i). [2]

Par commodité nous réécrivons (18) sous la forme :

$$\text{(20)} \qquad \begin{cases} \Delta d_h = v_h, \qquad h = 1, \ldots, \mathrm{N}, \\ \qquad \Delta \mathrm{D}' = \mathrm{V}', \end{cases}$$

où $d_h = (y_h', y_{h0}, z_h', x_h', u_{h0})'$, $h = 1, \ldots, \mathrm{N}$, $\Delta = (\mathrm{B}, -\alpha e_1, -(l \otimes \gamma'), -(\mathrm{I_T} \otimes \beta'), -\tau)$, $\mathrm{D} = (d_1, \ldots, d_\mathrm{N})'$, $\mathrm{V} = (v_1, \ldots, v_\mathrm{N})'$.

3.1. Estimation efficace

Si μ_{h0} était connu, alors comme la matrice des paramètres concernés par les variables endogènes de (18), B est une matrice triangulaire inférieure, l'estimation par MCGC serait donc convergente; voir Lahiri et Schmidt [1978]. Quand μ_{h0} est inconnu la procédure que nous suggérons consiste premièrement à estimer (19) par les MCO puis à utiliser les résidus des MCO $\dot{u}_{h0}$ à la place du terme d'erreur u_{h0} dans (18).

Nous nous intéressons tout d'abord à l'estimateur des MCG conditionnel et contraint (MCGCC) donné par :

$$\hat{\delta}_{\mathrm{mcgcc}} = \{ \mathrm{S}' (\dot{\Omega}_*^{-1} \otimes \dot{\mathrm{D}}' \dot{\mathrm{D}}) \mathrm{S} \}^{-1} \mathrm{S}' (\dot{\Omega}_*^{-1} \otimes \dot{\mathrm{D}}' \dot{\mathrm{D}}) c,$$

où $\dot{\mathrm{D}} = (\dot{d}_1, \ldots, \dot{d}_\mathrm{N})'$, $\dot{d}_h = (y_h', y_{h0}, z_h', x_h', \dot{u}_{h0})'$, $\mathrm{vec}\, \Delta' = c - \mathrm{S}\delta$, $c = \mathrm{vec} \{ \mathrm{I_T}, 0 \}'$, $\mathrm{S} = -\partial \mathrm{vec}\, \Delta / \partial \delta'$, et $\delta = (\alpha, \gamma', \beta', \tau')'$ et en utilisant $\hat{\mathrm{V}}' \hat{\mathrm{V}} / \mathrm{N}$ pour estimer $\dot{\Omega}$. $\hat{\mathrm{V}}$ sont les résidus de l'estimation de (20) par la méthode des variables instrumentales. Cette procédure fournit des estimateurs convergents et donc $\hat{\delta}_{\mathrm{mcgcc}}$ est la solution des conditions du premier ordre pour l'estimation du maximum de vraisemblance conditionnel de δ après substitution de Ω; l'estimateur FIML est obtenu en itérant sur $\dot{\Omega}$. Par ailleurs, l'efficacité asymptotique n'est pas obtenue pour $\hat{\delta}_{\mathrm{mcgcc}}$ puisque la matrice d'information n'est pas bloc diagonale entre δ et Ω (voir Lahiri et Schmidt [1978]).

L'estimation efficace de δ est possible en deux étapes : en estimant (18) et en utilisant le Système à Variables Instrumentales (SVI) pour obtenir

2. Nous pouvons traiter MS (ii) comme un cas particulier de MS (i) obtenu par l'exclusion des restrictions sur ψ. Par ailleurs, bien que les résultats décrits ci-dessous soient valides, la procédure n'est plus asymptotiquement efficace; voir Pesaran et Smith [1990].

l'estimateur SVI conditionnel contraint (SVICC) :

$$\hat{\delta}_{\text{svicc}}=\{S'(\dot{\Omega}_*^{-1}\otimes\dot{D}'P_{Z^*}\dot{D})S\}^{-1}S'(\dot{\Omega}_*^{-1}\otimes\dot{D}'P_{Z^*}\dot{D})c,$$

où P_{Z^*} est la matrice de projection orthogonale sur l'espace définit par $Z_*=(Z_{*1},\ldots,Z_{*N})'$, $Z_{*h}=(Z'_{h0},y_{h0})'$, $h=1,\ldots,N$. Des estimations consistantes des paramètres initiaux de (18) et donc de Ω_* peuvent être obtenues par la méthode des variables instrumentales contraintes (VIC) en utilisant Z_{h0} et y_{h0} comme instruments :

$$\hat{\delta}_{\text{vic}}=\{S'(I_T\otimes\dot{D}'P_{Z^*}\dot{D})S\}^{-1}S'(I_T\otimes\dot{D}'P_{Z^*}\dot{D})c,$$

L'estimation SVICC de δ est pleinement efficace comme cela est montré par le lemme suivant, (SMITH et BLUNDELL [1986] et aussi TROGNON [1988])

LEMME 1 : L'estimateur SVICC, $\hat{\delta}_{\text{svicc}}$, est asymptotiquement équivalent à l'estimateur par Maximum de Vraisemblance (MV) pour δ.

Preuve. — Nous avons seulement besoin de montrer que la distribution limite de $N^{1/2}(\hat{\delta}_{\text{svicc}}-\delta)$ est indépendante de celle de $N^{1/2}(\dot{\psi}-\psi)$. On a :

$$N^{1/2}(\hat{\delta}_{\text{svicc}}-\delta)=\{S'(\dot{\Omega}_*^{-1}\otimes\dot{D}'P_{Z^*}\dot{D}/N)S\}^{-1}$$
$$S'(\dot{\Omega}_*^{-1}\otimes\dot{D}'P_{z^*}\dot{D}/N^{1/2})\,\mathrm{vec}(\Delta').$$

Le dernier terme peut être écrit sous la forme :

$$S'\,\mathrm{vec}\{\dot{D}'P_{z^*}\dot{D}\Delta'\Omega_*^{-1}/N^{1/2}\}.$$

Considérons le terme $\dot{D}'P_{z^*}\dot{D}\Delta'/N^{1/2}$:

$$\begin{aligned}N^{-1/2}\{\dot{D}'P_{z^*}\dot{D}-D'P_{z^*}\dot{D}\}\Delta&=-2N^{-1/2}\begin{pmatrix}0\\ \cdots\cdots\\ (\dot{\psi}-\psi)'Z'\end{pmatrix}D\Delta'+0_p(1)\\&=-2\begin{pmatrix}0\\ \cdots\cdots\cdots\\ (\dot{\psi}-\psi)'Z'V/N^{1/2}\end{pmatrix}+0_p(1)\\&=0_p(1),\end{aligned}$$

où $Z=(z'_{10},\ldots,z'_{N0})'$. De plus comme $\dot{\Omega}_*-\Omega_*=0_p(1)$, on a :

$$N^{1/2}(\hat{\delta}_{\text{svicc}}-\delta)\xrightarrow{L}N(0,\{S'(\Omega_*^{-1}\otimes p\lim(D'P_{z^*}D/N))S\}^{-1});$$

ce qui signifie que l'estimation de ψ n'a pas d'effet sur la distribution limite de l'estimateur SVICC et que par conséquent $\hat{\delta}_{\text{svicc}}$ est asymptotiquement efficace. □

Nous considérons à présent l'estimation efficace de Ω_* sous ES(i) :

THÉORÈME 2 : Un estimateur asymptotiquement efficace de Ω_* est donné par :

$$\hat{\Omega}_* = \hat{V}' \hat{V}/N, \tag{21}$$

où $\hat{V}' = \dot{D}\hat{\Delta}'$, $\operatorname{vec}\hat{\Delta}' = c - S\,\hat{\delta}_{svic}$.

Preuve: semblable au lemme (1). □

Le théorème 2 permet d'obtenir un estimateur simple et efficace pour ψ. Considérons maintenant les équations de vraisemblance du premier ordre associée à ψ :

$$Z' \left\{ \frac{u_0}{\omega_{00}} - V\Omega_*^{-1}\tau \right\}, \tag{22}$$

où $u_0 = (u_{10}, \ldots, u_{N0})'$. Le vecteur des paramètres ψ intervient à la fois dans u_0 et V. En substituant δ par son estimation efficace $\hat{\delta}_{svicc}$, $\dot{\omega}_{00} = \dot{u}_0' \dot{u}_0/N$ pour ω_{00} dans (21) et en résolvant on obtient :

$$\hat{\psi} = (Z'Z)^{-1} Z' \{ y_0 - \hat{U}\hat{\Omega}_*^{-1}\hat{\tau}\dot{\omega}_{00}/(1 + \hat{\tau}'\hat{\Omega}_*^{-1}\hat{\tau}\dot{\omega}_{00}) \}, \tag{23}$$

où $\hat{U} = (\hat{u}_1, \ldots, \hat{u}_N)'$, $\hat{u}_h = (\hat{u}_{h1}, \ldots, \hat{u}_{hT})'$, $\hat{u}_{hT} = y_{ht} - \hat{\alpha}\ y_{ht-1} - \hat{\gamma}'\ Z_h - \hat{\beta}' x_{ht}$, $t = 1, \ldots, T$ et $\hat{\delta}_{svicc} = (\hat{\alpha}, \hat{\gamma}', \hat{\beta}', \hat{\tau}')'$.

COROLLAIRE 3 : L'estimateur $\hat{\psi}$ de (23) est un estimateur asymptotiquement efficace de ψ.

Preuve : Ce résultat se déduit directement puisque $\hat{\delta}_{svicc}$ et $\hat{\Omega}_*^{-1}$ sont tous deux asymptotiquement efficaces d'après respectivement le lemme 1 et le théorme 2. De plus, comme $p\lim(Z'U/N) = 0$, la distribution asymptotique de $N^{1/2}(\hat{\psi} - \psi)$ est indépendante de celle de $N^{1/2}(\dot{\omega}_{00} - \omega_{00})$. □

Par conséquent, un estimateur asymptotiquement efficace pour ω_{00} peut être construit par :

$$\hat{\omega}_{00} = N^{-1}\hat{u}_0' \hat{u}_0, \tag{24}$$

où $\hat{u}_0 = (\hat{u}_{10}, \ldots, \hat{u}_{N0})'$ et $\hat{u}_{h0} = y_{h0} - \hat{\psi}' z_{h0}$, $h = 1, \ldots, N$.

3.2. Les modèles contraints

Ayant obtenu des estimateurs efficaces des paramètres inconnus pour la spécification des erreurs la moins contrainte ES(i), nous considérons à présent l'estimation et les tests pour la spécification des erreurs la plus restrictive discutée dans la section 2. Premièrement, l'hypothèse implicite contenue dans (18) est celle de la stabilité des coefficients structurels associés aux y_{ht-1}, z_h et x_{ht}, sur toute la période. Étant donné des instruments suffisants z_{h0}, $h = 1, \ldots, N$, une version non contrainte du modèle conditionnel (18) pourra être estimée par SVI et permettra de tester cette stabilité par un test du rapport de vraisemblance.

Pour tester l'hypothèse d'invariance-temporelle ES(ii) (*a*):

$$\mathrm{H}_0^i : \tau_t = \tau, \qquad t=1, \ldots, \mathrm{T},$$

peut être testée plus facilement à l'aide d'un test de Wald basé sur la distribution asymptotique de $\hat{\delta}_{svicc}$; cette statistique a pour distribution limite une loi du χ^2 à $(\mathrm{T}-1)$ degrés de liberté sous H_0^i. Si H_0^i est acceptée, alors une nouvelle estimation contrainte SVI du modèle conditionnel incorporant H_0^i doit être réalisée, estimation qui, en utilisant les mêmes arguments que ceux employés dans le lemme 1, est asymptotiquement efficace. De plus, tester H_0^i est un précurseur naturel pour tester l'exogénéité de y_{h0}:

$$\mathrm{H}_0^e : \tau = 0 \text{ contre } \mathrm{H}_1^e : \tau \neq 0.$$

Ceci pourra être fait soit directement par un test de Wald de $\mathrm{H}_0^i \cap \mathrm{H}_0^e$ après SVICC sous H_0^i, ou par la différence des statistiques de Wald pour tester $\mathrm{H}_0^i \cap \mathrm{H}_0^e$ et H_0^i respectivement; voir AITCHISON [1962] et SARGAN [1980]. Ces deux statistiques ont un χ^2 à un degré de liberté pour distribution limite sous $\mathrm{H}_0^i \cap \mathrm{H}_0^e$. Évidemment, si $\mathrm{H}_0^i \cap \mathrm{H}_0^e$ était acceptée alors l'estimation de (1) par SVIC est efficace.

Pour la spécification des erreurs ES(ii) (*b*), avec la forme décomposée des erreurs pour les observations $t=1, \ldots, \mathrm{T}$ on a:

$$\mathrm{H}_0^{ec} : \Omega_* = \sigma_e^2 (\mathrm{I}_\mathrm{T} + \rho^2 ll') - \omega_{00} \tau\tau',$$

où $\rho^2 = \sigma_\eta^2 / \sigma_e^2$, de simples estimations initiales et consistantes pour σ_e^2 et ρ^2 peuvent être obtenues par:

$$\sigma_e^2 = \frac{1}{(\mathrm{T}-1)} \left\{ \mathrm{tr}\{ \mathrm{A}\hat{\Omega}_* \} + \hat{\omega}_{00} \hat{\tau}' \mathrm{A} \hat{\tau} \right\}, \tag{25}$$

qui est semi-définie positive, où $\mathrm{A} = \mathrm{I}_\mathrm{T} - ll'/\mathrm{T}$, et

$$\rho^2 = \frac{1}{\mathrm{T}\dot{\sigma}_e^2} \left\{ \mathrm{tr}\, \hat{\Omega}_* + \hat{\omega}_{00} \hat{\tau}' \hat{\tau} - \mathrm{T}\dot{\sigma}_e^2 \right\}. \tag{26}$$

Les estimateurs donnés dans (25) et (26) peuvent être utilisés comme valeurs initiales pour mieux localiser les estimateurs efficaces des paramètres inconnus dans ES(ii) (*b*). La spécification des erreurs ES(ii) (*b*) peut être testée contre la formulation non contrainte ES(i) au moyen d'un test de Wald, voir CHAMBERLAIN [1982] et SZROETER [1983]. Considérons la fonction critère suivante:

$$\begin{aligned} \min_{\sigma_e^2, \rho^2} \{ \mathrm{N}\, v\{ \hat{\Omega}_* + \hat{\omega}_{00} \hat{\tau}\hat{\tau}' - \sigma_e^2 (\mathrm{I}_\mathrm{T} + \rho^2 ll') \}' \\ \times \{ \mathrm{av\hat{a}r}\, (\mathrm{N}^{1/2} v\{ \hat{\Omega}_* + \hat{\omega}_{00} \hat{\tau}\hat{\tau}' - \sigma_e^2 (\mathrm{I}_\mathrm{T} + \rho^2 ll') \}) \}^{-1} \\ \times v\{ \hat{\Omega}_* + \hat{\omega}_{00} \hat{\tau}\hat{\tau}' - \sigma_e^2 (\mathrm{I}_T + \rho^2 ll') \} \}, \end{aligned} \tag{27}$$

où v est l'opérateur sélectionnant dans une matrice définie positive symétrique le vecteur des éléments distincts deux à deux. La valeur minimale de (27) fournit un test asymptotiquement optimal pour tester ES(ii) (*b*) contre

ES(i). Sous H_0^{ec}, cette statistique de Wald aura pour distribution limite une loi du χ^2 avec $T(T-1)/2-2$ degrés de liberté.

Considérons la spécificatin des erreurs ES(iii) qui contient à la fois les hypothèses H_0^i et H_0^{ec} d'invariance temporelle et d'effet aléatoire :

$$H_0^{iec} : \Omega_* = \sigma_e^2 (I_T + \rho_0^2 ll'),$$

où $\rho_0^2 = \rho^2 - \omega_{00} \tau^2/\sigma_e^2$; qui est $H_0^{ice} = H_0^i = H_0^i \cap H_0^{ec}$. Cette spécification des erreurs implique la ré-paramétrisation bijective de $\rho^2 \to \rho_0^2$. Par conséquent, nous utiliserons des résultats analogues à ceux détaillés dans le lemme 1 et dans le théorème 2. En particulier, en utilisant les erreurs composées, la procédure FIML sur le modèle conditionnel (18) après estimation de ψ par MCO sera asymptotiquement efficace :

LEMME 4 : FIML Conditionnel avec Erreurs Composées (FIMLECC) sur (18) après estimation de ψ par les estimateurs MCO, $\dot{\psi}$ est asymptotiquement optimal.

Preuve : Comme dans la preuve du lemme 1, nous devons montrer l'indépendance de la distribution limite de l'estimateur FIMLECC $\hat{\delta}_{\text{fimlecc}}$ avec $N^{1/2}(\dot{\psi}-\psi)$. □

L'hypothèse ES(iii) où H_0^{iec} est simplement testée contre la spécification ES(i) où la variance est non contrainte, au moyen d'un test du rapport de vraisemblance comparant les valeurs de la log-vraisemblance pour le modèle conditionnel sous H_0^{iec} et ES(i) après substitution de $\dot{\psi}$. Cette statistique aura pour distribution limite une loi du χ^2 à $(T+1)(T-1)/2-2$ degrés de liberté sous H_0^{iec}. La comparaison de H_0^{iec} avec H_0^i est enfin réalisée par le test du rapport de vraisemblance sur le modèle conditionnel (10) après substitution de $\dot{\psi}$; cette statistique a pour distribution limite une loi du χ^2 avec $T(T-1)/2-2$ degrés de liberté sous H_0^{iec}. Pour tester H_0^{iec} contre H_0^{ec} un test du rapport de vraisemblance ou un test de Wald semblent appropriés, l'un et l'autre sont asymptotiquement distribués comme un χ^2 à $(T-1)$ degrés de liberté sous H_0^{iec}.

Finalement nous revenons au modèle avec la spécification des erreurs la plus contraignante, l'hypothèse de stationarité ES(iv) :

$$H_0^s : \rho_0^2 = \omega_{00}\, \tau\, [(1-\alpha)-\tau],$$

qui est réécrit en terme de la re-paramétrisation donnée ci-dessus et implique la transformation $\sigma_\varepsilon^2 \to \omega_{00}$; voir (6), (7) et (9). Comme pour ES(ii) ou H_0^{iec}, un test de Wald fournit un test simple de $H_0^{siec} = H_0^{iec} \cap H_0^s$ contre H_0^{iec} qui est distribué asymptotiquement comme un χ^2 à 1 degré de liberté sous H_0^{iec} :

$$\frac{N\{\tilde{\rho}_0^2 - \tilde{\omega}_{00}\tilde{\tau}[(1-\tilde{\alpha})-\tilde{\tau}]\}^2}{\tilde{\text{avar}}[N^{1/2}\{\tilde{\rho}_0^2 - \tilde{\omega}_{00}\tilde{\tau}[(1-\tilde{\alpha})-\tilde{\tau}]\}]},$$

où $\tilde{\rho}_0^2$, $\tilde{\tau}$ et $\tilde{\alpha}$ sont les estimateurs FIMLECC et $\tilde{\omega}_{00}$ est obtenu de la même manière que $\hat{\omega}_{00}$, précisément par :

$$\tilde{\omega}_{00} = N^{-1}\tilde{u}_0'\tilde{u}_0,$$

où $\tilde{\cdot}$ désigne pour l'estimation FIMLECC, l'estimateur efficace de $\tilde{\psi}$ sous H_0^{iec} :

$$\tilde{\psi} = (Z'Z)^{-1}Z'\{y_0 - \tilde{\tau}\tilde{U}\tilde{\Omega}_*^{-1} l\dot{\omega}_{00}/(1+\tilde{\tau}^2 l'\tilde{\Omega}_*^{-1} l\dot{\omega}_{00})\}.$$

4 Une application a un panel d'entreprises

Pour notre illustration empirique nous considérons la modélisation du comportement d'investissement des entreprises en utilisant un panel d'entreprises « côtées ». La date d'entrée de chaque entreprise dans le panel dépend de l'année où elle a été cotée sur le marché boursier britannique. Au total nous avons 532 firmes couvrant la période 1975-1986. Le choix du modèle d'investissement est motivé par les spécifications analysées dans BLUNDELL, BOND, DEVEREUX et SCIANTARELLI [1988] où différentes versions du modèle Q-TOBIN ont été considérées pour des entreprises du secteur industriel de Grande Bretagne. Cohérente avec les données, la spécification retenue pour le modèle fait intervenir dans la régression, à la fois, la variable Q courante et retardée (Q étant une mesure de la valeur boursière du capital de la firme à son coût de remplacement) et le taux d'investissement retardé. D'autres variables importantes ont été trouvées comme des variables indicatrices invariantes dans la dimension temporelle et spécifiques aux secteurs industriels.

Dans la spécification (1) du modèle dynamique sur données de panel, y_{ht} représente maintenant le taux d'investissement (I_{ht}) de la i-ème entreprise au cours de l'année t et x_{ht} comprend les variables explicatives Q_{ht}.Les z_h, sont des variables indicatrices invariantes dans le temps et spécifiques aux secteurs. Clairement y_{ht} pour plusieurs entreprises dépend du processus stochastique définissant l'entrée sur le marché boursier. Pour cette raison, on s'attend aux spécifications ES (i) ou ES (ii) des erreurs plutôt qu'au modèle restrictif ES (iv). Ainsi, ce sera bien le cas pour lequel ω_{10} est proche de zéro et où les MCG standards sont convergents.

La première colonne du tableau 1 présente les estimations par MCO, et montre que les coefficients associés à Q et à sa valeur retardée sont fortement significatifs. Par ailleurs, on peut s'attendre à ce que le coefficient sur la variable retardée soit biaisé vers le bas. Ceci est confirmé dans la deuxième colonne dans laquelle sont fournis les résultats par MCG pour le modèle à erreurs composées. Comme cela était attendu il y a une importante modification du paramètre associé à la variable retardée; voir SEVESTRE et TROGNON [1985]. Dans la troisième colonne, les estimations Intra confirment le biais négatif pour « T fini » sur la variable dépendante limitée (voir NICKELL [1981]).

En considérant les estimateurs du système de variables instrumentales introduits dans les sections 2 et 3, la quatrième colonne FIMLECC du tableau 1 contient les estimations pour la spécification des erreurs ES (iii) dans laquelle il y a une corrélation constante ω_{10} pour chaque période. Ce coefficient est positif, petit, et significatif. Il réduit peu le coefficient associé à la variable dépendante retardée. Finalement la colonne intitulée SVICC contient les estimations conditionnelles pour ES (i) et montre une nouvelle

TABLEAU 1

Estimations de l'équation d'investissement

	MCO	MCG	I	FIMLECC	SVICC
Q_{ht}	0.0077	0.0076	0.0081	0.0081	0.0085
	(7.42)	(7.25)	(7.59)	(7.21)	(7.33)
Q_{ht-1}	−0.0025	−0.0016	−0.0002	−0.0018	−0.0018
	(−2.83)	(−1.75)	(−0.20)	(−1.77)	(−1.82)
I_{ht-1}	0.4387	0.3373	0.1957	0.3101	0.2931
	(16.5)	(12.2)	(6.69)	(10.7)	(10.10)
τ	–	–	–	0.0621	–
				(2.34)	–
τ_1	–	–	–	–	−0.0232
					(−0.24)
τ_2	–	–	–	–	0.1256
					(2.30)
τ_3	–	–	–	–	0.1339
					(2.97)
τ_4	–	–	–	–	0.1749
					(2.45)
τ_5	–	–	–	–	0.1395
					(2.72)
τ_6	–	–	–	–	0.0437
					(0.89)
τ_7	–	–	–	–	0.0879
					(1.86)
τ_8	–	–	–	–	0.0637
					(1.27)
τ_9	–	–	–	–	0.5316
					(1.54)
τ_{10}	–		–	–	0.0337
					(0.92)
τ_{11}	–	–	–	–	−0.0149
					(−0.33)

Les variables indicatrices par secteurs et par périodes sont inclues dans toutes ces estimations (obtenues à l'aide de Gauss 386).

baisse du coefficient associé aux variables dépendantes retardées ainsi qu'une variation significative pour tous les éléments τ_t.

5 Conclusion

Cet article propose une procédure simple pour obtenir des estimateurs adaptés pour des modèles dynamiques linéaires sur données de panel discutés par *inter alia* ANDERSON et HSIAO [1981] et BHARGAVA et SARGAN [1983]. La procédure est basée sur un système de variables instrumentales conditionnel qui est asymptotiquement efficace sous des hypothèses relativement faibles.

Une série d'estimateurs pour différentes spécifications stochastiques plus ou moins restrictives peut être aisément obtenue. Une stratégie de test pour ces différentes formulations de la variance, basée sur une séquence de test du rapport de vraisemblance ou de test de Wald, est proposée. Les techniques présentées sont illustrées par une application à une équation de dépense d'investissement estimée sur un panel d'entreprises de Grande-Bretagne.

• Références bibliographiques

AITCHISON, J., (1962). – « Large-Sample Restricted Parametric Tests », *Journal of the Royal Statistical Society* (Series B), 24, pp. 234-250.

AMEMIYA, T. et MACURDY, T. E. (1986). – « Instrumental-Variable Estimation of Error-Components Models », *Econometrica*, 54, pp. 869-881.

ANDERSON, T. W. et HSIAO, C. (1982). – « Formulation and Estimation of Dynamic Models using Panel Data », *Journal of Econometrics*, 18, pp. 570-606.

ARELLANO, M. et BOND, S. (1988). – « Dynamic Panel Data Estimation », *IFS Working Paper*, 88-16.

BHARGAVA, A. et SARGAN, J. D. (1983). – « Estimating Dynamic Random Effects Models from Panel Data Covering Short Time Periods », *Econometrica*, 51, pp. 1635-1659.

BLUNDELL, R. W., BOND, S., DEVEREUX, M. et SCIANTARELLI, F. (1988). – « Does Q Matter for Investment? Some Evidence from a Panel of UK Companies », IFS Working Paper, 88-12, forthcoming in *Journal of Econometrics*.

BLUNDELL, R. W. et SMITH, R. J. (1989). – « Estimation in a Class of Simultaneous Equations Limited Dependent Variable Models », *Review of Economic Studies*, 56, 37-58.

BREUSCH, T. S., MIZON, G. E. et SCHMIDT, P. (1989). – « Efficient Estimation using Panel Data », *Econometrica*, 57, pp. 695-700.

CHAMBERLAIN, G. (1982), « Multivariate Regression Models for Panel Data », *Journal of Econometrics*, 18, pp. 5-46.

CHAMBERLAIN, G. (1978). – « Omitted Variables Bias in Panel Data: Estimating the Returns to Schooling », *Annales de l'INSEE*, 30/31, pp. 49-82.

CHAMBERLAIN, G. et GRILICHES, Z. (1975). – « Unobservables with a Variance Components Structure: Ability, Schooling and the Econometric Issues of Brothers », *International Econometric Review*, 16, pp. 422-449.

CHAMBERLAIN, G. (1984). – « Panel Data », in *Handbook of Econometrics*, (Z. GRILICHES and M. INTRILIGATOR (eds)), North Holland.

HAUSMAN, J. (1978). – « Specification Tests in Econometrics », *Econometrica*, 46, pp. 1251-1272.

HAUSMAN, J. et TAYLOR, W. E. (1981). – « Panel Data and Unobservable Individual Effects », *Econometrica*, pp. 1377-1398.

HAYASHI, F. (1982). – « Tobin's Marginal q and Average q: A Neoclassical Interpretation », *Econometrica*, 50.

HOCH, I. (1955). – « Estimation of Production Function Parameters and Testing », *Econometrica*, 23, p. 235.

HSIAO, C. (1985). – « Analysis of Panel Data », Cambridge University Press, Cambridge.

LAHIRI, K. et SCHMIDT, P. (1978). – «On the Estimation of Triangular Structural Systems», *Econometrica*, 46, pp. 1217-1221.

MADDALA, G. S. (1971). – «The Use of Variance Components Models in Pooling Cross Section and Time Series Data», *Econometrica*, 39, pp. 341-358.

MUNDLACK, Y. (1961). – «Empirical Production Free of Management Bias», *Journal of Farm Economics*, 43, pp. 44-56.

MUNDLACK, Y. (1978). – «On the Pooling of Time Series and Cross Section Data», *Econometrica*, 46, pp. 69-86.

NERLOVE, M. (1979). – «A Note On Error Components Models», *Econometrica*, 39, pp. 359-382.

NICKELL, S. J. (1981). – «Biases in Dynamic Models with Fixed Effects», *Econometrica*, pp. 1417-1428.

PESARAN, M. H. et SMITH, R. J. (1990). – «A Unified Approach to Estimation and Orthogonality Tests in Linear Single Equation Econometric Models», *Journal of econometrics*, 44, pp. 41-66.

SARGAN, J. D. (1980). – «Some tests of Dynamic Specification for a single Equation», *Econometrica*, 48, pp. 879-897.

SEVESTRE, P. et TROGNON, A. (1985). – «A Note on Autoregressive Error Components Models», *Journal of Econometrics*, 28, pp. 231-245.

SMITH, R. J. et BLUNDELL, R. W. (1986). – «An Exogeneity Test for The Simultaneous Equations Tobit Model», *Econometrica*, 54, pp. 679-685.

SZROETER, J. (1983). – «Generalised Wald Methods for Testing Nonlinear Implicit Overidentifying Restrictions», *Econometrica*, 51, pp. 335-353.

TROGNON, A. (1988). – «Efficiency of Two Steps Procedures: the Case of Quasi-Generalised M-Estimators», Article présenté à ESEM 1988, Bologne, Italie.