REVUE DE STATISTIQUE APPLIQUÉE

J.-P. MARCIANO

Sur le biais introduit par les estimateurs de la *k*-classe dans l'estimation des modèles linéaires interdépendants

Revue de statistique appliquée, tome 24, n° 3 (1976), p. 71-90

<http://www.numdam.org/item?id=RSA_1976__24_3_71_0>

SUR LE BIAIS INTRODUIT PAR LES ESTIMATEURS DE LA K-CLASSE DANS L'ESTIMATION DES MODÈLES LINÉAIRES INTERDÉPENDANTS (1)

J.-P. MARCIANO
Directeur du Laboratoire de mathématiques appliquées
de la Faculté d'économie appliquée de l'université d'Aix-Marseille

SOMMAIRE

Après un préliminaire sur les systèmes économétriques interdépendants, une première partie actualise des résultats sur l'existence des moments des estimateurs de la k-classe de THEIL pour ces systèmes. Une deuxième partie présente l'approche de KADANE pour l'étude du biais, l'estimateur proposé par SAWA appelé ici pseudo k/un-zéro, délimite l'intérêt de cette proposition et les conditions de la dominance sur les estimateurs de la k-classe. Une k-classe généralisée d'estimateurs est même présentée, dans cette recherche, pour l'optimisation des méthodes d'estimation d'un système interdépendant.

---

(1) Article remis en Mai 1975, révisé en Avril 1976.

Mots-clés : Estimation ; régression ; Econométrie ; systèmes interdépendants.

0 – PRELIMINAIRE.

0.1 – Modèle à une équation :

Soit un modèle linéaire *à une équation* destiné par exemple à simuler la consommation Y d'un pays en fonction de la variable salaire X_1 et de la variable prix X_2 :

$$Y = \alpha_1 X_1 + \alpha_2 X_2 + E \tag{1}$$

ou, plus généralement

$$Y = \alpha_1 X_1 + \alpha_1 X_2 + \ldots + E$$

Y est la variable expliquée, dite endogène, dont on a un vecteur de n observations.

Chaque X_i est une variable explicative, dite exogène, prédéterminée, dont on a un vecteur de n observations.

On pose : $X = (X_1, X_2, \ldots)$ et, soit α le vecteur des α_i ; X est de dimensions (n.q) s'il y a q variables exogènes.

Les coefficients α_i sont des paramètres inconnus à estimer. E est un vecteur colonne aléatoire de n résidus de loi λ, d'espérance nulle, de variance σ^2 inconnue, sans autocorrélation d'une observation à une autre, avec X indépendant du résidu.

Le triplet $\{(Y \quad X), \alpha, \lambda\}$ définit, par une structure statistique, le modèle que nous avons contruit à partir de la situation économique concrète.

On rappelle qu'on obtient pour le vecteur α des α_i le même estimateur A sans biais par la méthode des moindres carrés ordinaire ou par celle du maximum de vraisemblance :

$$\hat{\alpha} = A = (X' X)^{-1} X' Y \text{ (S1)}$$

$$a = (x' \ x)^{-1} x' y$$

Les minuscules a , x , y , etc. représentent, dans cet article, les observations de A , X , Y , etc. L'estimation a est une valeur de l'estimateur A, obtenue pour l'échantillon observé x , y (si l'estimation d'un paramètre α est a, on note aussi $a = \hat{\alpha}$).

0.2 – Modèle à plusieurs équations :

Si dans le modèle on a *plusieurs équations* avec chaque fois une *autre* variable endogène expliquée, et de nouvelles variables exogènes, par exemple, après l'équation (1) la variable "importations" I en fonction de la production intérieure X_3, etc.

$$I = \alpha_3 X_3 + \ldots \text{ (1')}$$

On a, avec les équations (1), (1')..., un système linéaire d'équation *indépendantes* dont on peut estimer les paramètres pour chaque équation par la méthode ordinaire des moindres carrés citée déjà en 0.1. (en Anglais : Ordinary least square, O.L.S. ; en Français : M.C.O.).

Mais on peut aussi avoir, dans certaines équations, la variable endogène expliquée par des variables déjà utilisées dans d'autres, exogènes mais aussi endogènes, par exemple :

– dans l'équation (1) la consommation aussi fonction de la production X_3 et des importations I

– et dans la (1') les importations aussi fonction de la consommation $Y, \ldots,$

on ne peut plus considérer chaque équation séparément, avec dans chacune, indépendantes du résidu, les variables expliquant la variable endogène considérée. On a ici un système linéaire d'équations *interdépendantes* que l'on écrira sous sa forme la plus générale, s'il y a p équations et donc p variables endogènes :

$$Y_1 + \beta_{12} Y_2 + \beta_{13} Y_3 + \ldots + \beta_{1p} Y_p + \gamma_{11} X_1 + \ldots + \gamma_{1q} X_q + E_1 = 0$$

$$\beta_{21} Y_1 + Y_2 + \beta_{23} Y_3 + \ldots + \beta_{2p} Y_p + \gamma_{21} X_1 \ldots + \gamma_{2q} X_q + E_2 = 0$$

etc., et, ceci pour n observations dans chacune des p équations.

$$\boxed{\text{ou, aussi : } Y \beta + X \gamma + \sigma U = 0 \ (2)}$$

avec β et γ matrices des coefficients respectifs β_{ij} et γ_{ij} et de plus, chaque $\beta_{ii} = 1$ pour $i = (1, \ldots p)$ et $U = E/\sigma$ (variable réduite). La forme (2) est dite *structurelle,*

On a ainsi pour β , $p^2 - p$ paramètres à estimer et pour γ, pq paramètres à estimer, soit au total $p^2 - p + pq$ paramètres. La restriction $\beta_{ii} = 1$, $\forall_i$, sera dite normative.

Toutefois, dans chaque équation, dans un modèle donné, on n'explique pas toujours la variable endogène par toutes les autres variables endogènes et exogènes du système et des coefficients β_{ij} et γ_{ij} sont supposés nuls au départ, ce qui réduit d'autant le nombre de paramètres à estimer. Ce dernier type de restriction sera dit a priori.

On a donc les conditions suivantes :

Y est une matrice n.p d'observations de p variables endogènes.

β est une matrice non singulière p.p de coefficients, paramètres à estimer.

X est une matrice n.q d'observations de q variables exogènes.

On va être amené à poser sur la variable aléatoire U centrée, de variance inconnue, 2 types d'hypothèses. Ces hypothèses sont nécessaires pour que les méthodes d'estimation que l'on va utiliser soient valides, mais il faudra prendre garde qu'elles risquent aussi d'éloigner les modèles de la réalité qu'ils veulent simuler ou prévoir :

γ est une matrice q.p des coefficients, paramètres à estimer.

U est une matrice n.p de résidus structurels, variables aléatoires centrées.

σ est un scalaire positif, écart type inconnu à estimer du résidu U.

Hypothèses (H0) : Les lignes de U ont chacune une distribution multinormale dans R^p de vecteur espérance nulle et de matrice de variance définie positive $\Omega = \frac{1}{n} E(U'U)$ avec les éléments u_i indépendants 2 à 2 (pas d'autocorrélation).

Hypothèses (H1) : Les variables prédéterminées sont uniquement les variables exogènes ; on les suppose 2 à 2 indépendantes, et chacune indépendantes du résidu U (homoscédasticité).

Corrolaires 2 (H2) : Dans une équation k quelconque, une variable explicative peut être endogène, expliquée dans une autre équation ; dès lors elle n'est plus indépendante du résidu U. Dans une équation k quelconque, on ne peut plus supposer systématiquement l'indépendance linéaire 2 à 2 des variables explicatives (en effet, le système peut très bien présenter, dans une autre équation, une variable endogène, explicative dans l'équation k, comme combinaison linéaire des variables exogènes, explicatives dans k).

Forme réduite : Dans la forme (2) dite structurelle, en multipliant par β^{-1} à droite des deux membres on passe à la forme *réduite :*

$$\boxed{Y = X\alpha + \sigma V \qquad \text{avec} \qquad \alpha = -\gamma\beta^{-1} \ (3)}$$

Dans ce système, sous forme réduite (3), si l'on obtient une estimation de α par une méthode quelconque, se pose le problème de l'identification de γ et β à partir de α.

Au niveau des estimateurs $\hat{\alpha} = -\hat{\gamma}\hat{\beta}^{-1}$; cette écriture résume un système de pq équations avec $p^2 - p + pq$ paramètres à estimer (valeurs inconnues de β et γ) ; seules les estimations α_{ij} des α_{ij} seront alors connues. (On a supposé ici le modèle sous forme explicite avec la contrainte dite normative $\forall_i \beta_{ii} = 1$ dans la $i^{ème}$ équation).

Dans ce cas, il y a sous identification ou impossibilité de résoudre le système ; si $p(p-1)$ *ou plus* valeurs des β et γ sont supposées nulles au départ (restrictions a priori) il y a possibilités d'identification ou de *suridentification.* Ce dernier cas est pratiquement le plus fréquent. THEIL a défini pour ce cas une famille d'estimateurs, connues sous le nom de k-classe, à laquelle appartiennent les estimateurs MCO, ceux des doubles moindres carrés, etc. dits à information limitée ; il existe d'autres estimateurs que ceux de la k-classe de THEIL pour les modèles interdépendants, celui des triples moindres carrés, celui du maximum de vraissemblance, à information complète, d'autres obtenus par l'introduction de variables instrumentales, etc. Il s'agit dans la $1^{ère}$ partie de faire une étude synthétique et comparative des estimateurs bien connus de la k-classe, notamment au sujet des conditions d'existence de leurs moments et de biais, qui ont été beaucoup moins étudiées.

1 – LES ESTIMATEURS DE LA K-CLASSE DE THEIL ET LEURS MOMENTS.

1.0 – Les moments de l'estimateur des moindres carrés.

Soit un système interdépendant :

On a la forme structurelle : $Y\beta + X\gamma + \sigma U = 0$. Si l'on considère une seule équation du système, on peut aussi écrire :

$$\underline{Y} = Y^* \beta^* + X^* \gamma^* + \sigma U$$

ou

$$\underline{Y} = (Y^* , X^*)\, \delta + \sigma U \quad (4) \qquad \text{avec} \qquad \delta = \begin{bmatrix} \beta^* \\ \gamma^* \end{bmatrix}$$

$\underline{Y}$ est une colonne de Y à n éléments (n, nombres d'observations)

Y^* (resp. X^*) est une matrice $n.p_1$ des p_1 variables endogènes présentes dans l'équation, hormis $\underline{Y}$ (resp. $n.q_1$ des q_1 exogènes).

β^* (resp. γ^*) est le vecteur des coefficients paramètres à estimer, associés à Y^* (resp. X^*).

Une telle méthode pour estimer le système *équation par équation* sera dite à information limitée.

Soient p_2 et q_2 les nombres de variables endogènes et exogènes exclues de cette équation et q_1 le nombre de variables exogènes présentes dans le terme de droite de cette équation :

On a :

$$p_1 + p_2 + 1 = p$$

$$q_1 + q_2 = q$$

On suppose qu'on peut identifier les paramètres de cette équation et soit ℓ $(= q_2 - p_1)$ le degré de suridentification. L'estimateur D de la k-classe de THEIL est donné à partir d'équations normales du type

$$\hat{\delta} = \begin{bmatrix} Y^{*\prime} Y^* - k\, T' T\, ; & Y^{*\prime} X^* \\ X^{*\prime} Y^* & X^{*\prime} X^* \end{bmatrix}^{-1} \begin{bmatrix} (Y^* - k\, T)' \\ X^{*\prime} \end{bmatrix} \underline{Y} \qquad \text{avec} \qquad T = Y^* \bar{A}_x$$

$$\boxed{\begin{array}{lll} \text{ou} & \text{par : } D = \hat{\delta} = (Z' (I - k\bar{A}_x) Z)^{-1} Z' (I - k\bar{A}_x)\, \underline{Y} & \\ & \text{avec} \qquad Z = (Y^* , X^*) & \\ & \text{avec} \qquad A_x = X (X' X)^{-1} X' & \\ & \text{avec} \qquad \bar{A}_x = 1 - A_x & (S/0) \end{array}}$$

On ne peut, pour l'instant, que remarquer pour $k = 0$ l'identification avec l'estimateur des moindres carrés ordinaires (MCO) :

$$D = (Z' Z)^{-1} Z' \underline{Y}$$

Encore, un tel estimateur pour le vecteur des paramètres dans un système *suridentifié* suppose-t-il la levée de l'hypothèse 1 et ne s'appliquerait qu'à des systèmes *indépendants*. Toutefois, certains économètres pour des modèles macroéconomiques interdépendants, surtout de prévision à court terme, continuent à les estimer par MCO (Modèle de la Communauté Economique Européenne 1971, de la Banque de France 1972, . . .) malgré la contre indication des hypothèses posées.

En 1972, MARIANO a démontré dans ECONOMETRICA [4], le théorème suivant :

Théorème 0 :

Soit un système interdépendant juste identifiable (alors $q_2 \geqslant p_1$) et dans une équation de ce système :

q_1 le nombre de variables exogènes incluses

$p_1 + 1$ le nombre de variables endogènes incluses,

n le nombre d'observations, supposé supérieur ou égal à $p + q$.

Le moment d'ordre r de l'estimateur MCO existe si et seulement si $r < n - p_1 - q_1 + 1$.

Pour la démonstration de ce théorème, il suffit de remarquer que l'estimateur MCO de δ est une sous matrice de $F = Y' P_k Y$ avec $P_k = 1 - K \bar{A}$.

F suit une distribution de WISHART à $q_1 + 1$ degrés de liberté pour laquelle on étudie les conditions d'existence des moments.

1.1 – **L'estimateur des doubles moindres carrés et ses moments** – (En anglais : Two stage least square 2 T.L.S. – en français DMC).

De nombreux travaux ont été publiés à ce sujet, notamment (voir aussi bibliographie) :

* *A la suite des travaux de la COWLES Commission, entre les années 1940 et 1950, sur la résolution des modèles économétriques aux U.S.A., THEIL a été amené à imaginer cet estimateur.*
* *En 1953, BASMANN, par des voies différentes, a été amené à préciser le même estimateur.*
* *THEIL, en 1954, a publié un ouvrage où il est devenu plus persuasif sur l'usage de cet estimateur.*
* *BASMANN [0], en 1961, dans* "JASA" a intuitivement montré *que les moments de cet estimateur existaient pour une équation si l'ordre était inférieur à :* $q_2 - p_1 + 1$.
* *SAWA, en 1969, a confirmé cette intuition pour* $p_1 = 2$ *avec p, nombre d'équations, arbitraire.*
* *MARIANO, en 1972, dans* "Econométrica" a démontré *que les moments de cet estimateur existaient pour chaque équation avec la condition de BASMANN, soit avec l'ordre inférieur à :* $q_2 - p_1 + 1$ [4].

* *NAGAR, en 1974, dans* "Econométrica" a précisé *l'espérance de cet estimateur, quand* $p_1 = 2$ *[6].*

En France, face à ces travaux, il existe de très fines descriptions de cet estimateur, notamment dans :

- *"Méthodes statistiques de l'économétrie" du Professeur MALINVAUD en 1969 [3].*
- *"Logique de l'investigation économétrique" de BONNAFOUS en 1973, ouvrage complétant le 1° sur le point des comparaisons (les 2 cités en bibliographie). [1].*

Dans les modèles macro-économiques à court terme, il est couramment employé (modèle STAR de la Direction de la Prévision du Ministère des Finances, Modèle MATOUK de l'Université de PARIS, etc.).

NOUS POURRIONS APPELER L'ESTIMATEUR DMC ESTIMATEUR k-UN, PUISQU'IL EST OBTENU A PARTIR DE L'ESTIMATEUR DE LA k-CLASSE EN FAISANT $k = 1$: ON APPELLERAIT ALORS L'ESTIMATEUR MCO, ESTIMATEUR $k - ZERO$.

1.1.a – **Rappel sur l'estimateur des doubles moindres carrés** (DMC).

A partir de la forme structurelle (2) on a un nuage d'observations dans R^s ($s \leqslant p + q$). On ne peut, en général, avoir $s = p + q$ car les p variables endogènes sont peut-être, pour certaines, linéairement dépendantes avec les q variables exogènes (hypothèse H2).

Ces dernières sont 2 à 2 indépendantes, et chacune indépendante de U (H1), ce qui n'est pas le cas de Y*.

Dans toute méthode à information limitée, comme celle des DMC, on a pour une équation isolée un nuage d'observation dans :

$$R^{s1} \ (s_1 \leqslant p_1 + q_1 + 1)$$

On ne peut en général avoir $s_1 = p_1 + q_1 + 1$ en raison de l'hypothèse (H2) évoquée ci-dessus. L'hypothèse (H1), évoquée aussi ci-dessus, nécessaire pour une estimation par MCO, n'est respectée que pour X* et non pour le vecteur Y* des variables endogènes.

On pense alors à remplacer Y* par une variable instrumentale la plus proche possible de Y* pour laquelle l'hypothèse (H1) serait respectée ; on choisit pour cela $\hat{Y}^*$ obtenue en estimant chaque $Y_k \in Y^*$ en fonction de l'ensemble des q variables exogènes, selon la méthode des MCO. Comme les variables exogènes sont indépendantes du résidu, $\hat{Y}^*$ combinaison linéaire des variables exogènes sera aussi indépendant des U et l'on aura ainsi une variable instrumentale très proche des Y* qui vérifie l'hypothèse d'homoscédasticité (H1).

1.1.b – **1ère Etape.**

On a la conduite suivante, dans l'équation (4) :

$$\underline{Y} = (Y^* , X^*) \ \delta + \sigma U$$

On remplace préliminairement chacune des variables endogènes sauf $\underline{Y}$ par leur valeur estimée en fonction de l'ensemble des variables exogènes suivant le modèle :

$$\forall\, Y_k \in Y^* \qquad \hat{Y}_k = \sum_{i=1}^{q} \phi_i X_i + V_1$$

avec, ϕ $= (\phi_{1'}, \phi_{2'}, \ldots \phi_q)$ vecteur de paramètre à estimer

X $= (X_1, \ldots X_q)$ vecteur des variables exogènes

V_1 variable aléatoire centrée de variance résiduelle σ_v^2. par MCO, on obtient :

$\underline{\hat{\phi}}$ $= (X'\,X)^{-1}\,X'\,Y_k$

$\hat{Y}_k$ $= X\,\hat{\phi} = X\,(X'\,X)^{-1}\,X'\,Y_k$

1.1.c – 2ème **Etape.**

En substituant $\hat{Y}^*$, l'équation (5) devient 5' :

$$\boxed{\underline{Y} = (\hat{Y}^*\,,\hat{X}^*)\,\delta + \sigma\,U \quad (5')}$$

ou en posant : $\hat{Z}^* = (\hat{Y}^*\,X^*)$

$$\hat{Y}^* = \hat{Z}^*\,\delta + \sigma\,U$$

Les variables $\hat{Y}_{k'}$, $\hat{Y}^*$, $\hat{\underline{Z}}^*$, à l'origine de cette transformation, sont dites instrumentales. L'estimation par MCO donne, appliquée une 2ème fois :

$$\boxed{D = \hat{\delta} = (Z^{*\prime}\,Z^*)^{-1}\,Z^{*\prime}\,\underline{\underline{Y}}}$$

Après transformations, l'estimation se présente comme celle obtenue par l'estimateur de la k-classe de THEIL pour k = 1 :

$$\boxed{\begin{array}{c} d = \hat{\delta} = (z'\,az)^{-1}\,z'\,a\underline{y} \quad (S2) \\ \text{avec} \quad a = x\,(x'\,x)^{-1}\,x' \end{array}}$$

Cette écriture transformée de l'estimation est souvent considérée comme la plus pratique pour le calcul. MALINVAUD, dans son ouvrage, aboutit à la même estimation $\delta = \begin{pmatrix} \hat{\beta} \\ \hline \hat{\gamma} \end{pmatrix}$ pour la même méthode par résolution d'une équation,

du type déjà cité :

$$\begin{pmatrix} (Y^* - T)' \\ X^{*\prime} \end{pmatrix} \underline{\underline{Y}} = \begin{pmatrix} Y^{*\prime}\,Y^* - T'\,T\,;\, Y^{*\prime}\,X^* \\ X^{*\prime}\,Y^* \qquad ;\, X^{*\prime}\,X^* \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix}$$

avec $T = Y^* - (X'\,X)^{-1}\,X'\,Y^* = \bar{A}\,Y^*$

La matrice des variances-covariances asymptotiques est alors pour :

$\sqrt{n}\ (\hat{\gamma} - \gamma^{\circ})$ et $\sqrt{n}\ (\hat{\beta} - \beta^{\circ})$
$\text{Var. (D)} = \sigma^2 \begin{bmatrix} Y^{*\prime}\,Y^* - T'T\,; & Y^{*\prime}\,X^* \\ X^{*\prime}\,Y^* & ;\,X^{*\prime}\,X^* \end{bmatrix}^{-1}$ (S2)

On comparera avec la même matrice, dans le cas d'un système d'équations indépendantes estimées par MCO :

$$V(A) = \sigma^2 (X' X)^{-1} \quad \widehat{V(A)} = S^2 (X' X)^{-1}$$

$$\text{avec} \quad S^2 = \hat{\sigma}^2 = \frac{\Sigma (Y - \hat{Y})^2}{n - p - 1} \qquad \text{(S1')}$$

MALINVAUD propose, dans la méthode DMC, un test d'hypothèse nulle dit de suridentification sur les coefficients de variables exogènes $X_k \in X^*$. Mais le problème est préalablement celui de l'existence des moments de cet estimateur qui présente généralement un biais.

Théorème 1 :

Soit une équation d'un système interdépendant, tel qu'il a été défini jusqu'ici :

$$\underline{Y} = (Y^*, X^*)\ \delta + \sigma U$$

Soit q_2 le nombre de variables exogènes exclues de l'équation.

Soit p_1 le nombre de variables endogènes incluses dans l'équation, mis à part $\underline{Y}$.

Soit D l'estimateur de δ défini par DMC.

Le moment d'ordre r de D existe si, et seulement si,

$$r < q_2 - p_1 + 1$$

La démonstration de ce théorème opère selon la même démarche que pour le théorème 0 mais l'intérêt du résultat est ici bien plus grand.

– Le calcul de E (D) effectué pour $p_1 = 1$ nécessite, dans l'équation étudiée, au moins 1 variable exogène absente pour que E (D) existe, 2 absentes pour que les écarts-types de coefficients existent.

– Ce même calcul également effectué pour $p_1 = 2$ nécessite alors, dans l'équation étudiée, au moins 2 variables exogènes absentes pour que E (D) existe, 3 absentes pour que les écarts-types des coefficients existent.

L'expression E (D) n'est, dans aucun cas, d'allure simple

1.1' – L'estimateur des triples moindres carrés et ses moments

Il n'appartient pas à la k-classe de THEIL, mais figure ici car il se déduit de l'estimateur DMC. Il a été mis au point, en 1962, par ZELLNER qui a montré que l'estimation des triples moindres carrés (TMC) – (en anglais : Three stage least square 3 SLS) – était parfois plus efficace que l'estimation par DMC. On rappelle que la première phase est identique dans les deux méthodes, pour un même système interdépendant. L'on va ensuite estimer globalement le vecteur $\underline{\delta}$ des δ, non plus par les méthodes MCO mais par celle des moindres carrés généralisés en définissant Ω, matrice inconnue des variances-covariances des erreurs. L'estimation n'est alors plus à information limitée.

On estime d'abord Ω par la matrice des variances-covariances des résidus empiriques obtenues par les moindres carrés ordinaires dans (4).

Ayant ainsi obtenu $\hat{\Omega}$ on applique à (4) la méthode des moindres carrés généralisés :

$$\boxed{D = (\underline{Z}^{*\prime}\, \Omega^{-1}\, \underline{Z}^{*})^{-1}\, Z^{*\prime}\, \Omega^{-1}\, \underline{Y} \quad \text{(S3)}}$$

où $\underline{Z}^*$ est une matrice diagonale dont le $i^{\text{ème}}$ élément est le Z^* de la $i^{\text{ème}}$ équation.

On remarque que les méthodes DMC et TMC donnent les mêmes estimations si Ω est diagonale, c'est-à-dire si les résidus des différentes équations ne sont pas corrélés entre eux. Sinon, plus la corrélation est forte, plus l'emploi de la méthode TMC offre un intérêt pour l'estimation d'un système économétrique interdépendant.

On pourrait appeler l'estimateur obtenu pseudo k-un puisque l'estimateur DMC est appelé k-un. MALINVAUD, pour sa part, préfère en cas d'hétéroscédasticité une autre spécification du modèle et ce n'est pas par hasard que la méthode TMC est absente de son ouvrage déjà cité "Méthodes Statistiques de l'Econométrie".

En 1974, dans le cadre de la Cowles-Fundation, SARGAN a démontré le théorème suivant :

Théorème 1' :

Soit une équation d'un système interdépendant, tel qu'il a été défini jusqu'ici :

$$\underline{Y} = (Y^*\,,X^*)\,\delta + \sigma\,U$$

Soit q_2 le nombre des variables exogènes exclues de l'équation.

Soit p_1 le nombre de variables endogènes incluses dans l'équation, mis à part $\underline{Y}$.

Soit D l'estimateur de δ défini par TMC.

Le moment d'ordre r de D existe si, et seulement si,

$$r < q_2 - p_1 + 1$$

Nous proposerons ici le théorème suivant :

Théorème 2 : (Généralisation des théorèmes 0, 1 et 1')

Soit une équation d'un système avec les notations précédentes :

$$\underline{Y} = (Y^*, X^*)\ \delta + \sigma U$$

Soit q_2 le nombre de variables exogènes exclues de l'équation.

Soit p_1 le nombre de variables endogènes incluses dans l'équation, mis à part Y.

Soit D l'estimateur de δ défini par MCO, DMC ou TMC.

Le moment d'ordre r de D existe si, et seulement si,

$$r < f\,(n - q) + q_2 - p_1 + 1$$

avec f = 0 en cas de système interdépendant suridentifiable

f = 1 en cas de système interdépendant juste identifié, ou de système indépendant, estimé par M.C.O.

On a ainsi pu construire un théorème que les économètres auront intérêt à appliquer lors d'une estimation des paramètres d'un système interdépendant, ce qui leur permettra de rejeter la validité de certaines prévisions ; l'étude des biais des estimateurs doit maintenant pouvoir nous aider à préférer un estimateur à un autre dès que les dimensions du système spécifié sont bien connues. Pour l'étude de l'estimateur de Maximum de Vraisemblance à information limitée, (MVIL) estimateur de la k-classe non traité ici, on se reportera à [3].

2 – ETUDE DES BIAIS ET K-CLASSE GENERALISEE POUR L'ESTIMATION DES SYSTEMES INTERDEPENDANTS.

2.1 – L'approche de KADANE et l'estimateur pseudo k/un-zéro :

Les estimateurs de la k-classe, pour un système économétrique interdépendant, sont généralement biaisés, et, lorsqu'on cherche à diminuer le biais, c'est souvent au détriment de la variance de l'estimateur qui augmente. KADANE, en 1971, dans "Econométrica", à l'aide d'une série assymptotique d'un multiple scalaire σ de la variance résiduelle, a étudié le biais approché et l'erreur quadratique dans l'estimation d'un système interdépendant.

Il a montré que, pour certains estimateurs, le biais tend vers zéro quand, dans la série asymptotique sus-citée, les termes en σ^3 et d'ordre inférieur sont ignorés. Ces estimateurs sont appelés par KADANE "Small σ assymptotically unbiased" (S.S.A.U.).

Pour un estimateur de la k-classe d'un système présenté sous forme structurelle (4) $\underline{Y} = (Y^*, X^*)\ \delta + \sigma U$, les équations normales s'écrivent :

$$(8)\quad \begin{bmatrix} (Y^* - kT)' \\ X^{*\prime} \end{bmatrix} \underline{Y} = \begin{bmatrix} Y^{*\prime} Y^* - kT'T\,; & Y^{*\prime} X^* \\ X^{*\prime} Y^* & X^{*\prime} X^* \end{bmatrix} \hat{\delta}\,(k), \text{ noté aussi :}$$

$$H' \underline{Y} = (H' H) \hat{\delta} (k)$$

et var. (D) $= \sigma^2 (H' H)^{-1}$ avec le biais $e_k = (H' H)^{-1} H' \sigma U$ avec les notations déjà indiquées notamment en 1.1. pour l'estimateur D.M.C. (k = 1).

A partir du développement $(1 = h \Delta)^{-1} = I - h \Delta + h^2 \Delta^2 - h^3 \Delta^3 + \ldots$, KADANE a démontré le théorème suivant :

Théorème 3 :

Le biais asymptotique de l'estimateur de la k-classe pour le paramètre δ (k), dans une équation d'un système interdépendant sous forme (4), est tel que :

$$E (e_k) = \sigma^2 r_k \zeta + 0 (\sigma^3)$$

avec
$$e_k = \hat{\delta}_k - \delta_k$$

et

$$r_k = (1 - k) n + kq - (p_1 + q_1 + 1)$$

avec
$$0 \leqslant k \leqslant 1 \tag{9}$$

ζ est un vecteur dont les éléments sont fonction de β, γ et Ω. $0 (\sigma^3)$ est le terme limitant le développement, en fonction de σ. Si k n'est pas fixé (cas de l'estimateur MVIL) :

$$E (k) = - \sigma^2 \zeta + 0 (\sigma^3)$$

Corrolaire 1 :

Si l'on pose $k = 1 + \dfrac{\alpha}{n}$ et si l'on fait tendre n vers l'infini, on retrouve les résultats de NAGAR sur le biais.

Il faut rappeler qu'en 1959, NAGAR, pour examiner le biais et la matrice des variances dans des estimateurs de la k-classe, avait envisagé une approche trouvant le terme en $1/n$ dans le biais asymptotique pour un grand échantillon et le terme en $1/n^2$ de la matrice des moments ; ces résultats s'appliquaient à la k-classe, mais uniquement pour des valeurs de k proche de l'unite. Si, dans l'approche de KADANE, on fait tendre le paramètre n vers l'infini, on retrouve donc les résultats de NAGAR ; ceci est notamment vrai pour l'estimateur D.M.C.

L'AUTEUR A TENU ICI A PRESENTER DES CORROLAIRES AVEC DES REMARQUES QUI FIXENT UNE LIMITE ET UN PLUS GRAND INTERET A L'APPROCHE DE KADANE AU-DELA DES THEOREMES DEJA CONNUS 3 (KADANE) 4 et 5 (SAWA).

Corrolaire 2 :

Si $r_k = 0$ dans l'expression (9), le biais asymptotique espéré est nul :

$$r_k = 0 => k = \frac{n - (p_1 + q_1 + 1)}{n - q}$$

L'estimateur ainsi obtenu est alors appelé "small sigma assymptotically unbiased" (S.S.A.U.). On pourrait penser ainsi, à partir de la forme générale d'un estimateur de la k-classe pour un système interdépendant (S.O.), définir une valeur de k pour laquelle l'estimateur est asymptotiquement sans biais (S.S.A.U.).

On peut aussi, pour certaines valeurs de k (estimateurs des simples et doubles moindres carrés), chercher le nombre optimal de variables endogènes et exogènes présentes explicatives (resp. p_1 et q_1) dans une équation, et celui des absentes (resp. p_2 et q_2), pour avoir un estimateur asymptotiquement sans biais (S.S.A.U.).

KADANE attachait au moins autant d'importance à la mesure asymptotique d'un biais qu'à la présentation d'estimateurs S.S.A.U. Or, par la suite, c'est ce second aspect de son travail qui fut surtout retenu.

Corrolaire 3 :

L'application du corrolaire 2 aux estimateurs M.C.O. ($k = 0$) et D.M.C. ($k = 1$) a conduit de nombreux auteurs à présenter les résultats suivants : l'estimateur D.M.C. (resp. M.C.O.) est S.S.A.U. si $q_2 = p_1 + 1$ (resp. si $n = p_1 + q_1 + 1$). L'expression $q_2 - p_1$ a déjà été vue comme étant le degré ℓ de suridentification.

Or, l'application du théorème 1 (resp. théorème 0) montre que, *dans ce cas, la variance de l'estimateur tend vers l'infini* et l'absence de biais asymptotique n'offre plus aucun intérêt. Ce résultat s'étend à toutes les valeurs de k entre zéro et un, bien qu'aucune démonstration de ce fait, comme aucune condition d'existence des moments d'un estimateur quelconque de la k-classe, ne soit, à notre sens, entièrement satisfaisante à ce jour, dans le cas le plus général.

L'approche de KADANE peut être étendue à d'autres estimateurs que ceux de la k-classe. SAWA, en 1973, (International Economic Review), à mis au point, par certaines combinaisons linéaires d'estimateurs de la k-classe, des estimateurs asymptotiquement sans biais (S.S.A.U.).

Théorème 4 (SAWA) :

Soient deux estimateurs $\hat{\delta}(k)$ et $\hat{\delta}(k')$ de la k-classe définis comme précédemment, l'estimateur :

$$\frac{1}{1-\lambda}((\hat{\delta}(k) - \lambda\,\hat{\delta}(k')) \qquad \text{est} \qquad \text{S.S.A.U.}$$

avec,

$$\lambda = \frac{(1-k)\,n + kq - (p_1 + q_1 + 1)}{(1-k')\,n + k'q - (p_1 + q_1 + 1)} \qquad \text{et} \qquad k \geqslant 0 \;,\; k' \geqslant 1$$

Corrolaire 4 :

Alors, soit $\hat{\delta}(1)$ l'estimateur DMC et $\hat{\delta}(0)$ l'estimateur MCO, l'estimateur $\hat{\delta}^{10} = \dfrac{1}{1-\lambda}(\hat{\delta}(1) - \lambda\,\hat{\delta}(0))$ est S.S.A.U.

avec,

$$\lambda = \frac{q - (p_1 + q_1 + 1)}{n - (p_1 + q_1 + 1)} = \frac{q_2 - p_1 + 1}{n - (p_1 + q_1 + 1)}$$

Cet estimateur, combinaison linéaire des estimateurs des simples et doubles moindres carrés, n'appartient pas à la k-classe : il sera appelé ici pseudo-k/un-zéro. L'existence des moments de cet estimateur n'a été montré que dans certains cas particuliers, et l'auteur pense pouvoir généraliser la démonstration. Enfin, il faut noter que si l'on inverse l'ordre ($k = 0$ et $k' = 1$) on retrouve le même estimateur.

Il peut aussi s'écrire

$$\widehat{\delta^{10}} = \hat{\delta}(1) + \frac{\gamma}{1 - \gamma} (\hat{\delta}(1) - \hat{\delta}(0))$$

De plus, quand dans un modèle spécifié la taille de l'échantillon croît et devient très grande, λ devient très petit ainsi que le terme correctif à droite de $\hat{\delta}_1$; vice versa si, pour le même modèle n est faible, le terme correctif sera plus important. Quand la taille de l'échantillon d'observations est faible, cet estimateur, combinaison linéaire de ceux des simples et doubles moindres carrés, apporte donc un terme correctif non négligeable pour l'estimateur des doubles moindres carrés.

Théorème 5 (SAWA) :

L'estimateur $\hat{\delta}^{10}$, pseudo k/un-zéro est dominé par l'estimateur DMC si et seulement si :

$$(n - q - 2)(q_2 - p_1 - 7) \leqslant 12$$

On rappelle qu'en général $\hat{\delta}$ domine $\hat{\delta}'$ si, et seulement si :

$$\Omega(\hat{\delta}') \geqslant \Omega(\hat{\delta}) \qquad \text{avec} \qquad \Omega(\delta)$$

matrice des variances résiduelles.

La démonstration s'opère par une première étape pour établir que :

$$\Omega(\hat{\delta}^{10}) = \Omega(\hat{\delta}(1)) + \zeta_1 \sigma^4 + \theta(\sigma^6)$$

ou θ est le terme qui limite le développement, et ζ_1 fonction de

$$n, q, q_2, p_1, X, X^*, \beta, \gamma, \Omega$$

Par exemple, pour $n = 25$, $p_1 = q_1 = 2$, $q = 15$, l'estimateur pseudo k/un-zéro domine alors l'estimateur DMC. C'est souvent le cas dans les grands systèmes, sans que l'on puisse dégager une règle générale.

On peut, d'autre part, montrer que pour un degré de suridentification ℓ inférieur ou égal à six l'estimateur DMC domine l'estimateur MVIL, au sens de KADANE.

Par exemple pour $n = 6$ et un système d'équations linéaires :

$$y_1 = f_1 (y_2 x_1 x_2) \quad p = 2 \quad q_1 = 2$$

$$y_2 = f_2 (y_1 x_1 x_3) \quad q = 3 \quad p_1 = q_2 = 1$$

L'estimateur MVIL est dominé par l'estimateur DMC, lui-même dominé par l'estimateur pseudo k/un-zéro.

Lemme 1 :

On peut même définir une relation de préférence}, basée sur la dominance:

$$\delta' \} \; \delta \Leftrightarrow \delta' \quad \text{domine} \quad \delta$$

Corrolaire 5 :

Si on a la condition suffisante $\ell \leqslant 6$ *et* $n \geqslant q + 2$ (estimateur pseudo k/zéro-un)} (estimateur DMC)} (estimateur MVIL).

2.2 – **Vers une k-classe généralisée.**

Quelques tentatives non-significatives pour généraliser fructueusement la k-classe, invitent à considérer le problème suivant :
Soit pour un système interdépendant sous forme (4)

$$\underline{Y} = (Y^* , X^*)\, \delta + \sigma U$$

un estimateur à information limitée, obtenu à partir d'équations normales de type (8) généralisées :

$$(8') \quad \begin{bmatrix} Y^{*\prime}\, W^{*\prime} \\ X^{*\prime} \end{bmatrix} \underline{Y} = \begin{bmatrix} Y^{*\prime}\, W^{*\prime}\, Y^* \; ; & Y^{*\prime}\, W^{*\prime}\, X^* \\ X^{*\prime}\, Y^* & X^{*\prime}\, X^* \end{bmatrix} \hat{\delta}$$

avec les notations déjà utilisées en 1.0 et 1.1 et, notamment :

$$A_X = X\,(X'\,X)^{-1} \qquad \bar{A}_X = 1 - A_X$$

et, en plus :

$$(8'') \; \bar{W}^* = 1 - W^* = (kI + k'\,\Lambda\,(1 - \Lambda)^{-1})\, \bar{A}_X$$

avec Λ matrice diagonale formée de la diagonale de A_X.

W^* représente une variable instrumentale substituée à Y^* pour éliminer la colinéarité avec U dans une optique d'estimation finale par MCO.

a) $\underline{k' = 0}$, on obtient en (8'') :

$$(8'') \; W^* = I - k\,(\bar{A}_X) = Y^{*\prime}\, W^{*\prime} = Y^{*\prime}\,(I - k\,\bar{A}_X)' = (Y^* - k\,T)'$$

car : $$T = Y^*\, \bar{A}_X$$

On retrouve, en (8'), les équations normales donnant l'estimateur de la k-classe de THEIL si $0 \leqslant k \leqslant 1$.

b) $\underline{k' \neq 0}$, on a un estimateur dit de la k-classe généralisée ou kk' généralisé, que l'on peut comparer à ceux de la k-classe. On peut d'abord calculer le biais par l'approche de KADANE, en envisageant une extension du Théorème 3.

Théorème 6 :

Le biais asymptotique de l'estimateur de la k-classe généralisée pour le paramètre δ (kk'), dans une équation d'un système interdépendant sous forme 4, est tel que :

$$E(e_{kk'}) = \sigma^2 r_{kk'} \zeta + \theta(\sigma^3) \qquad \text{avec} \qquad e_{kk'} = \hat{\delta}_{kk'} - \delta_{kk'}$$

et,
$$r_{kk'} = n - k(n - q) - k' q - (p_1 + q_1 + 1) \tag{9'}$$

ζ est un vecteur dont les éléments sont fonction de β, γ et Ω; $\theta(\sigma^3)$ est le terme limitant le développement en fonction de σ

On a un corrolaire correspondant au corrolaire 2 du Théorème 3.

Corrolaire 2' :

Si $r_{kk'} = 0$ dans l'expression (9'), le biais asymptotique au sens de KADANE est nul. L'estimateur obtenu est alors S.S.A.U. si l'on a :

$$n(1 - k) + (k - k')q = p_1 + q_1 + 1$$

On retrouve, pour $k' = 0$, la condition trouvée pour la k-classe dans le corrolaire 2, aussi bien par KADANE que par NAGAR.

Corrolaire 6 :

Soit l'estimateur obtenu en faisant $k = k'$ dans (8''), et appelé k^2 généralisé. D'après les théorèmes (3) et (6) :

$$r_{k^2} = (I - k)n - (p_1 + q_1 + 1)$$

$$r_k = (I - k)n - (p_1 + q_1 + 1) + kq \qquad \text{donc} \qquad r_k > r_{k^2}$$

donc, le biais asymptotique espéré sera inférieur dans l'estimateur k^2 généralisé, à celui de tout estimateur de la k-classe, pour k fixé. L'approfondissement de cette étude passera par l'étude de la matrice des variances résiduelles pour l'estimateur k^2 généralisé et par les conditions de sa dominance sur d'autres estimateurs présentés ; mais on peut penser que l'estimation des paramètres d'un système interdépendant par les doubles moindres carrés et, a fortiori par les moindres carrés ordinaires, n'est, si elle est systématique, guère justifiée, et même peut-être source d'erreurs ; un tableau synoptique joint, aidera à choisir parmi quelques estimateurs pour une analyse statistique d'un système interdépendant.

Résume synoptique des estimateurs a information limitee dans les modéles linéaire interdependants $(Y\beta + X\gamma + \sigma U = 0)$

Rappel sur le modèle linéaire à une équation	Estimateurs dans une équation avec information limitée $Y = Z\delta + \sigma U$ avec $Z = (Y^*, X^*)$				
	Estimateurs de la k-classe de THEIL				Estimateurs de la k-classe généralisée
$\underline{Y} = \alpha X + E$	$\hat{\delta} = [Z'(I - k\bar{A})Z]^{-1} Z'(1 - k\bar{A})\underline{Y}$ avec $A = X(X'X)^{-1}X'$, $\bar{A} = 1 - A$			$(k' = 0 \Leftrightarrow$ k-classe	$\bar{W}^* = [kI + k'\Lambda(I - \Lambda)^{-1}\bar{A}^*$
$\boxed{\text{MC généralisés}}$	$\boxed{\text{k-zéro}}$	$\boxed{\text{k-un}}$	$\boxed{k = g}$	$\boxed{\text{pseudo k-un généralisé}}$	$\boxed{\text{k généralisé}}$
	MCO	DMC	MVIL	TMC	Cas particulier
	$k = 0$	$k = 1$	$k = g$	(= DMC avec MC généralisés pour la variable instrumentale)	$k = k'$
$\hat{\alpha} = (X'\Omega X)^{-1}X'\Omega\underline{Y}$	$\hat{\alpha} = (X'X)^{-1}X'\underline{Y}$	$\hat{\alpha} = (Z'AZ)^{-1}Z'A\underline{Y}$			
	Existence de l'espérance de l'estimateur $(m_1(D))$				
	$p_1 + q_1 < n$ (Théorème 0)	$p_1 < q_2$ (Théorème 1)		$p_1 < q_2$ (Théorème 2)	
	Existence de la variance de l'estimateur et de $m_2(D)$				
	$p_1 + 1 + q_1 < n$	$p_1 + 1 < q_2$		$p_1 + 1 < q_2$	S.S.A.U. si $k = n - \frac{p_1 + q_1 + 1}{n}$
	Estimateur pseudo k/un zéro = S.S.A.U. (Théorème 4, cor. 4)				(Théor. & cor. 6)

LISTE DES PRINCIPAUX SYMBOLES UTILISES

a) Variables aléatoires.

Y	– Variable ou ensemble de variables endogènes.
$X = (X_1, X_2 \ldots)$	– Ensemble de variables prédéterminées (confondues ici avec les exogènes).
E	– Variable résiduelle (ne pas confondre avec E (A) espérance de A, . . .).
U	– Variable résiduelle réduite dans un modèle de forme structurelle.
V	– Variable résiduelle réduite dans un modèle de forme réduite.
$\underline{Y}$	– Variable expliquée dans une équation d'un modèle interdépendant.
Y^*	– Variables endogènes explicatives présentes dans une équation d'un modèle interdépendant.
X^*	– Variables exogènes explicatives présentes dans une équation d'un modèle interdépendant.

Les valeurs de ces variables sont représentées en minuscules y , x, . . .

b) Paramètres.

	Estimateur Correspondant	
$\alpha = (\alpha_1, \alpha_2)$	A	Ensemble des coefficients d'un modèle linéaire (forme 1) ou, dans un système interdépendant d'une forme réduite (forme 3).
σ^2	S^2	Variance résiduelle.
β (matrice de β_{ij})		Matrice de coefficients pour les variables endogènes d'un modèle linéaire (forme 2)
γ (matrice de γ_{ij})		Matrice de coefficients pour les variables exogènes d'un modèle linéaire (forme 2).
δ	D	Matrice des coefficients pour une équation d'un modèle linéaire (forme 4).
Ω	W	Matrice des variances covariances des erreurs.
ϕ		Ensemble des coefficients du modèle linéaire introduisant la variable instrumentale dans les doubles moindres carrés.

Les valeurs des estimateurs sont notées en minuscules, et on trouve aussi pour ceux-ci la notation "paramètre surmonté d'un chapeau" $\hat{\alpha}$, $\hat{\beta}$, . . .

c) Formes du modèle.

(1) Cas d'une équation système interdépendant : $Y = \alpha_1 X_1 + \alpha_2 X_2 + \ldots + E$

(2) Forme structurelle : $Y\beta + X\gamma + \sigma U = 0$

(3) Forme réduite : $Y = X\alpha + \sigma V$

(4) Forme structurelle d'une équation : $\underline{Y} = (Y^* , X^*)\,\delta + \sigma U$

(5) Modèle pour une variable instrumentale : $Y_k = \phi_1 X_1 \;\; \phi_2 X_2 + \ldots + V_1 \,;\, \forall\, Y_k \in Y^*$

d) Indices de dimension.

n nombre d'observations

q nombre de variables exogènes

p nombre de variables endogènes

Pour une équation sous forme 4.

$p_1 + 1$ nombre de variables endogènes du modèle présentes

p_2 nombre de variables endogènes du modèle absentes

q_1 nombre de variables exogènes du modèle présentes

q_2 nombre de variables exogènes du modèle absentes

BIBLIOGRAPHIE

[0] R.L. BASMANN – A note on the exact Finite sample FrequencyFonctions of generalised classical linear Estimators, Journal of the American Statistical Association, 56, 1964.

[1] A. BONNAFOUS – La logique de l'investigation économétrique, Dunod, 1973.

[2] J.B. KADANE – Comparison of k class estimators when the disturbances are small, *Econometrica,* 39, 1971.

[3] E. MALINVAUD – Méthodes Statistiques de l'Econométrie, Dunod, 1969.

[4] R. MARIANO – The existence of moments of the ordinary least squares and two stage least squares estimators, *Econometrica,* 40, 1972.

[5] A.L. NAGAR – The bias and moment matrix of the general k-class of estimators of the parameters in simultaneous equations, *Econometrica,* 27, 1959.

[6] A.L. NAGAR – The exact mean of the two stage least squares estimators of the structural parameters in an equation having three endogenous variables, *Econometrica,* 42, 1974.

[7] J.O. SARGAN – The moments of the 3 SLS estimates of the structural coefficients of a simultaneous equation model, Cowles Foundation Paper, 1974.

[8] T. SAWA – Almost unbiased estimator in simultaneous equation systems, *International Economic Review,* 14, 1973.

[9] H. THEIL – Principles of Econometrics, North Holland, 1971.

[10] A. ZELLNER & H. THEIL – Three stage least squares, *Econometrica,* 30, 1962.