Le signe de l'erreur systématique dans l'estimation de l'élasticité de la demande par la méthode des moindres carrés

PAR

HENRI BLANPAIN

I

On sait qu'il n'est pas possible d'estimer directement les paramètres de la courbe de demande

$$q = ap + b \tag{1}$$

par simple considération de paires d'observations p et q. Cela tient au fait que les équations qui permettent de déterminer les coefficients de régression par la méthode des moindres carrés ne sont applicables que dans le cas où nous prenons la régression d'une seule variable dépendante sur une ou plusieurs variables indépendantes.

Cette condition n'est évidemment pas remplie dans l'équation (1) où les variables p et q dépendent toutes deux l'une de l'autre, au moins généralement. (Il pourrait en être autrement si les prix étaient, par exemple, fixés par voie d'autorité.)

Cette impossibilité d'ordre théorique a d'ailleurs une conséquence pratique : étant donné que, dans sa forme la plus simple, la courbe d'offre a exactement la même structure que (1), nous ne pourrons jamais savoir si le résultat obtenu est une courbe de demande ou une courbe d'offre, même approximative. La considération du signe du coefficient de p dans (1) est en effet un critère fort médiocre et qui ne peut évidemment pas servir dans certains cas qui, quoique rares, n'en sont pas moins intéressants (inclinaison positive de la courbe de demande par exemple).

II

Étant donné que toute équation « raisonnable » de la demande comporte au moins deux variables dépendantes, prix et quantité, il nous faut, en vertu du principe rappelé ci-dessus, considérer un système comprenant au moins deux équations.

https://doi.org/10.1017/S0770451800051174 Published online by Cambridge University Press

Soit, en effet, par exemple,

$$q = ap + bi + c + u \tag{2}$$

$$q = a'p + b't + c' + u' \tag{3}$$

où (2) est la courbe de demande et (3) la courbe d'offre avec « i » le revenu et « t » un quelconque facteur « temps », u et u' étant des variables purement aléatoires qui représentent les « écarts » dus au caractère statistique des relations envisagées.

Nous supposerons que les variables i et t sont indépendantes. Ceci ne présente aucune difficulté en ce qui regarde t et semble pouvoir être admis, en première approximation, pour i, à condition toutefois que l'étude ne porte que sur un produit ou un groupe de produits qui ne représente pas une fraction trop considérable du revenu de l'économie envisagée.

On pourra alors transformer (2)-(3) en

$$q = \alpha t + \beta i + \gamma + u, \tag{4}$$

$$p = \alpha' t + \beta' i + \gamma' + u', \tag{5}$$

Étant donné que chacune des équations (4) et (5) ne comporte qu'une seule variable dépendante, nous pouvons estimer les coefficients de ces équations par la méthode des moindres carrés. Ceci fait, il sera possible d'estimer les coefficients de (2)-(3) en fonction des valeurs trouvées pour (4)-(5). Le résultat sera différent de celui qui aurait été trouvé par application directe de la méthode des moindres carrés aux équations (2)-(3).

La dernière opération envisagée ci-dessus, à savoir la détermination des paramètres de (2)-(3) en fonction de ceux de (4)-(5), n'est cependant possible que parce que les variables indépendantes i et t apparaissent chacune dans une des deux équations (2)-(3) et dans une seule, ou, en d'autres termes, parce que le système (2)-(3) est exactement « identifié » [1]. La méthode indirecte d'estimation des paramètres qui a été exposée ci-dessus n'est que rarement possible et les calculs sont généralement difficiles.

Cette méthode d'estimation, qui fait appel à des hypothèses « a priori » à propos de la structure des équations envisagées est appelée « estimation structurelle ».

---

[1] Voir à ce sujet : W. C. Hood and T. Koopmans, *Studies in Econometric Method*, New-York, 1953 et L. Klein, *A Textbook of Econometrics*, New-York, 1953.

https://doi.org/10.1017/S0770451800051174 Published online by Cambridge University Press

III

Il résulte de la discussion ci-dessus que la méthode généralement employée pour estimer l'élasticité de la demande, à savoir l'estimation directe des paramètres d'une équation du genre de (1) est théoriquement inadéquate. Il n'est toutefois pas exclu que l'application directe de la méthode des moindres carrés donne des résultats qui ne soient pas fort différents de ceux qui sont obtenus par une méthode plus correcte. La comparaison des résultats obtenus par les deux méthodes peut donc être intéressante.

Un article de G. M. Kuznets (2) permet précisément de faire cette comparaison dans quatre cas. Il s'agit d'études portant sur la demande de produits alimentaires aux États-Unis. La détermination de l'élasticité de la demande par rapport au prix ayant été faite en utilisant les mêmes variables par la méthode directe et par la méthode « structurelle », les résultats sont les suivants :

	Estimation structurelle	Moindres carrés
I Girshick-Haavelmo	—0,25	—0,34 (—0,53)
II Hildreth-Jarrett	—0,89	—0,95
III idem	—0,76	—0,76
IV French	—0,24	—0,45 (—0,71)

Note : les parenthèses indiquent les résultats obtenus en prenant les prix comme variables dépendantes, les quantités comme variables indépendantes.

Il est assez remarquable que les résultats obtenus par la méthode des moindres carrés soient, en valeur absolue, plus grands dans trois cas sur quatre et équivalents dans l'étude III.

---

(2) G. M. Kuznets, Measurement of Market Demand with particular reference to consumer's demand for food, *Journal of Farm Economics*, 1953, p. 878.

Les produits considérés et les périodes étudiées dans les différents cas sont les suivants :

I : Tous produits alimentaires, 1922-41.

II et III : Produits de l'élevage, 1922-41.

IV : Viande, 1919-41.

Les modèles I et IV envisagent la demande au niveau du consommateur, les modèles II et III, la demande au niveau du producteur *(at farm level)*.

Les modèles II et III sont identiques à ceci près que II est linéaire et III linéaire après transformation logarithmique.

Quatre observations, ou même trois seulement, sont un échantillon assez faible pour en déduire une loi générale, mais les écarts observés sont relativement considérables.

Le but de cette note est de montrer que les différences observées ne s'expliquent, sauf « accident » dû à l'échantillonnage, que dans une hypothèse bien précise qui pourrait être susceptible d'interprétation en termes économiques ou pourrait, plus modestement, orienter des recherches empiriques ultérieures.

IV

Considérons le modèle suivant :

$$y_1 + \beta_{12} y_2 + \gamma_1 z_1 = v_1 \tag{6}$$

$$\beta_{21} y_1 + y_2 + \gamma_2 z_2 = v_2 \tag{7}$$

où :

y_1 et y_2 sont des variables dépendantes
z_1 et z_2 sont des variables indépendantes
v_1 et v_2 représentent les écarts statistiques observés.

Ce modèle correspond à (2)-(3) en posant :

$$y_1 = q \qquad y_2 = p \qquad z_1 = i \qquad z_2 = t \qquad v_1 = u$$

$$\beta_{12} = -a \qquad \gamma_1 = -b \qquad \beta_{21} = -\frac{1}{\alpha'}, \qquad \gamma_2 = \frac{b'}{a'}, \qquad v_2 = -\frac{u'}{a'},$$

Estimer l'élasticité de la demande revient à estimer « a » dans (2) ou β_{12} dans (6).

Asymptotiquement, c'est-à-dire pour un échantillon infiniment grand, la méthode « structurelle » nous permet d'obtenir la vraie valeur de β_{12}. Par contre, la méthode des moindres carrés donne une valeur, soit $\beta_{12}^{(1)}$, entachée d'une erreur systématique, même asymptotiquement.

Il est possible [3] de déterminer $\beta_{12}^{(1)}$ en fonction des vraies valeurs des paramètres du modèle (6)-(7).

On a :

$$\beta_{12}^{(1)} = \frac{\beta_{21}\,\sigma_1^2 - (1 + \beta_{12}\,\beta_{21})\,\sigma_{12} + \beta_{12}\,(\sigma_2^2 + \gamma_2^2\,\mu^2\,z_2^0)}{\beta_{21}^2\,\sigma_1^2 - 2\,\beta_{21}\,\sigma_{12} + \sigma_2^2 + \gamma_2^2\,\mu^2\,z_2^0}$$

---

[3] Voir W. C. Hood and T. Koopmans, *Op. cit.*, pp. 221-235.

où :

σ_1^2, σ_2^2 et σ_{12} sont respectivement la variance de v_1 et de v_2 et la covariance de v_1 et v_2)

et où $\mu^2 z_2^0$ est la variance des résidus de la détermination de z_2 par la méthode des moindres carrés étant donné z_1. (En d'autres termes, en unités normalisées, si r est le coefficient de corrélation entre z_1 et z_2, $\mu^2 z_2^0 = 1 - r^2$.)

Dans ces conditions, l'erreur systématique résultant de l'utilisation de la méthode des moindres carrés est, asymptotiquement,

$$\delta = \beta_{12}^{(1)} - \beta_{12} = \frac{(1 - \beta_{12}\beta_{12})(\beta_{21}\sigma_1^2 - \sigma_{12})}{\beta_{21}^2\sigma_1^2 - 2\beta_{21}\sigma_{12} + \sigma_2^2 + \gamma_2^2\mu^2 z_2^0}$$

Appelant $a^{(1)}$ (2) l'estimation, par la méthode des moindres carrés, de a dans (2), l'observation que l'élasticité par rapport aux prix est généralement plus grande, en valeur absolue, lorsque cette estimation est faite par la méthode des moindres carrés, se traduit par

$$|a^{(1)}| - |a| > 0$$

ou

$$a^{(1)} - a < 0$$

ou

$$\delta = \beta_{12}^{(1)} - \beta_{12} > 0$$

Pour simplifier la discussion posons :

$$A = 1 - \beta_{12}\beta_{21} \qquad B = \beta_{21}\sigma_1^2 - \sigma_{12}$$

$$1/C = \beta_{21}^2\sigma_1^2 - 2\beta_{21}\sigma_{12} + \sigma_2^2 + \gamma_2^2\mu^2 z_2^0$$

Nous avons alors $\delta = ABC$ et $\delta > 0$ entraîne nécessairement que zéro ou deux des quantités A, B et C sont négatives.

Nous pouvons généralement nous limiter au cas où la courbe de demande est descendante et la courbe d'offre montante, c'est-à-dire dans (2)-(3), $a < 0$ et $a' > 0$ et dans (6)-(7) $\beta_{12} > 0$ et $\beta_{21} < 0$. Dans ces conditions, $\beta_{12}\beta_{21} < 0$ et $A > 0$, nécessairement.

D'autre part, $B = \beta_{21}\sigma_1^2 - \sigma_{12} < 0$ correspond, dans le plan $(\beta_{21}\,\sigma_{12})$ à la région située en dessous et à droite de la droite d'inclinaison positive $\beta_{21}\sigma_1^2 - \sigma_{12} = 0$. Cette droite est d'inclinaison positive étant donné que $\sigma^2 > 0$.

Par conséquent, nous limitant toujours au cas général $\beta_{21} < 0$ (courbe d'offre montante), il est évident que tous les points du deuxième quadrant et une partie des points du troisième quadrant correspondent à $B < 0$.

Enfin, toujours dans le plan $(\beta_{21}\, \sigma_{12})$, $1/C = 0$ est représenté par une hyperbole dont les asymptotes sont $\beta_{21} = 0$ et $\sigma_{12} = \frac{1}{2}\, \sigma_1^2\, \beta_{21}$ et dont les deux branches sont situées dans les premier et troisième quadrants. Les points du plan situés à l'intérieur de chacune des deux branches correspondent à $C < 0$. Si nous limitons l'étude au cas $\beta_{21} < 0$ (courbe d'offre montante), seuls certains points du troisième correspondent à $C < 0$.

Par conséquent, il est impossible de trouver dans le deuxième quadrant des points tels que $BC > 0$ et δ ne peut être positif que si $\sigma_{12} < 0$.

V

Cette conclusion semble assez remarquable par son contenu et par sa généralité.

Elle signifie, en effet, qu'il existe une corrélation positive entre les variations non expliquées de « q » dans (2)-(3). (Cette corrélation est positive étant donné que $v_1 = u$ et $v_2 = -u'/a'$ avec $a' = -{}^1/\beta_{21} > 0$). Le fait de connaître le signe de cette corrélation peut être utile dans le choix d'hypothèses ultérieures quant à la forme de (2)-(3).

La conclusion de la discussion ci-dessus est d'autre part fort générale en ce qu'elle ne dépend que des hypothèses les plus classiques sur la forme des courbes d'offre et de demande et du fait que le système (2)-(3) est identifié. En particulier, elle ne dépend pas du choix qui est fait quant aux variables indépendantes qui figurent dans le modèle considéré.

Il faut signaler en terminant que l'évaluation de l'erreur systématique découlant de l'application directe de la méthode des moindres carrés est un problème assez difficile dont la solution exacte ne semble avoir été mise au point que dans le cas d'un système de deux équations exactement identifié. Certaines recherches préliminaires, assez incomplètes d'ailleurs, donnent à penser que la solution ci-dessus est approximativement valable dans des hypothèses plus générales.