Journal de la société statistique de Paris

Georges Prat

A propos de la structure autorégressive des résidus d’un modèle linéaire : deux interprétations économiques

Journal de la société statistique de Paris, tome 134, nº 1 (1993), p. 73-85

<http://www.numdam.org/item?id=JSFS_1993__134_1_73_0>

© Société de statistique de Paris, 1993, tous droits réservés.

L’accès aux archives de la revue « Journal de la société statistique de Paris » (http://publications-sfds.math.cnrs.fr/index.php/J-SFdS) 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/

A PROPOS DE LA STRUCTURE AUTORÉGRESSIVE DES RÉSIDUS D'UN MODÈLE LINÉAIRE : DEUX INTERPRÉTATIONS ÉCONOMIQUES

par Georges PRAT [1]
C.N.R.S., Université de Paris-X Nanterre

Cette brève note présente quelques réflexions à propos de l'interprétation **économique** pouvant être attribuée à un phénomène très fréquemment rencontré par les économètres travaillant sur des séries temporelles : la structure autorégressive d'ordre un des résidus d'un modèle linéaire. Lorsqu'il est en présence de ce phénomène, l'économètre conclut le plus souvent que le modèle est probablement mal spécifié, soit dans ses relations [2], soit en raison de variables omises [3]. En outre, lorsque les résidus issus d'un ajustement MCO sont autocorrélés, les tests usuels de signification des paramètres ne sont plus applicables car, si les espérances des paramètres restent en principe valables, leurs variances sont par contre généralement sous-estimées : bien que sans biais les estimateurs ne sont pas efficaces. C'est pourquoi, lorsqu'on ne parvient pas à déceler la nature du défaut de spécification, un certain nombre de méthodes sont malgré tout utilisées pour améliorer la précision des valeurs estimées des paramètres du model. Parmi ces méthodes, la plus usitée consiste à estimer **conjointement** les paramètres du modèle et l'autocorrélation des erreurs (**supposée d'ordre un**) [4] :

---

1. L'auteur tient à exprimer ses remerciements envers M^{me} C. Babusiaux et M R Uctum pour leurs observations sur une première version de ce texte. Il va sans dire que sa responsabilité reste entière.

2 Non seulement les **relations** peuvent ne pas être linéaires, mais encore les **variables** elles-mêmes peuvent être mal spécifiées (niveau versus variation ; spécification des effets retardés...)

3 En effet, comme la très grande majorité des variables économiques sont caractérisées par une structure autorégressive, il en résulte que si les résidus d'un modèle sont aussi caractérisés par une telle structure, une présomption émane pour qu'une ou plusieurs variables aient été oubliées.

4 Méthodes de type Cochrane-Orcutt ou Hildreth Lu. On remarque que le système constitué par les relations [1] et [2] est équivalent à la relation suivante faisant intervenir les « pseudo variations » des variables :

$$Y(t) - \rho Y(t-1) = \sum_{i=1}^{n} a_i [X_i(t) - \rho X_i(t-1)] + (1-\rho) b + \varphi(t)$$

Après avoir fait varier le coefficient ρ entre −1 et 1, on retient la valeur qui minimise la variance de l'erreur estimée, c'est à dire du résidu $\hat{\varphi}$. Notons par ailleurs que la méthode des moindres carrés généralisés, bien que généralisant les deux méthodes qui viennent d'être rappelées, est rarement utilisée car elle nécessite la connaissance du processus d'autocorrélation des erreurs, ce qui est très peu fréquent.

$$Y(t) = \sum_{i=1}^{n} a_i X_i(t) + b + \rho u(t-1) + \varphi(t) \qquad [1]$$

$$u(t) = Y(t) - \sum_{i=1}^{n} a_i X_i(t) - b \qquad [2]$$

L'objet de cette note ne porte naturellement pas sur la méthode d'estimation de ce système à deux équations, laquelle est exposée dans un très grand nombre de manuels d'économétrie. Son but est de suggérer deux interprétations **économiques** pouvant justifier la présence du résidu retardé d'une période $u(t-1)$ dans une relation telle que [1]. L'enjeu est important car si la présence du résidu retardé peut être justifié au plan économique, son introduction dans l'équation d'ajustement a pour objectif **non seulement** une estimation plus efficace des paramètres, **mais encore**, de justifier au plan de la théorie économique, **la présence du terme ρu (t – 1) dans la valeur calculée de Y**[1].

La première interprétation proposée, la plus simple, repose sur l'hypothèse d'un **Modèle à Correction d'Erreur** (MCE) simple dans lequel on introduit une contrainte particulière ; dans cette optique, le défaut de spécification visé ci-dessus correspond précisément à l'omission de ce MCE dans la formulation du modèle (§ 1). La seconde interprétation repose sur l'hypothèse suivant laquelle au moins une des variables exogènes est une **grandeur mémorisée** (ou « anticipée », ou « permanente »), estimée d'après une moyenne pondérée des valeurs passées d'une variable observable, cette dernière étant affectée par des **erreurs de mesure** ; dans cette seconde optique, la structure autorégressive des résidus n'est pas attribuable à un défaut de spécification ou incomplétude du modèle, mais résulte au contraire de la spécification même du modèle (§ 2).

1. Un modèle à correction d'erreur contraint

Le modèle complet se décompose en deux équations. La première donne l'expression de la cible $\overline{Y}$ (objectif de la politique économique, grandeur désirée, valeur d'équilibre dans le long terme)[2] :

---

1. Une démarche habituelle consiste à **estimer** les paramètres du modèle (a_i et b) d'après le système constitué par les équations [1] et [2] mais à **conserver pour valeur calculée** (qui représente l'expression numérique de la théorie économique soumises à l'épreuve des faits), la valeur suivante :

$$\hat{Y}(t) = \sum_{i=1}^{n} \hat{a}_i X_i(t) + \hat{b} \quad \textbf{et non} \quad \hat{Y}(t) = \sum_{i=1}^{n} \hat{a}_i X_i(t) + \hat{b} + \underline{\hat{\rho}\hat{u}(t-1)}$$

Cependant, il arrive aussi que le terme $\widehat{\rho u}$ $(t-1)$ soit pris en compte dans la valeur calculée ; on doit souligner que, dans ce cas, la prise en compte de ce terme n'est jamais justifiée plan de la **théorie économique**, à ma connaissance (elle est par contre souvent justifiée par des fins de prévision).

2. La cible est supposée être parfaitement connue dans la mesure où elle représente le résultat d'un programme d'optimisation d'un agent qui maximise sous contrainte une certaine fonction (utilité pour le consommateur, profit pour l'entreprise, « bien être » pour l'Etat...).

$$\overline{Y}(t) = \sum_{i=1}^{n} a_i X_i(t) + b \qquad [3]$$

La seconde équation décrit le **processus d'ajustement** de la valeur effective Y sur la cible $\overline{Y}$. Si ce processus correspond à un Modèle à Correction d'Erreur (MCE) [1], on peut écrire, en considérant la forme la plus simple (et la plus usitée jusqu'ici) de ce modèle :

$$Y(t) - Y(t-1) = \beta\,[\overline{Y}(t-1) - Y(t-1)] + \gamma\,[\overline{Y}(t) - \overline{Y}(t-1)] + \varepsilon(t) \qquad [4]$$
$$0 \leq \beta \leq 1 \qquad\qquad \gamma \geq 0$$

où la variation Y dépend de l'écart retardé entre la cible et la réalisation (déséquilibre de court terme) et de la variation de la cible (modification de la valeur d'équilibre), ε(t) correspondant au résidu du modèle [2]. Les équations [2], [3] et [4] permettent de déduire :

$$Y(t) + (\gamma - 1)\,\overline{Y}(t-1) = \gamma\overline{Y}(t) + (1-\beta)\,u\,(t-1) + \varepsilon(t) \qquad [5]$$

avec $u\,(t) = Y(t) - \overline{Y}(t)$.

La relation [5] montre immédiatement que si $\gamma = 1$, on obtient l'expression suivante du MCE-contraint (MCEC par la suite [3]) :

$$Y(t) = \overline{Y}(t) + (1-\beta)\,u\,(t-1) + \varepsilon(t) \qquad [6]$$

En posant $\rho = (1 - \beta)$, et en remplaçant $\overline{Y}(t)$ par son expression [3], $Y(t)$ peut être **estimé** conformément à la relation suivante :

---

1 Davidson et ali (1978, p. 680) relevaient qu'il n'y a pas vraiment de théorie complètement articulée permettant de spécifier sans arbitraire la forme des ajustements dynamiques. Cependant, Salmon (1982) a montré que certaines formes du MCE peuvent être déduites de la résolution d'un problème d'optimisation dynamique reposant sur des fonctions de coût particulières, ce qui permet d'attribuer quelques fondements **économiques** au MCE

On ne s'interrogera pas ici sur le bien fondé économétrique de ce processus Rappelons néanmoins que le résultat essentiel de la « nouvelle économétrie » adaptée aux séries non stationnaires est de justifier le MCE au plan de la théorie statistique, lorsque certaines conditions particulières prévalent sur le groupe de variables non stationnaires qui est considéré. En effet, Granger à montré que si ·

– chacune des variables du groupe est intégrée d'ordre 1 (i.e. si, pour chaque variable, la **dérivée** est stationnaire),

– les variables considérées sont cointégrées (i.e s'il existe une combinaison linéaire des ces variables qui soit stationnaire, c'est à dire intégré d'ordre 0),

alors il existe une représentation du processus d'ajustement fondée sur un MCE. Etant donné que ces deux conditions sont assez souvent observées (au moins lorsque la longueur des séries considérée est suffisante pour que l'analyse en terme de cointégration soit économiquement pertinente), on justifie le MCE au plan de la théorie **statistique** tout en comprenant les succès économétriques de ce modèle.

Pour un exposé de synthèse sur le MCE en liaison avec la théorie statistique de la cointégration, le lecteur peut se reporter à l'article de Maurel (1989).

2. L'erreur $\varepsilon(t)$ s'interprète comme l'erreur associée au modèle complet constitué par les hypothèses jointes [3] et [4] ; sa signification est donc la même que celle habituellement retenue, et l'on admet que ε est un aléa de moyenne nulle (ce qui revient à supposer la validité du modèle).

3 L'Appendice de cet article montre que le modèle adaptatif est un cas particulier du MCEC

$$Y(t) = \sum_{i=1}^{n} a_i X_i(t) + \rho u(t-1) + b + \varepsilon(t) \qquad [7]$$

qui est **équivalente** à [1], de sorte que les techniques usuelles peuvent permettre l'estimation des paramètres de ce modèle, **le paramètre ρ ayant une signification précise** : plus ρ est grand (i.e. plus β est petit), et plus le déséquilibre de court terme se résorbe lentement [1]. Quant à l'hypothèse $\gamma = 1$, elle revêt aussi une signification **économique** très précise. A titre d'illustration, supposons que les grandeurs Y et X_i représentent des logarithmes ; l'hypothèse $\gamma = 1$ impose qu'en régime d'équilibre dynamique, caractérisé paf l'égalité $Y = \overline{Y}$ (et donc $u = 0$), les taux de croissance de Y et de $\overline{Y}$ sont égaux, condition qui ne prévaudrait pas si γ était différent de l'unité dans le cadre du MCE considéré. En d'autres termes, la contrainte $\gamma = 1$ est, dans le cadre du MCE [4], la seule qui soit rigoureuse au plan de la théorie économique, dans la mesure où il apparait a priori incohérent d'admettre que les valeurs observées et désirées d'une même variable croissent à un taux différent en régime d'équilibre dynamique.

Au plan de l'estimation du modèle, on doit souligner deux points :

- la valeur calculée $\hat{Y}(t) = \sum \hat{a}_i X_i(t) + \hat{\rho}\hat{u}(t-1) + \hat{b}$ issue de [7] correspond bien à la valeur théorique estimée dans le cadre des hypothèses du modèle, puisque le résidu final correspond bien à l'erreur du modèle $\varepsilon(t)$. **La valeur calculée doit donc bien contenir le résidu retardé u_{-1}** ;

---

1 Si $0 < \beta < 1$, alors le coefficient ρ doit être positif et compris entre 0 et 1, conformément à ce qui est presque toujours observé.

Par ailleurs, on peut introduire une hypothèse moins limitative suivant laquelle l'*équation* de la cible se modifierait entre t et $t-1$, en moyenne, sur l'ensemble de la période. On a alors :

[i] $$\overline{Y}'(t-1) = \sum_{i=1}^{n} a'_i X_i(t-1) + b'$$

[i i] $$\overline{Y}(t) = \sum_{i=1}^{n} a_i X_i(t) + b$$

[i i i] $$Y(t) - Y(t-1) = \beta[\overline{Y}'(t-1) - Y(t-1)] + [\overline{Y}(t) - \overline{Y}'(t-1)]$$

Il est alors envisageable, dans une première étape, d'estimer l'équation de la cible en $t-1$ par les MCO (équation [i]), le résidu de cette équation correspondant donc à $\hat{u}(t-1)$. D'après les équations [i i] et [i i i], on peut alors monter que le modèle complet peut être estimé par l'équation réduite suivante :

$$\overline{Y}(t) = \sum_{i=1}^{n} a_i X_i(t) + b + \rho\hat{u}(t-1) + \varepsilon(t)$$

Puisque $\hat{u}(t-1)$ est estimé avec les paramètres a_i' et b', lesquels peuvent être différents des paramètres a_i et b_i, la contrainte supportée ici est moins grande qu'avec la méthode de Cochrane-Orcutt qui impose les égalités $a_i' = a_i$ et $b' = b$. Les différences entre (a_i', b') et (a_i, b) permet d'apprécier dans quelle mesure la cible peut être considérée comme stable ou non, **en moyenne**, entre t et $t-1$. En fait, une différence significative pourrait s'interpréter comme une dérive structurelle constante de la fonction représentant la cible. Mais l'interprétation économique d'un tel phénomène reste cependant entière et propre au phénomène analysé.

– non seulement l'estimation ne semble pas poser de problème statistique particulier [1], mais encore elle conduit à des estimateurs plus efficaces. Un avantage matériel non négligeable est que les techniques nécessaires pour estimer le modèle (méthodes de type Cochrane-Orcutt) sont banalisées dans les logiciels économétriques.

Les remarques extrêmement simples qui viennent d'être présentées partent de considérations économiques pour aboutir à une vérification économétrique. Dans le déroulement des travaux de recherche, une démarche allant en sens inverse peut naturellement être adoptée. En effet, les développements ci-dessus montrent que si les résidus d'une équation économétrique ne faisant pas intervenir un processus d'ajustement sont caractérisés par une structure autorégressive simple, c'est qu'il existe **peut-être** un MCEC sous-jacent, non pris en compte dans la formulation. Par conséquent, un tel résultat suggère un retour sur la spécification du modèle, en associant à ce dernier un processus d'ajustement de type MCEC, dans la mesure où cet ajustement peut être justifié au plan des comportements économiques, naturellement.

Une illustration de cette approche peut être donnée par un modèle de demande de monnaie, lorsque l'encaisse désirée dite « de court terme » (égale à l'encaisse détenue) est supposée s'ajuster sur la demande de monnaie dite « de long terme » (cible résultant d'un programme d'optimisation) suivant un modèle à correction d'erreur. La contrainte $\gamma = 1$ implique qu'en régime d'équilibre dynamique, l'encaisse détenue est égale à la cible, et donc croît au même taux que cette dernière. Dans Prat (1988), on montre ainsi que le MCEC s'accorde bien avec un modèle de demande de monnaie.

Naturellement, la **condition** $\gamma = 1$ n'est nécessaire pour assurer la cohérence économique **que dans le cadre du MCE simple traduit par la relation [4]**. En fait, on peut distinguer deux possibilités (pouvant d'ailleurs jouer simultanément) susceptibles d'assouplir cette condition. En premier lieu, une version plus générale ajoutant un terme proportionnel à la variation **retardée** de l'endogène à droite de l'équation [4] conduit à une condition de cohérence dans un régime d'équilibre dynamique qui est plus souple que dans le MCE retenu ; mais cette condition n'est pratiquement jamais vérifiée par les résultats économétriques [2]. Une seconde possibilité consiste à **ajouter** au membre de droite de la relation [4], **une constante** α_0 s'identifiant à $(1-\gamma)g_0$; cette constante a pour effet de capturer l'écart entre le taux de croissance g_0 de la cible en régime d'équilibre dynamique et le taux de croissance de la valeur

1. Ce qui n'est pas toujours le cas pour le MCE traditionnel.

2 La version plus générale du MCE utilisée notamment par Hendry et Ericsson (1990) s'écrit, en ne retenant qu'un seul retard :

$$Y(t) - Y(t-1) = \alpha[Y(t-1) - Y(t-2)] + \lambda[\bar{Y}(t-1) - Y(t-1)] + \gamma[\bar{Y}(t) - \bar{Y}(t-1)]$$

Suivant cette formulation, la condition de cohérence dans un régime d'équilibre dynamique est plus souple que dans le MCE traditionnel, puisqu'elle s'écrit $\alpha + \gamma = 1$. Cependant, cette condition n'est pas vérifiée par les résultats économétriques présentés par les auteurs dans le cadre d'un modèle de demande de monnaie estimé sur les données séculaires du Royaume Uni (à vrai dire, elle n'est qu'exceptionnellement vérifiée).

observée qui serait donné par l'équation [4] (c'est à dire γg_0), afin d'assurer une solution cohérente dans un tel régime [1]. Cependant, d'une part la signification **économique** de cette constante n'est pas vraiment précisée, et d'autre part il faudrait montrer qu'elle compense effectivement l'écart indiqué, c'est à dire montrer que l'on a bien l'égalité $\alpha_0 = g_0 (1 - \gamma)$ [2]. En tout état de cause, le fait qu'il existe des formulations alternatives au MCEC restant cohérentes avec l'équilibre dynamique, n'enlève rien au fait que le MCEC fournit une interprétation économique **possible et très simple** de la structure autorégressive d'ordre un des résidus d'un ajustement linéaire.

2. Effets de mémoire et erreurs de mesure

On suppose ici que l'une au moins des variables X_t est le résultat d'un processus adaptatif appliqué sur une variable exogène Z_t [3], soit, s'il s'agit de X_1 :

$$X_1 (t) = (1 - \lambda) X_1 (t - 1) + \lambda Z_1 (t) \qquad 0 \leq \lambda \leq 1 \qquad [8]$$

où Z_1 représente la vraie mesure de la variable, c'est-à-dire celle qui est effectivement utilisée par les **agents économiques**. Soit Z_1^* la mesure utilisée par le modélisateur. On peut alors être certain que dans la très grande majorité des cas, il existe un écart **variable** entre Z_1 et Z_1^*, de sorte que l'on peut écrire :

$$Z_1 (t) = Z_1^* (t) + \eta_1 (t) \qquad [9]$$

où l'erreur η_1 est supposée être un bruit blanc.

Par exemple, si Z_1 représente le taux d'inflation, le modélisateur calculera ce taux par la relation $Z_1^*(t) = \log P (t) / P (t - 1)$, où P représente l'indice général des prix à la consommation, formule qui admet a priori non seulement l'hypothèse d'absence d'erreurs de perception sur les prix de chaque bien, mais encore l'hypothèse que la procédure d'agrégation par les indices est bonne. En d'autres termes, le modélisateur utilise une information ex-post, traitée et centralisée, tandis que les consommateurs utilisent non seulement ces informations médiatisées, mais encore l'information décentralisée et informelle « du moment », pouvant impliquer une confusion entre l'évolution générale des prix et les variations des prix relatifs [4]. Bien d'autres exemples pourraient être donnés, montrant l'existence d'écarts entre les mesures utilisées par le modélisateur et celles utilisées par les agents économiques. Même si en première approximation l'erreur η_1 semble pouvoir être négligée dans la majorité des

1. Sur ce point, voir Davidson et ali (1978), pp. 680 et suiv.

2. La constante α_0 doit bien sûr être distinguée de la constante pouvant résulter de l'équation de la cible. On admet ici que la technique d'estimation utilisée permet d'identifier des deux constantes.

3. On suppose soit que Z_t est stationnaire (ex : un taux de variation) soit qu'une correction est introduite afin d'évacuer toute possibilité de biais systématique si Z_t n'est pas stationnaire (voir Appendice).

4. Par exemple, la ménagère perçoit la hausse des prix autant par les annonces faites par les médias qu'en faisant son marché.

cas, les **effets cumulés** qu'elle exerce sur X_1 par l'intermédiaire de la relation [8] peuvent d'autant moins l'être qu'ils peuvent générer une structure autorégressive des erreurs, sans que l'on puisse nécessairement invoquer un défaut de spécification du modèle. En effet, les relations [8] et [9] sont équivalentes au système suivant :

$$X_1(t) = X_1^*(t) + E(t) \tag{10}$$

$$X_1^*(t) = (1-\lambda)\, X_1^*(t-1) + \lambda Z_1^*(t) \tag{11}$$

$$E(t) = (1-\lambda)\, E(t-1) + \lambda \eta_1(t) \tag{12}$$

La relation [12] montre que même si η_1 est un aléa de moyenne nulle, l'effet cumulé E peut être caractérisé par des fluctuations pouvant ressembler à celles que connaissent un grand nombre de variables observables (le processus générateur de E correspondrait alors à une « chaine de Markov »).

Revenons au modèle théorique de base qui s'écrit, en isolant la variable X_1 :

$$Y(t) = a_1 X_1(t) + \sum_{i=2}^{n} a_i X_i + b + \varepsilon(t) \tag{13}$$

Naturellement, on ne peut estimer le modèle au moyen de la relation [13], puisque X_1 est inconnue (on suppose ici que tous les X_i sont connus pour $i > 1$). On ne peut estimer le modèle qu'en remplaçant X_1 par sa mesure connue par l'économètre X^* (relation [11]), soit :

$$Y(t) = \hat{a}_1 X_1^*(t) + \sum_{i=2}^{n} \hat{a}_i X_i + \hat{b} + \hat{u}(t) \tag{14}$$

Cette démarche implique que l'erreur E est contenue dans le résidu empirique $\hat{u}$ (relations [10], [13] et [14]) :

$$\hat{u}(t) = \hat{a}_1 E(t) + \hat{\varepsilon}(t) \tag{15}$$

où le résidu $\hat{\varepsilon}(t)$ correspond bien à l'erreur du modèle. Les relations [12] et [15] montrent que le résidu $\hat{u}$ a une structure autorégressive de nature stochastique, puisque $\hat{u}$ contient E, et que E est caractérisé par une telle structure.

Plus précisément, on peut remplacer dans [15] $E(t)$ par son expression [12] ; on obtient [1] :

$$\hat{u}(t) = \hat{a}_1(1-\hat{\lambda})\, E(t-1) + \hat{a}_1 \hat{\lambda} \hat{\eta}_1(t) + \hat{\varepsilon}(t) \tag{16}$$

1. Puisqu'en principe on ne peut estimer que les erreurs u et v, seules ces dernières devraient être surmontées d'un chapeau (^). Bien que les erreurs ε et η ne puissent être estimées, on a considéré les grandeurs $\hat{\varepsilon}$ et $\hat{\eta}$ pour plus de cohérence dans l'écriture ; il est clair que dans ce cas, il faut supposer que $\hat{\varepsilon}$ et $\hat{\eta}$ sont implicitement contenu dans $\hat{v}$

En écrivant [15] en (t – 1), on a :

$$\hat{a}_1 E(t-1) = \hat{u}(t-1) - \hat{\varepsilon}(t-1) \qquad [17]$$

En reportant [17] dans [16], on obtient :

$$\left.\begin{aligned} &\hat{u}(t) = (1-\hat{\lambda})\hat{u}(t-1) + \hat{v}(t) && \text{(a)} \\ \text{avec :}\quad &\hat{v}(t) = \hat{\varepsilon}(t) - (1-\hat{\lambda})\,\hat{\varepsilon}(t-1) + \hat{a}_1\hat{\lambda}\hat{\eta}_1(t)^{1} && \text{(b)} \end{aligned}\right\} \qquad [18]$$

Compte tenu de [18], le modèle [13] peut être estimé par la relation suivante :

$$Y(t) - [(1-\hat{\lambda})\hat{u}(t-1)] = \hat{a}_1 X_1^*(t) + \sum_{i=2}^{n} \hat{a}_i X_i(t) + \hat{b} + \hat{v}(t) \qquad [19]$$

L'équation [19] doit être estimée simultanément avec la relation [11] qui permet le calcul de $X_1^*(t)$, cette variable dépendant aussi de λ. La valeur estimée de λ peut être déterminée par balayage de manière à minimiser la variance de $\hat{v}$. On voit donc que la relation [13] reposant sur les vraies valeurs de X qui sont **inobservables**, conduit à l'ajustement suivant fondé sur les valeurs **observables** :

$$Y(t) = a_1 X_1^*(t) + \sum_{i=2}^{n} \hat{a}_i X_i(t) + \hat{b} + \hat{\rho}\hat{u}(t-1) + \hat{v}(t) \qquad [20]$$

en **imposant la condition** $\rho = (1-\lambda)$, contrainte qui n'existe pas avec l'hypothèse d'un MCEC (relation [7]). Un test sur la validité de l'hypothèse d'une erreur de mesure pourrait donc consister à déterminer librement le couple $(\hat{\lambda}, \hat{\rho})$ de manière à minimiser la variance de $\hat{v}$ sans imposer l'égalité $\rho = 1-\lambda$, et à tester ensuite si les valeurs obtenues pour ρ et $(1-\lambda)$ sont significativement différentes l'une de l'autre. Si l'égalité est admissible, on pourrait conclure que « tout se passe comme si » les résidus retardés reflètaient bien une erreur de mesure sur la variable mémorisée. La présence du résidu retardé dans la valeur calculée serait ainsi regardée comme traduisant le biais résultant des erreurs commises par le **modélisateur** dans la mesure de la vraie valeur mémorisée par les agents économiques. Il faut cependant souligner qu'il s'agit ici d'une interprétation **possible** et non d'une **preuve** sur la réalité du phénomène [2].

1. Dans le cas où il y aurait j variables mémorisées (au même taux η), on obtiendrait :

$$v(t) = \varepsilon(t) - (1-\lambda)\,\varepsilon(t-1) + \lambda \sum_{i=1}^{j} a_i\,\eta_i(t)$$

où chaque X_i $(i \leq j)$ est une moyenne pondérée des valeurs passées des variables Z_i, chaque variable Z_i étant affectée par une erreur de mesure η_i

2. Par exemple, dans le cas particulier où la condition $\beta = \lambda$ est satisfaite, on ne sait plus si on estime un MCEC ou bien si c'est la variable mémorisée qui contient des erreurs.

Il importe de souligner que, d'après la relation [18b], le **résidu empirique $\hat{v}$ ne correspond pas à l'estimation de l'erreur ε du modèle théorique**, puisque $\hat{v}$ dépend non seulement de $\varepsilon(t)$ mais encore de $\varepsilon(t-1)$ et de l'erreur de mesure $\eta_1(t)$ commise sur Z_1 [1]. Cependant, cette dernière relation montre que si ε et η_1 sont des aléas de moyenne nulle restant petits (i.e. si le modèle satisfait aux critères habituels de validité et si Z_1^* est une mesure sans biais de Z_1), on peut considérer $\hat{Y}^*$ comme une **approximation** acceptable de $\hat{Y}$. En effet, sous les conditions indiquées, Y, $\hat{Y}^*$ et $\hat{Y}$ auront **à chaque instant** même espérance mathématique [2], ce qui signifie encore que $\hat{Y}^*$ est une estimation sans biais de $\hat{Y}$. Cet argument permet de justifier au plan économique la confrontation entre les valeurs de $Y(t)$ et celles de $\hat{Y}^*(t)$ afin d'apprécier la capacité du modèle à rendre compte de la réalité [3]. Par contre, si la valeur calculée ne prenait pas en compte le terme $\hat{\rho}\hat{u}(t-1)$ (i.e. si le modèle [13] était estimé d'après l'équation [14]), l'espérance de Y diffèrerait à chaque instant t de celle de la valeur calculée du terme autorégressif $a_1(1-\lambda)E(t-1)$ (relations [12] et [15]) [4].

Une illustration de cette approche est donnée dans Prat (1992). Dans le cadre d'une formulation du cours des actions (USA, 1956 à 1989), il est supposé que le taux d'actualisation utilisé par les investisseurs sur le marché des actions résulte d'une mémoire des valeurs passées « du » taux d'intérêt de référence (utilisé par le modé-

1. En effet, la valeur calculée d'après [20] s'écrit :

$$\hat{Y}^*(t) = \hat{a}_1 X_1^*(t) + \sum_{i=1}^{n} \hat{a}_i X_i(t) + \hat{b} + \hat{\rho}\hat{u}(t-1) \qquad [21]$$

de sorte que la valeur effectivement estimée $\hat{Y}^*$ diffère de la valeur $\hat{Y}$ qui aurait été estimée si les vraies valeurs de X_1 avaient été connes :

$$\hat{Y}(t) - \hat{Y}^* = \hat{\lambda}\hat{a}_1\hat{\eta}(t) - (1-\hat{\lambda})\hat{\varepsilon}(t-1) \qquad [22]$$

2. En fait, la valeur effectivement calculée $\hat{Y}^*$ pourra d'autant plus être considérée comme une bonne approximation de la vraie valeur théorique estimée $\hat{Y}$ que le résidu empirique $\hat{v}$ est petit.

3. Ajoutons un résultat très intuitif : la variance V des résidus v est plus grande que la variance des résidus ε, puisque l'on déduit de [18]b

$$V(v) = [1 + (1-\lambda)^2]\,V(\varepsilon) + |\,a_1\,|^2 \lambda^2 V(\eta)$$

Comme $0 \leq \lambda \leq 1$, on a bien $V(v) > V(\varepsilon)$. On peut donc en déduire que si on avait pu estimer $\hat{Y}$ (i.e. si X_1 avait pu être estimé sans erreur), l'ajustement aurait été en principe de meilleure qualité que l'ajustement obtenu entre Y et $\hat{Y}^*$. Par conséquent, si ce dernier ajustement est satisfaisant, on peut en déduire que le modèle est vérifié.

4. Naturellement, même avec l'ajustement [14], la moyenne de Y **sur l'ensemble de la période d'estimation** reste égale à la moyenne de la valeur estimée (c'est là une conséquence banale des MCO). Par ailleurs, relevons que le résidu v est (négativement) autocorrélé, puisque l'on a :

$$\hat{v}(t) = \hat{\varepsilon}(t) - (1-\hat{\lambda})\,\hat{\varepsilon}(t-1) + \hat{a}_1\hat{\lambda}\hat{\eta}_1(t)$$

$$\hat{v}(t-1) = \hat{\varepsilon}(t-1) - (1-\hat{\lambda})\,\hat{\varepsilon}(t-2) + \hat{a}_1\hat{\lambda}\hat{\eta}_1(t-1)$$

Ce phénomène sera d'autant plus important que λ est petit, que la variance de $\hat{\eta}_1$ est grande par rapport à celle de $\hat{\varepsilon}$ et que le coefficient a_1 est grand. Cependant, des simulations ont montré qu'en règle générale, on peut raisonnablement supposer que l'autocorrélation reste faible. En effet, en réalisant des simulations stochastiques fondées sur les hypothèses $\lambda = 0{,}02$, $a_1 = 1$, et en retenant une même variance pour $\hat{\varepsilon}$ et $\hat{\eta}_1$, on a constaté que l'autocorrélation de $\hat{v}$ était pratiquement nulle.

lisateur, lequel sait que cette référence change au cours du temps, d'où une erreur de mesure). De même, le taux de croissance attendu de long terme du « revenu des actionnaires » est supposé dépendre de la mémoire que les investisseurs conservent des taux de croissance passés du revenu ; compte-tenu de l'incertitude pesant sur la définition du « revenu » (dividendes, bénéfices... ?), le taux de croissance observé ne peut également être représenté qu'avec une erreur de mesure [1]. Les processus de mémorisation étant supposés de type adaptatif (taux λ), la prise en compte de la valeur retardée du résidu s'est ainsi révélée pertinente au plan empirique, la valeur du coefficient associé au résidu retardé étant imposée à $(1 - \lambda)$.

Conclusions

Suivant l'hypothèse du modèle à correction d'erreur contraint (MCEC), la structure autorégressive **d'ordre un** des résidus d'un modèle linéaire – qui est observée lorsque le MCEC est omis – peut s'interpréter comme un défaut de spécification du modèle dans son ensemble, et l'adjonction du MCEC apporte une solution possible pour corriger ce défaut de spécification, lorsque naturellement l'hypothèse d'un ajustement partiel peut être justifiée au plan de l'analyse économique [2]. Par contre, suivant l'hypothèse d'une variable mémorisée avec erreur de mesure, le modèle économique n'est pas mal spécifié, mais les résidus observés sont autocorrélés en raison du processus stochastique caractérisant l'erreur de mesure commise par le modélisateur sur la variable mémorisée, de sorte qu'il devient souhaitable de prendre en compte cette autocorrélation dans l'ajustement pour traduire le comportement des agents.

Dans le cadre du MCEC proposé, le coefficient d'autocorrélation $\hat{\rho}$ du résidu $\hat{u}$ doit être optimisé au sein du modèle, et la valeur calculée $\hat{Y}^*$ (qui tient compte du terme $\hat{\rho}\,\hat{u}$ (t – 1)) s'interprète sans ambiguïté comme la valeur théorique calculée dans le cadre du modèle. Par contre, avec l'hypothèse de variables mémorisées au taux λ (la mémorisation s'effectuant suivant un processus adaptatif, la variable observée étant affectée par une erreur de mesure), d'une part le coefficient d'autocorrélation ρ doit être contraint ($\rho = 1 - \lambda$) et d'autre part la valeur calculée (qui tient compte du terme $(1 - \hat{\lambda})\,\hat{u}$ (t – 1)) ne peut être regardée que comme une estimation sans biais de la vraie valeur théorique calculée dans le cadre du modèle.

Un autre point important me semble devoir être souligné. Qu'il s'agisse de l'une ou de l'autre des deux approches qui viennent d'être présentées, un test ex-post doit être réalisé dans le but de vérifier qu'il n'existe pas de décalage temporel systématique entre la valeur observée Y et la valeur calculée $\hat{Y}^*$. Cette condition, nécessaire pour

1. En d'autres termes, le taux d'intérêt de référence et le taux de croissance du revenu jouent le rôle de la variable Z_1 des relations [8] et [9].

2. Soulignons que cette solution n'est que **possible**. Rien ne permet d'affirmer qu'elle sera confirmée par les applications empiriques.

conclure qu'un modèle est bien spécifié, devrait d'ailleurs être analysée systématiquement, quel que soit le modèle. Cependant, elle prend ici une importance particulière, dans la mesure où la présence du résidu retardé est en principe susceptible de provoquer un retard systématique de la valeur calculée sur la valeur observée [1], ce qui serait pour le moins gênant dans la mesure où la valeur calculée est censée représenter une explication possible de la valeur observée. Pour étayer la bonne spécification du modèle, il paraît donc nécessaire de vérifier que la corrélation entre les valeurs observées et calculées de Y est maximale lorsque les deux grandeurs sont considérées en phase. En fait, l'expérience montre que cette condition de validité est très sévère [2].

Enfin, on peut se demander dans quelle mesure les deux interprétations proposées peuvent être considérées simultanément, puisque, d'un point de vue économique les deux hypothèses peuvent parfaitement prévaloir ensemble. Dans ce cas, le « bon » modèle de base correspond à la cible à laquelle le MCEC est associé, mais la cible est déterminée par au moins une variable mémorisée au taux λ avec erreurs de mesure. Dans le cas particulier où $\beta = \lambda$ (β caractérisant l'ajustement décrit par le MCEC), le modèle peut être estimé en optimisant directement le coefficient d'autocorrélation $\rho = (1 - \beta) = (1 - \lambda)$, mais la valeur calculée n'est assimilable à la valeur théorique qu'en espérance, si l'erreur de mesure η est un aléa. Par contre, dans le cas général où β et λ prennent des valeurs différentes, il devient nécessaire de traiter le résidu de second rang v.

Ces quelques réflexions avaient pour seule ambition d'indiquer deux directions possibles de recherche pour interpréter la structure autorégressive d'ordre un des résidus d'un modèle linéaire, ces interprétations permettant, le cas échéant, de justifier économiquement la présence du résidu retardé dans la valeur estimée. Les techniques économétriques actuelles peuvent permettre de confronter ces deux hypothèses aux données de l'observation. L'hypothèse d'un MCEC peut aisément donner lieu à un test empirique simple afin d'être infirmée ou confirmée [3] ; par contre, si l'hypothèse d'une variable mémorisée avec erreurs peut aussi être soumise à des tests empiriques, les conclusions de ces tests nous semblent actuellement plus incertaines que pour la première hypothèse.

1. En effet le résidu retardé $\hat{u}(t-1)$ contient $Y(t-1)$

2 En effet, on constate le plus souvent, dans les modèles faisant intervenir l'erreur retardée, que la valeur calculée retarde sur la valeur observée. Dans les deux exemples mentionnés dans cet articles (Prat, 1986 et 1989 pour le MCEC appliqué à la demande de monnaie ; Prat, 1992 pour l'optique fondée sur des effets de mémoires avec erreur sur les variables, appliquée au cours des actions), cette condition est satisfaite.

3. Sans doute, de la même manière que, dans le bourgeois gentilhomme de Molière, Monsieur Jourdain faisait de la prose sans le savoir, un bon nombre d'économètres ont-ils estimés des MCEC sans en être complètement conscients !

APPENDICE

Le MCEC et le modèle adaptatif : comparaison

Le processus d'ajustement [6] peut être utilement comparé au processus adaptatif traditionnel suivant très souvent usité :

$$Y(t) - Y(t-1) = \lambda\,[\overline{Y}(t) - Y(t-1)] + \varepsilon(t) \qquad [23]$$

alors que le MCEC [6] peut s'écrire :

$$[Y(t) - Y(t-1)] - [\overline{Y}(t) - \overline{Y}(t-1)] = \beta\,[\overline{Y}(t-1) - Y(t-1)] + \varepsilon(t) \qquad [24]$$

Le modèle adaptatif [23] et le MCEC [24] présentent clairement deux différences qui sont liées :

- d'après le modèle adaptatif, c'est la variation de Y qui dépend de l'écart entre la cible et la valeur observée, alors que d'après le MCEC c'est l'écart entre la variation de Y et la variation de la cible qui dépend de l'écart entre cette dernière et la valeur observée. Sur ce point, le MCEC est moins restrictif car la cible $\overline{Y}$ n'y est pas supposée stable au cours du processus d'ajustement, contrairement au processus adaptatif [1]. Ajoutons qu'en régime d'équilibre dynamique, le modèle adaptatif conduit à un biais systématique dans la mesure où la réalisation Y est systématiquement en dessous de la cible $\overline{Y}$ si cette dernière est caractérisée par un trend ascendant, ce qui est économiquement inconsistant ;
- dans le MCEC présenté ici, c'est l'écart retardé entre $\overline{Y}$ et Y qui intervient, alors que dans le modèle adaptatif fait intervenir l'écart entre la cible $\overline{Y}$ en t et Y en $t-1$. Cette différence peut être attribuée au fait que dans le MCEC, la cible varie elle-même au cours du processus d'ajustement, ce qui n'est pas le cas pour le modèle adaptatif.

Rappelons que le processus adaptatif [23] conduit à la relation suivante très souvent utilisée par les économètres :

$$Y(t) = (1-\lambda)\,Y(t-1) + \sum_{i=1}^{n} \lambda\, a_i X_i(t) + \lambda\, b + \varepsilon(t) \qquad [25]$$

Ainsi, alors que le **modèle adaptatif** conduit à adjoindre aux variables exogènes l'**endogène retardée** (relation [25] [2]), le **MCEC** conduit à adjoindre à ces dernières le **résidu retardé** (relation [7]).

---

1. Remarquons que le modèle adaptatif correspond au MCE classique (équation [4]) dans le cas particulier où $\beta = \gamma$.

2. En fait, la présence de l'endogène retardée souffre d'une ambiguïté au niveau de l'interprétation théorique. En effet, la relation [25] peut en fait traduire deux hypothèses radicalement différentes :

– l'hypothèse d'un ajustement adaptatif entre la grandeur observée et la grandeur désirée, conformément aux équations [23] et [24] ;

– l'hypothèse suivant laquelle ce seraient non pas les X_i qui interviendraient dans le comportement des agents, mais leurs **valeurs mémorisées** suivant un processus adaptatif au taux λ ·/..

RÉFÉRENCES CITÉES

DAVIDSON J.E.H., HENDRY D., SRBA F., YEO S. (1978) "Econometric Modelling of the Aggregate Time-series Relationship between Consumers' Expenditure and Income in the United Kingdom", *Economic Journal*, vol. 88, 661-692.

HENDRY D. et ERICSSON N.R. (1991) "An Econometric Analysis of U.K. Money Demand, in Monetary Trend in the United States and the United Kingdom by Friedman M. and Schwartz A.J.", *American Economic Review*, mars.

MAUREL F. (1989) "Modèles à correction d'erreur : l'apport de la théorie de la co-intégration", *Économie et Prévision* n° 88-89, 105-125.

PRAT G. (1986) *Demande de monnaie et incertitude : ambiguité théorique et réalité empirique*, CEMA, document miméo., septembre, 34 p.

PRAT G. (1988) "Note à propos de l'influence de l'incertitude sur la demande de monnaie", *Revue Économique*, mars, 451-460.

PRAT G. (1992) "Anticipation et évaluation des actions", in *Monnaie, taux d'intérêt et anticipations*, MAROIS W. et KEMPF H. éd., Economica.

SALMON T. (1982) "Error Correction Mechanisms", *Economic Journal*, septembre, 615-629.

Ces deux phénomènes peuvent intervenir simultanément, et on ne peut les discriminer économétriquement ; seules les valeurs estimées du taux λ peuvent permettre une présomption d'interprétation sur le phénomène dominant, dans la mesure où, le plus souvent, on peut penser que les ajustements sont plus rapides que les effets de mémoire. Les modèles de demande de monnaie illustrent parfaitement cette difficulté.