

REVUE DE STATISTIQUE APPLIQUÉE

H. MADINIER

M. MOUILLART

Les méthodes d'estimation des modèles à retards échelonnés en économie

Revue de statistique appliquée, tome 31, n° 4 (1983), p. 53-73

http://www.numdam.org/item?id=RSA_1983__31_4_53_0

© Société française de statistique, 1983, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

LES METHODES D'ESTIMATION DES MODELES A RETARDS ECHELONNES EN ECONOMIE

H. MADINIER et M. MOUILLART

Université de Paris-X-Nanterre

L'usage des modèles à retards échelonnés (M.R.E.) en économie se justifie en général pour deux raisons principales⁽¹⁾ : l'introduction de retards échelonnés dans les relations économiques peut tout d'abord rendre compte de la formation des anticipations des agents, compte tenu des observations passées des variables qui président à leur prise de décision ; ou cette introduction peut se proposer de rendre compte des délais de réalisation de décisions ou d'actions économiques ; les délais de réalisation et la structure globale des retards n'étant pas connus, l'économétrie des M.R.E. se propose alors de les rechercher et de les estimer au mieux.

Sur le plan strictement théorique, la justification de l'introduction des M.R.E. en économie ne présente guère de difficulté, l'abondance de la littérature consacrée à ce sujet pouvant le cas échéant le confirmer (par exemple D.W. JORGENSON (1966), P.J. DHRYMES (1971), ou D. HERNARD et M. MOUILLART (1980).

En revanche, et tel sera l'objet de cet article, l'estimation des M.R.E. ne semble pas entièrement satisfaisante, telle qu'elle est généralement envisagée par l'économiste.

En effet, la plupart des séries économiques se caractérise par la présence d'un effet de mémoire⁽²⁾ qui affecterait la structure de la fonction aléatoire associée au M.R.E. (D. HERNARD, M. MOUILLART et D. STRAUSS-KAHN (1979). Dans ces conditions, si les spectres des séries analysées exhibent bien une forte puissance à la fréquence zéro (symptôme de la forme typique du spectre), les techniques économétriques couramment utilisées par l'économiste semblent en défaut sur nombre des hypothèses nécessaires à leur emploi. Tel sera le cas si l'on envisage les deux méthodes d'estimation des M.R.E. les plus répandues parmi les économistes : la méthode des moindres carrés ordinaires (M.C.O.) et la méthode polynomiale d'ALMON (M.P.A.).

De manière plus générale, et quelle que soit la taille de la mémoire, nous allons essayer de montrer que les deux méthodes d'estimation (M.C.O. et M.P.A.) ne semblent pas des plus satisfaisantes⁽³⁾. Et qu'il serait plus souhaitable de leur substituer l'emploi de la méthode de la régression spectrale (M.R.S.) dont les propriétés remarquables lui permettent de s'affranchir de ces contraintes.

(1) Pour une analyse plus précise de ces deux raisons, P.A. MUET (1979) p. 12-15.

(2) L'effet de mémoire sur une série temporelle se signale par une fonction d'autocorrélation qui décroît lentement et donc par une forte part de la puissance spectrale totale portée par la bande de fréquence zéro.

(3) Pour une première analyse de cette incapacité relative des méthodes "traditionnelles", H. MADINIER et M. MOUILLART (1981).

Pour ce faire, nous allons comparer les trois méthodes d'estimation en les analysant sous l'angle de leurs performances relatives en ce qui concerne les résultats obtenus avec chacune d'elle. Notre démarche s'organise donc comme suit :

– Dans une première étape, nous simulons un certain nombre de séries temporelles au moyen de M.R.E. dont nous connaissons toutes les caractéristiques (en particulier les valeurs "vraies" des coefficients que les méthodes se proposent d'estimer).

– Puis nous procéderons dans une seconde étape à l'estimation de chacun de ces modèles à l'aide des trois méthodes. Pour chacune de ces méthodes, nous apprécierons ses performances sur quatre plans distincts :

- 1) La méthode permet-elle d'appréhender la forme exacte (le profil) de la distribution des coefficients du M.R.E. ?
- 2) Les coefficients estimés sont-ils bien significatifs lorsqu'il le faut et la méthode ne risque-t-elle pas de nous conduire à retenir (ou à rejeter) à tort certains d'entre eux ?
- 3) Le nombre de retards (de coefficients) peut-il être facilement et correctement recherché et appréhendé par la méthode ?
- 4) La qualité générale des résultats (coefficients de détermination R^2 , test de DURBIN-WATSON⁽⁴⁾. . .) est-elle satisfaisante ? Nous pensons, ce faisant, apprécier le pouvoir de représentation et les qualités prédictives des méthodes.

Ces éléments relatifs à notre démarche vont nous permettre de structurer la progression de cet article. Il nous semble important, afin d'apprécier justement certains des résultats que nous présenterons, de rappeler quelques points de méthode : ceci constituera l'essentiel des développements qui suivent maintenant.

1. LES METHODES D'ESTIMATION DES M.R.E.

L'estimation des M.R.E. en économie se fonde sur le recours à trois méthodes : la méthode des moindres carrés ordinaires (M.C.O.), la méthode polynomiale d'ALMON (M.P.A.) et la méthode de la régression spectrale (M.R.S.).

La M.C.O. est d'un usage suffisamment courant pour ne pas avoir à en rappeler les principes⁽⁵⁾. En revanche, une présentation rapide de la M.P.A. et de la M.R.S. nous permettra d'insister sur les aspects principaux et de souligner certaines de leurs caractéristiques nécessaires à une bonne intelligibilité de nos résultats.

(4) La statistique associée au test de DURBIN-WATSON (à savoir le rapport de VON NEUMANN sur les résidus) sera notée DW par la suite. Cette valeur, comprise entre 0 et 4, est de l'ordre de 2 en l'absence de non-corrélation des résidus (par exemple, E. MALINVAUD (1972).

(5) Si l'on choisit de recourir à la M.C.O., le M.R.E. est ramené à un modèle de régression multiple (pour des raisons d'ordre théorique ou par manque d'informations précises quant à la structure des retards).

Nous savons que dans ces conditions, les variables exogènes seront très fortement colinéaires : mauvaise estimation des coefficients. . . En outre, le choix du nombre de retards n'est pas toujours très aisé.

1.1. La méthode polynomiale d'ALMON

La méthode polynomiale d'ALMON que nous noterons par la suite MPA pour plus de commodité, a été mise au point en 1964 par S. ALMON. Les deux objectifs principaux poursuivis par cette méthode sont les suivants :

– d'une part la réduction de la colinéarité des variables exogènes entre elles ;

– et d'autre part, une spécification suffisamment souple de la distribution des retards pour permettre de restituer tout profil d'évolution temporelle des coefficients.

Pour cela, S. ALMON a choisi d'exprimer la distribution des retards par une fonction polynomiale. En effet, on sait que pour toute suite de n points, il est possible de trouver une fonction polynomiale de degré $(n - 1)$ passant par tous les points, donc de rendre compte de profils fort complexes.

Soit le M.R.E.

$$Y_t = \sum_{i=IR1}^{IR2} \alpha_i X_{t-i} + \eta_t + B$$

où IR1 et IR2 sont respectivement les premier et dernier retards considérés,

η_t la variable d'erreur,

α_i le coefficient affectant la variable exogène X_{t-i} ,

B la constante.

On peut alors considérer que les coefficients α_i s'expriment par :

$$\alpha_i = \beta_0 + \beta_1 i + \dots + \beta_{q+1} i^{q+1}$$

$i = IR1 \dots IR2$,

où $q + 1$ sera le degré de polynome, avec $q + 2 \leq r$ et $r = IR2 - IR1$, $IR2 \geq 0$ (6)

Au lieu d'avoir r coefficients à estimer, il n'en faudra plus estimer que $q + 2$:

$$\beta_0 \dots \beta_{q+1}.$$

Toutes les valeurs que peut prendre la fonction des coefficients α_i se déduisent donc d'une combinaison linéaire de polynômes d'interpolation de LAGRANGE $\phi_j(i)$ de coefficients β_j :

$$\alpha_i = \sum_{j=0}^{q+1} \phi_j(i) \beta_j$$

avec

$$\phi_j(i) = \frac{\prod_{k=0, k \neq j}^{q+1} (i - k)}{\prod_{k=0, k \neq j}^{q+1} (j - k)}$$

(6) Remarquons, en fait, que S. ALMON suppose $IR1 \geq 0$.

dont l'usage permet une estimation plus commode des coefficients auxiliaires β_j si l'on impose des restrictions a priori supplémentaires (par exemple S. ALMON considère des situations avec $\alpha_{IR1-1} = \alpha_{IR2-1} = 0$).

Le MRE que permet d'estimer la MPA se réécrit donc :

$$Y_t = \sum_{j=0}^{q+1} \beta_j \sum_{i=IR1}^{IR2} \phi_j(i) X_{t-i} + \eta_t + B$$

Les estimations des coefficients β_j sont obtenues en appliquant la méthode des moindres carrés à cette équation transformée. Nous en déduisons alors l'estimation des α_i :

$$\hat{\alpha}_i = \sum_{j=0}^{q+1} \phi_j(i) \hat{\beta}_j$$

Si l'on note Z la matrice de terme générique $Z_{jt} = \sum_{i=IR1}^{IR2} \phi_j(i) X_{t-i}$, la matrice des variances et covariances des $\hat{\beta}_j$ sera égale à $\sigma_n^2(Z'Z)^{-1}$ et celle des $\hat{\alpha}_i$,

$$S^2 = \phi(i) \sigma_n^2 (Z'Z)^{-1} \phi'(i)$$

avec

$$\phi(i) = (\phi_0(i), \dots, \phi_{q+1}(i)), \sigma_n^2 \text{ désignant la variance de } \eta_t.$$

Cette méthode appelle plusieurs remarques :

– Faisant appel aux techniques de régression habituelles, on se retrouve confronté aux problèmes de celles-ci : hypothèse de stationnarité du modèle, estimateur à variance non minimale si il y a autocorrélation des erreurs, etc. . . Nous utiliserons comme cela se pratique couramment la méthode de COCHRANE-ORCUTT pour éliminer une éventuelle autocorrélation du premier ordre⁽⁷⁾.

– On remarque également que cette technique nécessite non seulement la spécification a priori du nombre de retards, mais aussi la spécification du degré du polynôme. Comme nous le verrons par la suite, le choix de ce dernier paramètre est particulièrement délicat et une mauvaise spécification peut introduire un biais important —qui n'est pas toujours décelé par les tests statistiques classiques— lors de l'estimation de certains coefficients.

1.2. La méthode de la régression spectrale

S'il est admissible et justifiable de déterminer au préalable le degré du polynôme, puis d'utiliser le critère de la variance résiduelle pour rechercher la meilleure estimation des coefficients, en revanche, il semble difficile de suivre la M.P.A. lorsqu'il s'agit de spécifier la structure (le profil) des retards.

(7) Nous montrerons les limites de cette méthode qui, par itérations successives, minimise la somme des carrés des résidus ; ce qui a pour effet une amélioration mécanique du R^2 mais n'entraîne pas systématiquement une meilleure estimation des coefficients. Pour des exemples d'application, H. MADINIER et M. MOUILLART (1981). Pour une présentation générale, E. MALINVAUD (1972).

Nous allons en effet montrer, en nous appuyant sur des simulations, que, ce faisant, la M.P.A. ne peut appréhender que certaines structures bien particulières de retards⁽⁸⁾.

En revanche, la M.R.S. permet de rechercher et d'estimer les coefficients d'un M.R.E. sans que soient posées des hypothèses contraignantes sur la structure des distributions de retards.

La M.R.S., à la différence de la M.P.A., s'accommode de l'estimation de M.R.E., sans aucune information a priori ni sur la distribution des retards ni sur la structure du terme aléatoire.

1.2.1. Rappels sur l'estimation des spectres et cospectres⁽⁹⁾

Soit une série temporelle stationnaire X_t satisfaisant aux conditions de DIRICHLET :

$$E(X_t) = 0$$

$$E(X_t^2) = \sigma^2$$

Sa fonction d'autocovariance admet comme représentation de CRAMER :

$$\mu_\tau = E(X_t \cdot X_{t-\tau}) = \int_{-\pi}^{\pi} e^{i\omega\tau} dF(\omega) \quad (10)$$

Puisque nous intéressons à un processus réel X_t , μ_τ peut s'écrire :

$$\mu_\tau = 2 \int_0^{\pi} \cos \tau\omega dF(\omega)$$

Puisque d'autre part, X_t n'est pas déterminable (il n'est pas possible de reproduire totalement X_t par un processus autorégressif), la fonction de densité spectral $dF(\omega)$ s'écrit :

$$dF(\omega) = f(\omega) d\omega$$

où précisément $f(\omega)$ est appelé spectre de puissance du processus et s'obtient en inversant l'équation précédente :

$$f(\omega) = \frac{1}{2\pi} \left(\mu_0 + 2 \sum_{\tau=1}^{\infty} \mu_\tau \cos \tau\omega \right) \quad \omega \in [-\pi, \pi]$$

En pratique, l'estimation des spectres nécessite l'estimation des fonctions d'autocovariance C_τ ⁽¹¹⁾ pour un nombre restreint de valeur de τ (en pratique entre 5 % du nombre d'observations de la série X_t (JENKINS); 15 % (PARZEN) et 30 % (GRANGER)). Ce nombre m correspond bien évidemment au nombre de bandes de fréquence sur lesquelles porte l'estimation.

(8) Sur l'estimation spectrale en économie et les problèmes qu'elle soulève, cf. GRANGER et HATANAKA, D. HERNARD et M. MOUILLART (1980), FISHMAN (1969).

(9) $F(\omega)$ est une fonction de saut. Les sauts ont pour amplitude σ_j^2 aux points $\omega = \omega_j$.

(10) $C_\tau = \frac{1}{n-\tau} \sum_{t=1}^{n-\tau} (x_t - \bar{x})(x_{t+\tau} - \bar{x})$

(11) Pour les questions d'interprétation et de mise en oeuvre pratique, nous renvoyons aux références précédemment mentionnées.

De plus, et d'après le théorème de PARZEN présenté dans P.J. DHRYMES (1971), il est nécessaire de lisser les estimations C_τ au moyen d'une fenêtre de décalage notée λ_τ . Finalement, l'estimation spectrale sera ⁽¹²⁾ :

$$\hat{f}(\omega_j) = \frac{1}{2\Pi} (C_0 \lambda_0 + 2 \sum_{\tau=1}^m C_\tau \lambda_\tau \cos \tau \omega_j)$$

$$\omega_j = \frac{\Pi j}{m}, j = 0 \dots m \text{ et } \lambda_\tau = 1 + \cos \frac{\Pi \tau}{m}$$

si l'on choisit la fenêtre de TUKEY-HAMMING (G.S. FISHMAN (1969).

En ce qui concerne maintenant l'estimation des cospectres⁽¹³⁾ la démarche théorique est comparable : le cospectre entre les séries X_t et Y_t , de spectre de puissance $f_x(\omega)$ et $f_y(\omega)$, est :

$$f_{xy}(\omega) = c(\omega) + iq(\omega) \quad \omega \in [-\Pi, \Pi]$$

où $C(\omega)$ représente la partie réelle du cospectre et $q(\omega)$ la partie imaginaire ou spectre de quadrature.

Ces fonctions vérifient toujours l'inégalité de cohérence :

$$c^2(\omega) + q^2(\omega) \leq f_x(\omega) \cdot f_y(\omega)$$

On appelle $C(\omega)$ la cohérence entre les séries X_t et Y_t pour la bande de fréquence ω :

$$C(\omega) = \frac{c^2(\omega) + q^2(\omega)}{f_x(\omega) \cdot f_y(\omega)}$$

Seule cette cohérence nous intéresse pour la suite de notre exposé. Elle s'interprète de la même manière que le carré du coefficient de corrélation linéaire, mais elle permet ici d'apprécier les liaisons bande de fréquence par bande de fréquence entre deux séries temporelles.

Enfin, comme nous avons estimé $\hat{f}(\omega_j)$, nous estimons :

$$\hat{c}(\omega_j) = \frac{1}{\Pi} \left(\frac{a_0 \lambda_0}{2} + \sum_{k=1}^{m-1} \lambda_k a_k \cos k \omega_j \right)$$

$$\hat{q}(\omega_j) = \frac{1}{\Pi} \sum_{k=1}^{m-1} \lambda_k b_k \sin k \omega_j$$

avec :

$$a_k = \frac{1}{n-k} \sum_{j=1}^{n-k} (X_j Y_{j-k} + Y_j X_{j-k})$$

$$b_k = \frac{1}{n-k} \sum_{j=1}^{n-k} (X_j Y_{j-k} - Y_j X_{j-k})$$

(12) Le cospectre correspond donc à la transformée de FOURIER des covariances théoriques croisées entre les séries X_t et Y_t . L'absence de parité des fonctions de covariance croisée justifie le caractère complexe du cospectre.

(13) Pour une représentation détaillée, D. HERNAD et M. MOUILLART (1980).

d'où nous déduisons :

$$\hat{C}(\omega_j) = \frac{\hat{c}^2(\omega_j) + \hat{q}^2(\omega_j)}{\hat{f}_x(\omega_j) \cdot \hat{f}_y(\omega_j)}$$

1.2.2. L'estimation par la MRS

La MRS se propose d'estimer les coefficients du modèle :

$$Y_t = B + \sum_{i=IR1}^{iR2} \alpha_i X_{t-i} + \eta_t$$

avec (1) X, Y et η , processus stationnaires et normaux, et de plus X et η indépendants :

$$f_{X\eta}(\omega) = 0 \quad \forall \omega \in [-\Pi, \Pi]$$

On cherche donc l'estimateur $\hat{A} = [\alpha_{1R1}, \dots, \alpha_{iR2}]$ de A. Pour ce faire, on utilise la méthode du maximum de vraisemblance. Si nous posons $E_T = [\eta_1 \dots \eta_T]$, avec η_t qui représente l'erreur du modèle et T le nombre d'observations, la densité de probabilité de E_T s'exprime :

$$f(E_T) = \frac{\exp \{(-1/2) E_T' C^{-1} E_T\}}{(2\Pi)^{T/2} |C|^{1/2}}$$

où C est la matrice des variances-covariances des η_t qui est définie positive.

En passant aux logarithmes, on cherche à minimiser les erreurs. On obtient alors \hat{A} :

$$\hat{A} = \hat{D} \cdot \hat{G}^{-1}$$

où la matrice G admet pour terme générique :

$$\hat{g}_{k,\ell} = \frac{1}{2m} \sum_{j=1-m}^m \hat{f}_\eta^{-1}(\omega_j) \hat{f}_{xy}(\omega_j) \exp(i\omega_j(\ell - k))$$

et le vecteur \hat{D} :

$$d_k = \frac{1}{2m} \sum_{j=1-m}^m \hat{f}_\eta^{-1}(\omega_j) \hat{f}_x(\omega_j) \exp(i\omega_j k)$$

$k, \ell = IR1 \dots IR2$

avec $\hat{f}_\eta(\omega_j) = \hat{f}_y(\omega_j) [1 - \hat{C}^2(\omega_j)]$

2. COMPARAISON DES METHODES D'ESTIMATION

La comparaison entre les méthodes s'est effectuée en recourant à la simulation de séries temporelles qui nous ont ensuite permis de construire des M.R.E. Nous avons alors comparé les estimations des coefficients avec les valeurs vraies de ces coefficients.

Cette démarche nous a permis d'apprécier réellement la précision de ces méthodes pour l'estimation des coefficients et pour la détermination de la structure des retards pour différents types de M.R.E. Ce qui est difficile voire impossible, quand on travaille sur des données réelles.

Nous allons donc présenter rapidement ces M.R.E. que nous avons construits afin de procéder à ces comparaisons.

2.1. Les M.R.E. analysés

Les séries temporelles que nous avons utilisées pour comparer les méthodes ont été obtenues en simulant un certain nombre de M.R.E. Ces M.R.E. ont tous la même structure :

$$Y_t = \sum_{i=IR1}^{IR2} \alpha_i X_{t-i} + \eta_t + B, \quad (1)$$

nous avons alors choisi $IR1 = 0$ et $IR2 = 6$, et défini des valeurs pour les coefficients α_i et les variables X et Y.

2.1.1. Les distributions des retards retenues

Distribution	Retards i (IR1)	0	1	2	3	4	5	6 (IR2)
α^1		2.	-3.	1.2	0.	3.	-4.	17
α^2		.1	.3	.5	.7	.5	.3	.1
α^3		.8	.64	.512	.410	.328	.262	.210

Les distributions α^2 et α^3 sont théoriquement facilement identifiables par les M.C.O. ou la M.P.A. ; elles peuvent être représentées par des fonctions polynomiales de degré peu élevé ou par des progressions géométriques. Nous les qualifierons d'orthodoxes.

En revanche, α^1 est étrangère à ce que le "bon sens" ou la théorie économique recommande de considérer. Nous la qualifierons d'hétérodoxe.

2.1.2. Forme de la variable explicative et de la variable aléatoire

Nous avons reconstitué la variable explicative X en la composant par une fonction temporelle simple notée g(t), à laquelle nous avons adjoint une fonction aléatoire $f_t(\xi)$:

$$X_t = g(t) + f_t(\xi) \quad (2)$$

avec :

$$f_t(\xi) = \sum_{j=0}^{K1-1} \beta_j \xi_{t-j} \quad \xi_t \sim \mathcal{N}(0, 2)$$

$$\text{et } E(\xi_t \cdot \xi_{t-k}) = 0 \quad \text{si } k \neq 0$$

Lorsque $K1 = 1$, $f(\xi)$ n'est rien d'autre qu'une variable aléatoire gaussienne. Lorsque $K1 > 1$, $f(\xi)$ est une fonction aléatoire dont la longueur de la

mémoire sera fonction croissante de K1 (D. HERNARD, M. MOUILLART et D. STRAUSS-KAHN (1978)).

De la même façon, nous avons défini η_t sous la forme :

$$\eta_t = \sum_{e=0}^{K2-1} \gamma_e \epsilon_{t-e} \quad (3)$$

$$\epsilon_t \sim \mathcal{G}(0, 1)$$

$$\text{et } E(\epsilon_t \cdot \epsilon_{t-k}) = 0 \quad \text{si } k \neq 0$$

Lorsque $K2 = 1$, η_t est une variable aléatoire ϵ_t . Sinon, lorsque $K2 > 1$, une autocorrélation des résidus apparaît dont l'effet sera d'autant plus néfaste sur les estimations que l'effet de mémoire $-K2-$ est grand.

De nombreuses structures de distributions pour β et γ et un grand nombre de couples de valeur pour K1 et K2 ont été testés. Les résultats sont très stables; nous ne retenons ici que deux configurations extrêmes :

- $K1 = K2 = 1$, avec $\beta_1 = 1$ et $\gamma_1 = 1$.
- $K1 = K2 = 51$, avec deux types de distributions pour β_j et γ_j .

La première est telle que $\beta_j = \gamma_j = 1$ pour $j = 1 \dots 51$.

La seconde est plus proche du cas que se proposent de traiter les méthodes d'élimination de l'autocorrélation du type de celle d'HILDRETHLIU ou de COCHRANE-ORCUTT :

$$\beta = (.5, .2, .5, .8, 1.1, .9, .7, .5, .3, -.1, -.3)$$

Modulo 11

$$\gamma = (.1, .2, .3, .4, .5, .6, .7, .8, .9, 1., 1., .9, .8, .7, .6, .5, .4, .3, .2, .1)$$

Modulo 20

Les analyses que nous avons ainsi conduites nous ont permis de constater que seule la présence d'un effet de mémoire était un critère discriminant dans nos résultats : la forme des distributions β et γ est sans importance⁽¹⁴⁾.

Il n'en va pas de même en ce qui concerne la forme de la distribution α . Celle-ci et la présence d'un effet de mémoire vont nous permettre de structurer la présentation des résultats que nous allons maintenant aborder. Mais, avant de ce faire, rappelons la structure générale du M.R.E. simulé :

$$Y_t = \sum_{i=IR1}^{IR2} \alpha_i [g(t-i) + \sum_{j=0}^{K1-1} \beta_j \xi_{t-i-j}] + \sum_{l=0}^{K2-1} \gamma_l \epsilon_{t-l} + B$$

dont nous cherchons à estimer la distribution α .

(14) Remarquons que tous les coefficients des polynômes sont significatifs et que leur degré de significativité sera d'autant plus élevé que le degré du polynôme est lui-même élevé.

2.2. Comparaison des simulations effectuées sur des séries sans mémoire

2.2.1. Distribution hétérodoxe α^1

Pour ces types de modèle et de distribution, les moindres carrés ordinaires fournissent une bonne estimation des coefficients (tableau 1). Sauf, bien sûr si le nombre de retards retenu à fin d'estimation est plus (moins) important que le nombre "exact", et que nous utilisons ou non la méthode de COCHRANE-ORCUTT, afin d'éliminer une éventuelle autocorrélation des résidus.

Ainsi, si nous spécifions correctement le nombre de retards la méthode restitue bien les coefficients de la distribution α^1 = coefficient hautement significatif, R^2 élevé, absence d'autocorrélation des résidus. . .

Afin d'estimer les coefficients avec la M.R.S., nous avons fait varier les nombres de bandes de fréquence et de coefficients (tableau 1). Il ressort des comparaisons que nous avons effectuées que d'une part, les résultats sont peu sensibles au choix du nombre de bandes de fréquence, que d'autre part, et comme en ce qui concerne la M.C.O., la méthode permet de sélectionner le "bon" nombre de coefficients, en utilisant pour ce faire, les indicateurs traditionnels (écarts-types des coefficients, R^2 , DW); et qu'enfin, le profil très perturbé de ce type de distribution est bien restitué par la méthode.

Avec la M.P.A., nous devons faire varier le degré du polynôme et le nombre de coefficients (tableau 2). Nous savons que pour la M.P.A., le profil des coefficients est fortement influencé par le degré du polynôme. Il n'est possible de restituer un bon profil que si le degré du polynôme est supérieur ou égal au degré de la courbe approximant les coefficients introduits. Or, si, comme le préconise S. ALMON, nous utilisons le critère du R^2 pour retenir la meilleure estimation, nous sommes amenés à retenir un polynôme de degré 5 qui ne peut restituer un bon profil puisque la série de coefficients hétérodoxes correspond à une courbe de degré 6. L'estimation des coefficients est donc très mauvaise.

Si nous utilisons un polynôme de degré 8, non seulement le profil de la distribution n'est pas restitué, mais en outre, les écarts type augmentent fortement, ôtant toute significativité statistique aux valeurs des coefficients.

Il est donc difficile de choisir le nombre de retards, car, contrairement aux autres méthodes déjà étudiées, et pour lesquelles les coefficients devenaient négligeables si l'on dépasse le nombre de coefficients introduits, nous remarquons que toute sur-spécification du nombre de coefficients perturbe totalement l'estimation de ceux-ci.

En conclusion, pour une série sans mémoire et avec une distribution hétérodoxe des coefficients, la M.P.A. est celle qui fournit les moins bons résultats et qui présente le plus grand nombre d'inconvénients lorsqu'il faut choisir entre différentes spécifications d'un M.R.E.

En effet, pour les M.C.O., seul le nombre de retards doit être spécifié et la comparaison entre les différentes estimations se fait facilement: pour la M.R.S., le choix du nombre de retards et de bandes de fréquence influence peu l'estimation des coefficients qui reste toujours satisfaisante, alors que pour la M.P.A., le choix des degrés du polynôme et du nombre de retards est déterminant pour la qualité des estimations.

TABLEAU 1

Estimation des coefficients de la distribution α^1 (absence de mémoire)

Comme pour l'ensemble des résultats présentés dans cet article, nous avons reporté la valeur du coefficient estimé et, entre parenthèses, l'écart-type associé

Méthode des MCO	Retards	0	1	2	3	4	5	6	7	8	9	R ²	DW			
	α_j	2	-3	1.2	0	3	-4	17	0	0	0					
	4 coefficients estimés	-0.877 (1.822)	-1.958 (1.821)	-0.430 (1.820)	.189 (1.798)							.016	2.35			
	7 coefficients estimés	2 (.009)	-2.994 (.009)	1.216 (.009)	.001 (.009)	3 (.009)	-3.998 (.009)	17.01 (.009)				1.	1.79			
	10 coefficients estimés	IDEM										.016 (1.600)	.008 (.872)	.002 (.185)	1.	1.83
MRS (30 bandes de fréquence)	4 coefficients estimés				.277 (.267)	2.70 (.267)	-3.81 (.267)	16.9 (.267)					.963			
	7 coefficients estimés	1.368 (.221)	-2.321 (.235)	1.357 (.231)	.422 (.225)	2.60 (.231)	-3.73 (.235)	16.83 (.221)				.997				
	10 coefficients estimés	1.47 (.242)	-2.40 (.250)	1.44 (.250)	.333 (.242)	2.66 (.240)	-3.74 (.236)	16.83 (.221)	-2.44 (.240)	.136 (.236)	-1.07 (.221)	.998				

TABLEAU 2
Estimation de la distribution α^1 par la MPA (absence de mémoire)

$\hat{\rho}$: coefficient d'autocorrélation de COCHRANE-ORCUTT.

Entre parenthèses figurent les écarts-type des coefficients estimés.

Pour chaque couple (degré du polynôme-nombre de coefficients estimés), nous avons réalisé l'estimation d'abord sans le recours à la méthode de COCHRANE-ORCUTT (A) puis avec (B), afin d'éliminer l'autocorrélation des résidus qui est indûment décelée dans le premier cas (statistique DW).

Retards	0	1	2	3	4	5	6	7	8	9	R ²	DW	$\hat{\rho}$
α_j	2	-3	1.2	0	3	-4	17	0	0	0			
Degré du polynôme													
Nbr de coeff. estimés													
2	-1.013 (1.773)	-1.404 (1.329)	-.950 (1.328)	.352 (1.754)							0	2.35	A
4	-.802 (1.702)	-1.678 (1.212)	-.994 (1.212)	1.252 (1.704)							.030	1.88	B (-.202 (0.98))
5	1.600 (.440)	-2.205 (.431)	-.809 (.389)	2.673 (.348)	.969 (.436)	-3.261 (.551)	16.65 (.696)				.942	3.63	A
7	1.244 (.190)	-1.569 (.204)	-.816 (.158)	2.206 (.130)	.951 (.187)	-2.533 (.263)	16.23 (.307)				.986	2.80	B (-.873 (.049))
10	-.508 (1.51)	.880 (1.298)	-1.278 (.941)	-1.216 (.994)	1.717 (.896)	5.255 (.895)	6.413 (.981)	3.689 (.884)	-.74 (1.169)	1.182 (1.225)	.316	3.23	A
8	-.504 (-.940)	1.280 (.721)	-1.001 (.470)	-1.271 (.501)	1.432 (.537)	5.058 (.459)	6.640 (.537)	4.394 (.580)	-.179 (.875)	-.197 (1.134)	.649	2.18	B (-.710 (.038))
10	-.089 (1.175)	-1.310 (1.100)	-.975 (1.757)	4.076 (5.152)	-2.807 (10.14)	1.93 (16.50)	12.58 (20.50)	2.069 (25.62)	-1.876 (96.19)	2.244 (225.2)	.590	3.69	A
	.578 (.388)	-.675 (.534)	-2.151 (1.863)	4.326 (3.908)	-1.910 (6.039)	1.185 (3.324)	11.75 (23.96)	3.694 (57.09)	-3.205 (111.5)	1.185 (193.5)	.939	2.89	B (-.925 (.038))

2.2.2. Distributions orthodoxes α^2 et α^3

Dans les cas de distribution sans mémoire, les M.C.O. semblent peu sensibles aux types des distributions étudiées⁽¹⁵⁾. Les estimations qu'ils permettent d'obtenir sont en général très bonnes lorsque le nombre de retards a été correctement spécifié. Dans le cas contraire, les M.C.O. permettent de sélectionner le "bon" nombre de retards à l'aide des indicateurs habituels (R^2 , DW. . .)

La M.R.S. fournit aussi de bons résultats, quel que soit le nombre de bandes de fréquence : les profils des distributions sont bien restitués et les valeurs des coefficients sont justement évaluées.

Ce type de distribution des retards particulièrement simple (et l'absence de mémoire sur la série) devait permettre d'obtenir de très bons résultats par la M.P.A. Quel que soit le degré du polynôme choisi, le profil des coefficients est bien restitué. La qualité des estimations est bonne, mais reste inférieure à celle des M.C.O. D'autre part, si l'on augmente le degré du polynôme et le nombre des retards, alors que le R^2 et le DW restent constants, on constate que les coefficients au delà du septième ne sont pas négligeables, et pour un degré du polynôme égal à 8, l'apparition d'écarts types élevés ôte toute signification aux coefficients estimés.

En conclusion, au vu de ces résultats d'estimations sur une série sans mémoire, l'utilisation de la M.P.A. ne se justifie en aucun cas. Les M.C.O. et M.R.S. permettent d'obtenir des résultats au moins aussi bons pour les distributions de coefficients simples et toujours meilleurs pour des distributions de coefficients plus complexes.

En outre, ces deux méthodes sont plus fiables face à une mauvaise spécification des paramètres. Ce point est particulièrement important pour tous les économistes qui ne connaissent pas, en général et a priori, la longueur et la forme des retards dans leur modèle.

2.3. Comparaisons de simulations effectuées sur des séries avec mémoire

Nous nous proposons ici d'étudier un M.R.E. dont la forme est :

$$Y_t = \sum_{i=IR1}^{IR2} \alpha_i [g(t-i) + \sum_{j=0}^{K1-1} \beta_j \xi_{t-i-j}] + \sum_{l=0}^{K2-1} \gamma_l \epsilon_{t-l} + B$$

et dont nous chercherons à estimer les coefficients α_i .

Pour construire la série simulée, nous avons retenu deux types de distributions pour β_j et γ_l :

– cas I : la première est la plus simple

$$\beta_j = \gamma_j = 1 \quad j = 1 \dots 51$$

– cas II : la seconde se présente sous une forme plus complexe⁽¹⁶⁾.

(15) Les résultats sont, pour les distributions α^2 et α^3 , résumés par le tableau 3 : seules sont retenues les estimations de 7 coefficients. Nous pouvons ainsi alléger les commentaires qui, pour l'essentiel, demeurent proches de ceux présentés pour α^1 . Remarquons cependant que ce faisant, nous ne présentons que les "meilleures estimations".

(16) Pour la valeur des coefficients, se reporter en 2.1.2.

TABLEAU 3
Absence de mémoire (K1 = K2 = 1) : 100 observations

Distribution α^1	Méthode	Retards α_j	0						1						2						3						4						5						6						R ²	DW	$\hat{\rho}$
			2.						-2.994 (.009)						1.216 (.009)						.001 (.009)*						3.						-3.998 (.009)						17.01 (.009)								
(Rappel des tableaux 1 et 2)	M.C.O.		2.						-2.994 (.009)						1.216 (.009)						.001 (.009)*						3.						-3.998 (.009)						17.01 (.009)						1.	1.79	
	M.P.A. (5,7)		1.6 (.440)						-2.205 (.431)						-.809 (.389)						2.673 (.348)						.969 (.436)						-3.261 (.551)						16.65 (.696)						.942	3.63	
	M.R.S.		1.368 (.221)						-2.321 (.235)						1.357 (.231)						.442 (.285)*						2.60 (.231)						-3.73 (.235)						16.83 (.221)						.997		
Distribution α^2	Méthode	α_j	.1						.3						.5						.7						.5						.3						.1						R ²	DW	$\hat{\rho}$
	M.P.A. (5,7)		.105 (.018)						.3 (.023)						.552 (.024)						.646 (.024)						.531 (.024)						.287 (.024)						.103 (.020)						.973	3.14	.937 (.035)
	M.R.S.		.124 (.023)						.284 (.041)						.483 (.046)						.614 (.046)						.356 (.046)						.168 (.041)						.052 (.023)						.959		
	Méthode	α_j	.8						.64						.512						.410						.328						.262						.210						R ²	DW	$\hat{\rho}$
	M.C.O.		.800 (.009)						.646 (.009)						.528 (.009)						.411 (.009)						.327 (.009)						.264 (.009)						.217 (.009)						.995	1.79	
	M.P.A. (5,7)		.790 (.012)						.638 (.015)						.518 (.016)						.407 (.015)						.314 (.015)						.250 (.013)						.201 (.001)						.992	2.64	.976 (.022)
	M.R.S.		.826 (.022)						.617 (.028)						.481 (.033)						.374 (.033)						.252 (.033)						.178 (.028)						.125 (.022)						.986		

* non significatif au seul de 5 %

Nous procéderons donc à deux séries d'estimations pour chaque type de distribution. Nous ne présenterons ici que les meilleures estimations pour chacune des méthodes.

Seules les estimations effectuées dans le cas d'une distribution hétérodoxe α^1 sont sensibles au choix des distributions β_j et γ_1 (cas I ou II). Nous ne présentons donc que les résultats obtenus dans le cas II en ce qui concerne les distributions orthodoxes.

Les limites des M.C.O. apparaissent clairement dès que l'on introduit une mémoire sur les séries (tableau 4). Les plus mauvais résultats sont obtenus dans le cas I où nous constatons simultanément une impossibilité de reconstituer le profil des coefficients, l'apparition de très mauvais écarts-types, ainsi qu'un R^2 faible, même pour les M.C.O. avec COCHRANE-ORCUTT. Dans les cas II, les estimations s'améliorent pour tous les types de distribution mais les écarts-types restent importants si l'on n'utilise pas la méthode de COCHRANE-ORCUTT.

Dans le cas I les résultats sont aussi mauvais pour la M.P.A. que pour les M.C.O. aussi bien pour la représentation du profil que pour la qualité des estimations. Dans le cas II les estimations de la M.P.A. se dégradent nettement par rapport à celles des M.C.O. Pour la distribution hétérodoxe α^1 alors que les M.C.O. restituent correctement le profil, la M.P.A. n'arrive pas à estimer certains coefficients; de plus, les écarts-types sont beaucoup plus élevés que l'on utilise COCHRANE-ORCUTT ou non; enfin, le R^2 est bon alors que les estimations ne le sont pas.

Pour les estimations effectuées sur les séries orthodoxes (α^2, α^3), les résultats sont sensiblement les mêmes si l'on utilise COCHRANE-ORCUTT.

Il apparaît donc que, aussi bien pour les séries sans mémoire que pour celles avec mémoire, nous obtenons de meilleures estimations avec les M.C.O. qu'avec la M.P.A. Mais pour les séries avec mémoire, ces estimations sont comme on pouvait l'attendre, beaucoup moins bonnes, et parfois inutilisables (cas I, série α^1): il est donc intéressant d'étudier les résultats fournis par la M.R.S. dans tous les cas que nous venons de présenter.

D'une manière générale, et quel que soit le type de distribution considéré, la M.R.S. fournit des résultats d'une qualité supérieure à ceux que l'on obtient à l'aide des autres méthodes (et même supérieure à ceux obtenus lorsque les séries étaient sans mémoire) (tableau 4).

Cette conclusion générale doit cependant être nuancée. Et c'est précisément en envisageant cette nuance que nous allons pouvoir présenter les caractéristiques de la M.R.S., tant du point de vue de son utilisation que des résultats auxquels elle conduit; si la M.R.S. semble meilleure que les autres, cette qualité sera d'autant plus marquée que :

- le nombre de bandes de fréquence utilisé afin d'estimer les spectres sera grand⁽¹⁷⁾
- les distributions seront plus hétérodoxes
- et enfin, le nombre d'observations sera important.

En ce qui concerne la longueur des séries analysées, considérons le cas de la distribution α^2 . Quel que soit le nombre de bandes de fréquence utilisé, et malgré la forte valeur du R^2 , la qualité des estimations fournies par la M.R.S.

(17) En effet, et comme en ce qui concerne les séries sans mémoire, plus le nombre de bandes de fréquence est grand (dans les limites du possible) et meilleure sera l'estimation puisqu'alors les fuites entre bandes de fréquence seront faibles.

TABLEAU 4

Effet de mémoire (K1 = K2 = 51) : 100 observations

Distribution α^1	Méthode	2.	-3.	1.2	0.	3.	-4.	17.	R ²	DW	$\hat{\rho}$
cas I	M.C.O.	.195 (1.965)*	3.364 (2.123)*	-.448 (2.120)*	-1.473 (2.123)*	-1.205 (2.121)*	-.114 (2.125)*	1.464 (1.967)*	.317	2.20	.515 (.090)
	M.P.A.	1.557 (1.877)*	4.070 (1.931)	.074 (1.703)*	-1.691 (1.240)*	-.767 (.888)*	-.050 (1.535)*	-.854 (3.177)*	.252	2.25	.604 (.080)
	M.R.S.	1.47 (.35)	-3.1 (.497)	1.01 (.498)	-.07 (.501)*	3. (.498)	-3.9 (.497)	17.2 (.358)	.999		
Distribution α^1	Méthode	2.	-3.	1.2	0	3.	-4.	17.	R ²	DW	$\hat{\rho}$
cas II	M.C.O.	2.039 (.076)	-3.003 (.072)	1.113 (.073)	-.054 (.073)	3.010 (.074)	-4.012 (.073)	17.074 (.078)	.999	1.46	.929 (.037)
	M.P.A.	2.304 (.361)	-2.687 (.435)	-.926 (.373)	2.907 (.298)	.959 (.459)	-3.664 (.661)	17.25 (.935)	.991	2.56	.408 (.092)

M.R.S. : les résultats sont semblables à ceux obtenus dans le cas I.

Distribution α^2	Méthode	α_1	.1	.3	.5	.7	.5	.3	.5	.7	.5	.3	.1	R ²	DW	$\hat{\rho}$
cas I	M.C.O.	.139 (.076)*	.297 (.072)	.413 (.073)	.646 (.073)	.510 (.073)	.288 (.073)	.174 (.078)	.984	1.46	.929 (.038)					
	M.P.A.	.118 (.073)*	.286 (.07)	.445 (.062)	.586 (.058)	.545 (.092)	.279 (.143)*	.154 (.221)*	.983	1.53	.975 (.022)					
	M.R.S.	.154 (.048)	.346 (.048)	.520 (.047)	.664 (.046)	.420 (.047)	.169 (.048)	.006 (.048)*	.991							
Distribution α^3	Méthode	α_1	.8	.64	.512	.410	.328	.262	R ²	DW	$\hat{\rho}$					
cas II	M.C.O.	.859 (.076)	.637 (.072)	.425 (.073)	.356 (.073)	.338 (.074)	.249 (.073)	.284 (.078)	.989	1.46	.929 (.097)					
	M.P.A.	.833 (.073)	.633 (.069)	.417 (.062)	.353 (.056)	.333 (.080)	.249 (.106)	.278 (.129)	.988	1.51	.975 (.022)					
	M.R.S.	.866 (.034)	.669 (.064)	.527 (.064)	.401 (.065)	.277 (.064)	.171 (.064)	.089 (.034)	.994							

TABLEAU 5
Caractéristiques de la M.R.S. (Distribution α^2)
K1 = K2 = 51, cas II

100 points séries brutes	.1	.3	.5	.7	.5	.3	.1	R ²
48 bandes de fréquence	.077 .098	-.312 .098	.431 .096	.791 .100	.892 .096	.036 .098	.194 .098	.983
30 bandes de fréquence	-.033 .104	-.417 .104	.315 .103	.728 .107	.854 .103	.228 .103	.339 .104	.969
24 bandes de fréquence	-.086 .105	-.453 .105	.241 .105	.663 .109	.815 .105	.318 .105	.432 .105	.955
400 points séries brutes								
48 bandes de fréquence	.007 .049	.169 .048	.421 .047	.664 .047	.520 .047	.346 .048	.154 .049	.991
100 points diff. premières								
48 bandes de fréquence	.107 .021	.275 .028	.456 .032	.632 .033	.472 .032	.299 .028	.105 .021	.987
400 points diff. premières								
48 bandes de fréquence	.082 .009	.270 .013	.463 .016	.657 .016	.467 .016	.283 .013	.098 .009	.988

Caractéristiques de la M.R.S.
(K1 = K2 = 51, cas II)
48 bandes de fréquence

α^3	.8	.64	.512	.410	.328	.262	.210	R ²
100 points diff. premières	.817 .017	.637 .020	.483 .022	.374 .023	.296 .023	.239 .030	.163 .017	.992
α^1	17	-4	3	0	1,2	-3	2	R ²
100 points diff. premières	16.853 .260	-3.573 .286	2.341 .289	-.024 .283	1.021 .289	-2.106 .286	1.579 .260	.996

est mauvaise lorsque l'effet de mémoire est important (tableau 5) et le nombre d'observations peu élevé. Si nous augmentons la longueur des séries analysées, la qualité des estimations s'améliore jusqu'à devenir convenable si nous accroissons simultanément le nombre de bandes de fréquence.

Si nous transformons maintenant les séries par le filtre "différences premières", nous vérifions que les résultats s'améliorent considérablement dans tous les cas envisagés (tableau 5) : les estimations sont même très bonnes et les R² se sont accrus.

L'explication de cette évolution se trouve dans l'analyse de la cohérence carrée entre variable explicative et résidus : cette cohérence doit en théorie être nulle pour toute bande de fréquence.

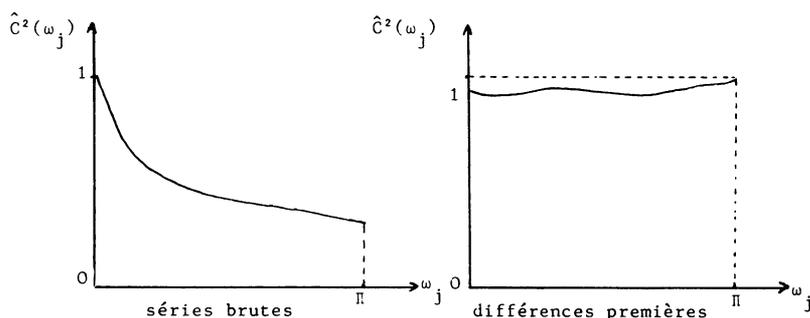
En réalité, tel n'est pas le cas. Quelle que soit la longueur des séries analysées, —si celles-ci sont en valeurs brutes—, la cohérence est élevée sur les basses fréquences, décroît lentement et éventuellement se relève sur les hautes fréquences.

Une analyse des spectres des séries expliquée et explicative nous montrerait alors (et ceci est bien sûr évident compte tenu du schéma de la construction) que ces séries exhibent une forte puissance à la fréquence zéro. La forme typique qui les caractérise a pour conséquence une cohérence élevée et maximale autour de la fréquence zéro et pour les fréquences adjacentes, compte tenu des fuites qui caractérisent les fenêtres spectrales utilisées.

Il en découle, si l'on se rapporte aux définitions analytiques des estimateurs spectraux, que les coefficients seront mal estimés. En effet, puisque $\hat{f}_\eta(\omega_j)$ est estimé par $\hat{f}_y(\omega_j) [1 - \hat{C}^2(\omega_j)]$ l'usage d'estimations biaisées de \hat{c} par suite de la présence de formes typiques ne pourra qu'entraîner une mauvaise estimation des coefficients et par suite, une cohérence élevée entre variable explicative et résidus.

Cohérence entre variable explicative et résidus élevée, forme typique des spectres des séries analysées et mauvaise estimation des coefficients vont de pair.

La réduction de la forme typique par un filtrage approprié (les différences premières dans le cas présent) permet de réduire le biais dans les estimations et d'améliorer la qualité des représentations. En effet, en réduisant la forme typique, nous pourrions constater par l'étude de la cohérence carrée entre variables expliquée et explicative que l'importance de la fréquence zéro a décliné. Mais ceci ne signifie pas qu'elle a décliné en niveau absolu : l'atténuation de l'effet de mémoire sur les séries a provoqué une diminution relative dans la cohérence en 0, les cohérences des autres bandes de fréquence retrouvant leur niveau théorique normal (graphique 1).



Graphique 1. – Cohérence carrée entre variables expliquée et explicative

Le filtrage préalable, en réduisant la forme typique de la cohérence, permet une meilleure estimation des coefficients et donc une réduction de la cohérence entre variable explicative et résidus.

L'utilisation préalable des analyses spectrale et cospectrale et l'étude des résultats obtenus par la M.R.S. devraient donc permettre de spécifier de façon efficiente les transformations préalables que l'on fait subir aux séries analysées. En présence de formes typiques marquées, les différences premières semblent souhaitables. En dehors de ce cas, les simulations que nous avons étudiées semblent montrer que tout filtrage préalable aura pour seule conséquence de réduire la valeur de R^2 sans pour autant améliorer la qualité des estimations.

En bien sûr, les résultats seront d'autant meilleurs que la longueur des séries sera accrue.

3. REMARQUES CONCLUSIVES

Au terme de ces comparaisons de méthodes, nous pouvons rappeler les quatre interrogations que nous avons soulevées au début de ce travail :

- a) Les méthodes d'estimation des M.R.E. permettent-elles d'appréhender la forme exacte de la distribution des coefficients à estimer ?
- b) Les coefficients estimés sont-ils en nombre exact et significativement estimés lorsqu'ils existent ?
- c) Le nombre de ces coefficients peut-il facilement être décelé ?
- d) La qualité générale des résultats (R^2 , DW. . .) est-elle satisfaisante ?

Afin de leur donner une réponse, nous avons construit des M.R.E. que nous avons ensuite estimés à l'aide de chacune des méthodes. La comparaison des résultats obtenus nous a permis de souligner les imperfections et les avantages de chaque méthode, dans un grand nombre de situations et pour des configurations des fonctions aléatoires fort diverses.

Nous pouvons finalement regrouper nos conclusions dans le tableau suivant, où chaque méthode a été notée, compte tenu des résultats que nous avons présentés.

TABLEAU 6
Qualité des résultats des estimations
(très bonne, bonne, moyenne, médiocre, mauvaise)

Méthode d'estimation		CARACTERISTIQUES DES SERIES ANALYSEES			
		Absence d'effets de mémoire		Présence d'effets de mémoire	
		Distribution α		Distribution α	
		Orthodoxe	Hétérodoxe	Orthodoxe	Hétérodoxe
M.C.O.	A	très bonne	très bonne	bonne	moyenne
	B	très bonne	très bonne	bonne	médiocre
	C	bonne	bonne	moyenne	moyenne
	D	très bonne	très bonne	bonne	médiocre
M.P.A.	A	moyenne	mauvaise	moyenne	mauvaise
	B	bonne	moyenne	moyenne	mauvaise
	C	très bonne	bonne	bonne	médiocre
	D	très bonne	bonne	bonne	médiocre
M.R.S.	A	très bonne	très bonne	très bonne	très bonne
	B	bonne	très bonne	bonne	très bonne
	C	très bonne	très bonne	très bonne	très bonne
	D	bonne	très bonne	très bonne	très bonne

A : reconnaissance par la méthode de la forme exacte de la distribution des coefficients

C : Capacité de la méthode à reconnaître le nombre exact des coefficients à estimer

B : capacité de la méthode à estimer significatifs les bons coefficients

D : qualité générale des résultats (R^2 , DW . . .)

D'une manière générale, la M.R.S. semble meilleure que les M.C.O., et largement préférable à la M.P.A. Il apparaît donc que l'utilisation de la M.P.A. doit être déconseillée: c'est la seule méthode à estimer significatifs des coefficients qui n'existent pas.

Entre la M.R.S. et les M.C.O., les performances sont comparables lorsqu'il s'agit de distributions de coefficients orthodoxes. En revanche, les M.C.O. semblent peu propices à la détection et à l'estimation de distributions hétérodoxes, surtout lorsque les effets de mémoire sont importants. Et surtout les M.C.O. nécessitent-ils un grand nombre d'analyses pour parvenir aux résultats, chaque analyse précédant pour un nombre de retards différents et l'utilisateur sélectionnant finalement la meilleure d'entre elles.

Ainsi, dans la mesure où l'on ne connaît pas, a priori, la structure exacte des M.R.E. que l'on estime sur séries réelles (caractéristiques de α , longueur de la mémoire. . .), l'utilisation des M.C.O. ou de la M.P.A. risque de conduire à des résultats erronés. Par contre, et puisqu'elle s'accorde avec la présence d'effets de mémoire et de distribution hétérodoxe, la M.R.S. devrait être largement utilisée.

Cette utilisation sera d'autant plus fructueuse, pour l'économiste, qu'elle sera doublée du recours aux analyses spectrales et cospectrales⁽¹⁸⁾. En effet, la présence de formes typiques des spectres, et donc l'existence de dépendance temporelle dans les séries, risque de conduire à des résultats difficilement interprétables si l'on n'a pas, au préalable, filtré les séries étudiées. L'étude des spectres et des cohérences entre séries est ici d'un précieux secours et indique la nécessité du filtrage préalable.

BIBLIOGRAPHIE

- S. ALMON (1965). — The Distributed lag Between Capital Appropriations and Expenditures, *Econometrica*, janvier, Vol. 33, n° 1, p. 178-196.
- P.J. DHRYMES (1971). — *Distributed Lags*, Oliver and Boyd, Edimburgh.
- G.S. FISHMAN (1969). — *Spectral Methods in Econometrics*, Harvard University Press, Cambridge.
- C.W. GRANGER et M. HATANAKA (1969). — *Analyse spectrale des séries temporelles en économie*, Dunod.
- Z. GRILICHES (1967). — Distributed Lags: A Survey, *Econometrica*, janvier, Vol. 35, p. 16-49.
- D. HERNAD, M. MOUILLART, D. STRAUSS-KAHN (1978). — Du bon usage de R/S, *Revue Statistique Appliquée*, n° 4, p. 61-80.
- D. HERNARD, M. MOUILLART, D. STRAUSS-KAHN (1979). — Forme typique du spectre et dépendance temporelle en économie, *Revue de Statistique Appliquée*, n° 4, p. 49-76.

(18) Une application économique a été réalisée dans le domaine de la production de logements neufs. Les résultats de cette étude peuvent être obtenus auprès des auteurs.

- D. HERNARD et M. MOUILLART (1980). – Estimation spectrale des modèles à retards échelonnés en économie, *Cahiers Economiques de Nancy*, n° 2-3, p. 89-138.
- D.W. JORGENSON (1966). – Rational Distributed Lag Functions, *Econometrica*, Vol. 34, n° 1, janvier, p. 135-149.
- H. MADINIER et M. MOUILLART (1981). – *Estimation des modèles à retards échelonnés: Quelques éléments méthodologiques*, ronéoté, 13 pages – Contribution au Colloque *Structures Economiques et Econométrie*, 21 mai 1981.
- E. MALINVAUD (1972). – *Méthodes statistiques de l'économétrie*, Dunod.
- P.A. MUET (1979). – La modélisation macroéconomique : une étude de la structure et de la dynamique des modèles macroéconométriques, *Statistiques et Etudes Financières, Série Orange*, Hors série 1979.
- P.A. MUET et ZAGAME (1976). – Fonction d'investissement et retards échelonnés, *Annales de l'INSEE*, n° 21, janvier-mars, p. 85-133.