



The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

Efficacité
dans la production céréalière :
une approche
non paramétrique

Isabelle PERRIGNE
Michel SIMIONI

**Efficient configuration
of grain production :
a nonparametric
approach**

Key-words:

cost scale economies, grain
production, nonparametric
regression

**Efficacité dans
la production
céréalière :
une approche non
paramétrique**

Mots-clés:

économies d'échelle,
production céréalière,
régression non
paramétrique

Summary – The analysis of the minimum efficient scale plays a crucial role in the study of sector concentration as emphasized by Sutton (1992). This paper uses Baumol, Panzar, and Willig (1982) framework to deal with the measurement of minimum efficient scale by estimating cost scale economies. The estimation of these last ones involves not only the estimation of a multiproduct cost function, but also the estimation of its partial derivatives with respect to input prices. In this paper, we propose the use of the nonparametric regression methods to deal with these two issues. Indeed, these methods do not require a priori the specification of the functionals of interest, i.e. the cost function and its derivatives, nor strong assumptions on the error structure. This « freedom » is particularly interesting with regard to the relevance of the results in empirical applications.

The preceding set up allows us to compute the scale economies for each possible combination of outputs, and thus, to determine the set of these combinations at which the scale economies are exhausted. From this set, named M-locus, we can derive the efficient structure of the studied sector by computing the optimal number of firms that can provide the same total amounts of outputs in a context where firms are cost minimizers.

This approach is implemented using a cross-section of 329 French farms in 1986. These farms produce mainly grains. From the nonparametric estimation of a two-output (grain, other productions) cost function and its partial derivatives with respect to input prices, we find then a large majority of farms with increasing scale economies. Moreover, the farms have not, on average, a sufficient size. Such results agree with the observation that the French grain production should move to stronger concentration. However, the determination of the optimal number of cost-minimizing farms is less conclusive. Indeed, the total number of farms in the sample is within the two bounds that can be given to this optimal number from the estimated M-locus. First, we can note that, according with the previous results, configurations with a smaller number of farms are possible. Second, the determination of the optimal number is sensitive to the chosen number of outputs. For instance, the estimation of a single output cost function gives rise to an optimal number of cost-minimizing farms which is smaller than the observed number of farms.

Résumé – La détermination de la taille minimale efficace d'une unité de production joue un rôle crucial dans l'étude de la concentration d'un secteur productif. Cet article propose d'utiliser le cadre d'analyse développé par Baumol, Panzar et Willig (1982), en vue de mesurer, via l'estimation des économies d'échelle, la taille minimale efficace d'une telle unité. Le recouvrement des économies d'échelle y est réalisé en utilisant la technique non paramétrique appelée méthode du noyau de convolution afin d'estimer une fonction de coût multiproduit et ses dérivées partielles par rapport aux prix des inputs.

Cette approche est illustrée à partir d'un échantillon d'exploitations céréalières françaises en 1986. Celles-ci sont ainsi en majorité caractérisées par des économies d'échelle croissantes. Ce résultat suggère qu'en moyenne, ces exploitations ont une taille insuffisante. De plus, leur nombre apparaît trop élevé par rapport à celui correspondant à une structure efficace du secteur pour la production totale observée. Ces résultats semblent corroborer la tendance observée vers une plus grande concentration dans la production céréalière.

* Station d'économie et sociologie rurales de l'INRA, Chemin de Borde Rouge, Auzeville, BP 27, 31326 Castanet-Tolosan cedex.

DANS un numéro récent des *Cahiers d'Economie et de Sociologie Rurales* consacré à la production céréalière, Blom (1991) met en évidence le fait que cette production est, d'une part, relativement concentrée dans la CEE, un quart des exploitations produisant des céréales détenant 75 % des superficies emblavées, et, d'autre part, est appelée à évoluer vers une plus grande concentration durant les dix prochaines années. Blom base cette dernière constatation sur l'incapacité à dégager un revenu suffisant d'une proportion importante d'exploitations céréalières dans certaines régions de la CEE, du fait de leur faible taille. La mesure de la taille minimale à partir de laquelle une exploitation céréalière serait viable en termes de revenus, ou taille minimale efficace, semble donc être cruciale en vue d'analyser les évolutions observées dans cette production.

L'objet de cet article* n'est pas de fournir une analyse complète des déterminants de la taille minimale efficace d'une exploitation céréalière. Il est de montrer comment certains outils développés en économie industrielle peuvent être utilisés afin d'analyser l'évolution de la production céréalière. Ainsi, récemment, Sutton (1992) a montré le rôle crucial joué par la détermination de la taille minimale efficace d'une unité de production dans l'étude de la concentration d'un secteur productif. Le degré de concentration d'un secteur serait donc lié positivement à la taille minimale efficace d'une entreprise y figurant, cette taille étant définie à partir du niveau de production auquel le coût moyen est soit égal à son minimum, soit le plus proche de ce dernier. Or un tel niveau de production correspond à des économies d'échelle dans la production proches de un. Si l'on dispose de données sur les coûts de production des entreprises dans le secteur étudié, il est alors possible d'utiliser le cadre d'analyse développé par Baumol, Panzar et Willig (1982) afin de déterminer, via l'estimation des économies d'échelle, la taille minimale efficace d'une entreprise. Ce cadre d'analyse, qui repose sur l'étude de la fonction de coût pour des technologies multiproduits, permet de recouvrir l'ensemble des combinaisons possibles dans la production telles que les économies d'échelle soient égales ou proches de un. La connaissance de cet ensemble rend possible la détermination de la taille minimale efficace d'une entreprise, ou, plus précisément, celle de bornes limitant le nombre d'entreprises efficaces dans la configuration étudiée.

* La réalisation de cet article a bénéficié d'un financement du Conseil scientifique du Département Economie et sociologie rurales de l'INRA. Les auteurs remercient Yves Léon, David Martimort et Quang Vuong, ainsi que les deux lecteurs anonymes, pour leurs commentaires sur cet article.

Les études empiriques faisant appel au cadre théorique développé par Baumol, Pangar et Willig⁽¹⁾ se sont largement appuyées sur une spécification paramétrique de la fonction de coût multiproduits. Ainsi, par exemple, Kaserman et Mayo (1991) utilisent une fonction de coût quadratique en vue d'étudier la structure efficace de la production d'électricité aux Etats-Unis. Les résultats empiriques alors obtenus dépendent fortement de la forme fonctionnelle choisie: translogarithmique, quadratique, ... Pour se libérer de cette contrainte, nous proposons une approche alternative basée sur l'utilisation des techniques d'estimation non paramétrique de la régression⁽²⁾. Ces techniques statistiques n'exigent ni la spécification de la forme fonctionnelle étudiée (ici, la fonction de coût), ni des hypothèses restrictives sur le terme d'erreur du modèle de régression. Ce sont les données qui «révèlent» le modèle de régression qui les a générées. Cette «liberté» envers le choix de la forme fonctionnelle est particulièrement intéressante quant à la pertinence des résultats empiriques obtenus et de leur interprétation.

Dans cet article, nous utilisons la technique d'estimation non paramétrique appelée méthode du noyau de convolution (ci-après, méthode du noyau) afin de: 1 – estimer une fonction de coût multiproduit à partir d'un échantillon de producteurs céréaliers français, 2 – construire l'ensemble des combinaisons d'outputs telles que les économies d'échelle sont égales à un, ou *M*-lieu, et déterminer à partir de la connaissance de cet ensemble, la taille minimale efficace d'un producteur céréalier, ou, plus précisément, donner les deux bornes limitant le nombre d'exploitations efficaces dans notre échantillon.

Le choix de la méthode du noyau tient au fait qu'elle permet d'estimer non seulement la fonction de coût, mais aussi ses dérivées premières par rapport aux outputs, ces dernières intervenant, en effet, dans le calcul des économies d'échelle.

Dans une première partie, nous présentons le cadre théorique de l'approche proposée par Baumol, Pangar et Willig. L'application de ce cadre d'analyse à un échantillon de producteurs céréaliers français fait l'objet de la seconde⁽³⁾. Cette application montre que la taille moyenne d'une exploitation dans notre échantillon est inférieure à la taille minimale efficace déterminée à partir du *M*-lieu. Il en résulte que la production totale de notre échantillon pourrait être assurée par un plus petit nombre d'exploitations. Ces résultats semblent confirmer les observations empiriques faites par Blom (1991).

⁽¹⁾ Voir le tour d'horizon de Panzar (1989).

⁽²⁾ Voir le tour d'horizon récent de Härdle (1991). Un exemple d'application à des données agricoles figure dans Moschini (1990).

⁽³⁾ La présentation de la méthode du noyau est donnée dans l'annexe 1.

STRUCTURE EFFICACE D'UNE PRODUCTION: RAPPELS THÉORIQUES

Sous l'hypothèse de minimisation du coût de la part du producteur, la fonction liant le coût de production aux quantités produites des différents biens et aux prix des facteurs de production, ou fonction de coût, constitue une *statistique suffisante* (McFadden, 1978) pour analyser toutes les caractéristiques importantes d'une technologie: élasticités de substitution, parts des facteurs, ... Si les quantités produites des M biens sont représentées par le vecteur y de dimension $(M \times 1)$ et les prix des P facteurs de production par le vecteur p de dimension $(P \times 1)$ et le coût par C , la fonction de coût s'écrit:

$$C = C(p, y) \quad (1)$$

La fonction $C(.,.)$ ainsi définie doit alors satisfaire les propriétés suivantes: elle doit être (i) non négative et non décroissante, (ii) homogène de degré un, et (iii) concave et continue dans les prix⁽⁴⁾.

Parmi les caractéristiques d'une technologie qu'une fonction de coût permet de mesurer, considérons les économies d'échelle. Celles-ci jouent, en effet, un rôle prépondérant dans l'analyse des déterminants technologiques de la structure d'une entreprise et/ou d'un secteur donné (cf. Baumol *et al.*, 1982). Usuellement, les économies d'échelle sont introduites dans un cadre monoproduit comme mesurant la réponse de la quantité produite d'outputs à un accroissement proportionnel des quantités d'inputs utilisées. On parle alors de rendements d'échelle. Ainsi, si toutes les quantités d'inputs sont multipliées par k et si l'on obtient une quantité d'outputs égale à k' , les rendements d'échelle sont dits croissants, constants ou décroissants selon que k' est respectivement supérieur, égal ou inférieur à k .

Il est possible de définir une mesure équivalente aux rendements d'échelle basée non plus sur la connaissance de la technologie mais sur celle de la fonction de coût. Ainsi, toujours dans le cadre monoproduit, les économies d'échelle peuvent être mesurées par le rapport suivant:

$$S(p, y) = \frac{CM(y)}{Cm(y)} = \frac{C(p, y)}{y \partial C(p, y) / \partial y'} \quad (2)$$

où $CM(.)$ et $Cm(.)$ représentent respectivement la fonction de coût moyen et la fonction de coût marginal. Les économies d'échelle sont alors croissantes, constantes ou décroissantes selon que $S(p, y)$ est respectivement supérieur, égal ou inférieur à un. Sous l'hypothèse de minimisation du coût, cette mesure est équivalente à la mesure usuelle des rendements d'échelle⁽⁵⁾. De plus, cette mesure est égale à un quand les coûts margi-

⁽⁴⁾ Voir Varian (1984) pour une discussion de ces propriétés.

⁽⁵⁾ Une démonstration de ce résultat figure dans Panzar (1989).

nal et moyen sont égaux, *i.e.* lorsque le coût moyen est minimum. Cette égalité correspond à une situation de long terme de concurrence pure et parfaite où le profit des entreprises est nul et où le prix de l'output est égal au coût marginal. Pour un vecteur donné des prix des inputs, la quantité d'outputs correspondant à l'égalité à un de $S(p,y)$ donne alors une indication sur la taille minimale efficace d'une entreprise dans le secteur étudié (cf. Sutton, 1992).

La mesure des économies d'échelle présentée ci-dessus peut être aisément généralisée dans le cadre d'une technologie permettant de produire plusieurs outputs. Cette mesure s'écrit alors :

$$S(p,y) = \frac{C(p,y)}{\sum_{i=1}^M y_i \frac{\partial C}{\partial y_i}} \quad (3)$$

Pour faciliter l'interprétation économique de cette mesure, nous allons nous intéresser au cas où $M = 2$, tout ce qui suit se généralisant aisément à des valeurs $M > 2$. De même, nous omettons dans l'écriture des fonctions de coût le vecteur de prix p , tout le raisonnement étant à prix fixés. Considérons maintenant les couples (y_1, y_2) tels que $y_1/y_2 = k$ où k est fixé de façon arbitraire. Pour chaque valeur de k , on définit ainsi un rayon dans l'orthant positif du plan (y_1, y_2) . Le long d'un tel rayon, on peut définir le coût radial moyen de la sorte :

Définition 1 (Baumol *et al.*, 1982) : Le coût radial moyen, noté $CRM(y_1, y_2)$, se définit comme :

$$CRM(y_1, y_2) = \frac{C(y_1, y_2)}{y_1} \quad (4)$$

où $y_1/y_2 = k$, k fixé.

Soit $y = (y_1, y_2)$. Le coût radial moyen $CRM(y)$ est une fonction croissante (resp. décroissante) en y si $CRM(ty)$ est une fonction croissante (resp. décroissante) du scalaire t , en $t = 1$. De plus, le coût radial moyen est minimum en y si $CRM(y) < CRM(ty)$ pour toute valeur positive $t \neq 1$.

Remarquons alors que le point où le coût radial moyen $CRM(y)$ est minimum correspond à celui où la mesure d'économie d'échelle $S(p,y)$ est égale à un. En effet, considérons le rayon défini par $y_2/y_1 = k$, avec k fixé. Minimisons le coût moyen radial, soit :

$$\min_{y_1} \frac{C(y_1, ky_1)}{y_1} \quad (5)$$

La condition de premier ordre s'écrit :

$$\frac{1}{y_1^2} \left[\left(\frac{\partial C}{\partial y_1} + k \frac{\partial C}{\partial y_2} \right) y_1 - C(y_1, y_2) \right] = 0 \quad (6)$$

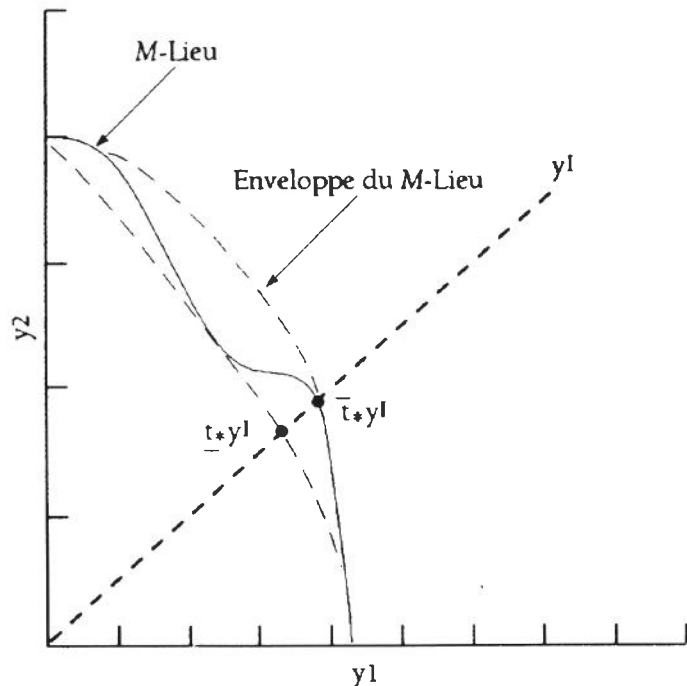
car $y_2 = ky_1$. Alors,

$$\frac{C(y_1, y_2)}{y_1 \frac{\partial C}{\partial y_1} + y_2 \frac{\partial C}{\partial y_2}} = 1 \quad (7)$$

De plus, $S(p, y)$ est supérieur ou inférieur à 1 selon que le coût radial moyen est respectivement décroissant ou croissant.

L'ensemble des points (y_1, y_2) qui minimisent $CRM(y)$ est appelé le *M-lieu*. Le *M-lieu* désigne donc l'ensemble des vecteurs d'outputs qui minimisent le coût radial moyen sur chacun de leurs rayons, ou, de façon équivalente, tels que les économies d'échelle sont égales à un. Remarquons que les points se situant à l'intérieur du *M-lieu* correspondent à des économies d'échelle supérieures à un et les points se situant à l'extérieur à des économies d'échelle inférieures à un (voir figure 1).

Figure 1.
M-lieu, \underline{t} et \bar{t}



L'interprétation économique du *M-lieu* est la suivante. Les points lui appartenant correspondent à une situation de long terme d'équilibre concurrentiel et de profit nul pour les entreprises, avec des prix égaux aux coûts marginaux. Si les entreprises opèrent à des niveaux d'outputs auxquels les économies d'échelle ne sont pas égales à un, ceci peut s'ex-

pliquer par une taille insuffisante du marché relativement aux niveaux d'outputs auxquels le coût radial moyen atteint son minimum. Dans ce cas, une situation de concurrence avec prix égaux aux coûts marginaux n'est pas viable financièrement. Par exemple, quand $M = 1$, si $S(p, y) > 1$ (ou encore, $C(y) > y \partial C(y)/\partial y$), un prix égal au coût marginal ne permet pas de couvrir le coût de production.

Le M -lieu possède une seconde interprétation. En effet, il permet de déterminer la structure efficace d'une branche via le calcul de la taille minimale efficace d'une entreprise. Cette taille s'obtient de la façon suivante. Soit $[\ddot{M}]$ l'enveloppe convexe du M -lieu. Notons $y^I = (y_1^I, y_2^I)$, le vecteur de production totale de cette industrie avec $(y_1^I = \sum_{i=1}^N y_{1i}$ et $y_2^I = \sum_{i=1}^N y_{2i}$, où N est le nombre d'entreprises dans cette industrie.

Soit :

$$\bar{t} \equiv \max \{t; ty^I \in [\ddot{M}]\} \quad (8)$$

et

$$\underline{t} \equiv \min \{t; ty^I \in [\ddot{M}]\} \quad (9)$$

La figure 1 permet de déterminer ces deux nombres. Connaissant ces nombres, on peut alors donner une borne inférieure et une borne supérieure à la taille minimale efficace d'une entreprise dans la branche considérée. Ces deux bornes indiquent que l'entreprise de taille efficace produit dans un voisinage du M -lieu et est ainsi caractérisée par des économies d'échelle proches de un.

Connaissant t et \bar{t} , il est alors possible de calculer le nombre d'entreprises efficaces pour produire les quantités totales d'outputs y^I .

Proposition 1 (Baumol *et al.*, 1982)⁽⁶⁾: Le nombre d'entreprises efficaces pour produire les quantités totales d'outputs y^I , noté $m(y^I)$, est tel que :

1. $m(y^I) \geq [1/2 \bar{t}]$
2. $\begin{cases} m(y^I) = 1 & \text{si } \underline{t} \geq 1 \\ 1 \leq m(y^I) \leq [2/\underline{t}] & \text{sinon,} \end{cases}$

où $[.]$ dénote la partie entière.

Connaissant $m(y^I)$, il est alors possible de savoir si la configuration de secteur étudiée est naturellement monopolistique, oligopolistique ou concurrentielle pour un niveau d'output donné y^I . De plus, lorsque la structure efficace est de nature concurrentielle, *i.e.* avec un grand nombre d'entreprises $m(y^I)$, on peut comparer ce nombre au nombre observé d'entreprises dans le secteur. Si ce dernier est plus grand que $m(y^I)$, cela signifie qu'il y a trop de firmes pour produire y^I . Les mêmes quantités

⁽⁶⁾ La démonstration se trouve dans l'ouvrage de Baumol *et al.* (1982; voir appendice III du chapitre 5).

d'outputs pourraient alors être produites par un nombre moindre d'entreprises. La présence d'un trop grand nombre de firmes dans le secteur étudié pourrait provenir d'une protection trop importante de celui-ci. De même, si le nombre observé est plus petit que $m(y^I)$, cela signifie que les mêmes quantités d'outputs pourraient être produites par un plus grand nombre d'entreprises. Dans ce cas, l'existence de barrières à l'entrée et un comportement collusif entre les entreprises déjà présentes sur le marché peuvent expliquer le nombre insuffisant d'entreprises observé.

APPLICATION AUX EXPLOITATIONS CÉRÉALIÈRES

Comme nous l'avons vu dans la section précédente, la détermination du M -lieu permet d'analyser la structure efficace d'une branche donnée. L'estimation de cet ensemble peut être obtenue via l'estimation des économies d'échelle pour toutes les combinaisons d'outputs envisageables. Une telle estimation nécessite celle de la fonction de coût multiproduit et de ses dérivées premières par rapport aux quantités. Nous proposons d'utiliser les méthodes d'estimation non paramétrique en vue de l'estimation de ces fonctions. En effet, remarquons que la fonction de coût peut s'écrire comme la forme fonctionnelle d'intérêt dans le modèle de régression suivant :

$$C_i = C(p_i, y_i) + \varepsilon_i, \quad i = 1, \dots, N \quad (10)$$

où C_i , y_i et p_i représentent respectivement l'observation du coût du i ème individu, celles des quantités d'outputs produites par cet individu et les prix des inputs auxquels il fait face. ε_i est un terme aléatoire tel que :

$$E(\varepsilon_i | p_i, y_i) = 0 \quad (11)$$

L'objet des méthodes d'estimation non paramétrique est l'estimation d'une fonction d'intérêt, ici, la fonction de coût, sans faire d'hypothèses paramétriques quant à la forme de celle-ci⁽⁷⁾. De plus, l'hypothèse faite en (11) est la seule hypothèse émise quant au terme d'erreur ε_i . Une telle démarche libère ainsi l'économètre du choix de la forme de la fonction à étudier⁽⁸⁾ et est cohérente avec la théorie économique qui, en général, n'établit que les propriétés générales de la fonction étudiée⁽⁹⁾.

L'application présentée ici a pour objet d'étudier les économies d'échelle dans les exploitations céréalières. Nous présentons donc d'abord les données que nous avons utilisées. Les résultats des estimations sont en-

⁽⁷⁾ Pour une discussion de l'approche paramétrique usuelle en économétrie de la production, voir Bonnieux (1989).

⁽⁸⁾ Comme le souligne Bonnieux (1989), ce choix porte en général sur une fonction translog (voir, par exemple, Bureau et Vermersch, 1988).

⁽⁹⁾ Voir, par exemple, les propriétés d'une fonction de coût telles qu'elles ont été énoncées ci-dessus.

suite donnés et commentés. Les aspects techniques, concernant soit la méthode d'estimation non paramétrique que nous avons utilisée (la méthode du noyau de convolution), soit sa mise en œuvre sur l'échantillon dont nous disposons, font l'objet des annexes 1 et 2.

Les données

Les données que nous utilisons proviennent du Réseau d'information comptable agricole (RICA). Pour l'année 1986, nous disposons d'un échantillon de 329 exploitations dont la principale production est céréalière (en moyenne sur l'échantillon 73 % de leur production). Les céréales produites par ces exploitations sont en majorité du blé et du maïs (voir tableau 1).

Trois inputs sont considérés ($P = 3$): le capital, le travail et les consommations intermédiaires, notés respectivement K , L , et M . La méthode des indices de Törnqvist permet d'agréger les données de comptabilité en trois indices de prix des facteurs⁽¹⁰⁾, notés respectivement p_K , p_L , p_M . Les prix des différents biens pris en compte dans le capital sont évalués à partir des bilans, puis agrégés. Notons que nous avons considéré la terre comme l'un de ces biens. L'indice de prix du travail tient compte du travail salarié et du travail familial. En effet, le RICA fournit pour chaque exploitation les nombres d'heures de travail familial et de travail salarié, ainsi que le montant des salaires. Nous supposons alors que le prix du travail familial peut être évalué à partir du prix moyen du travail salarié du quantile des exploitations ayant les plus hauts salaires. Quant aux consommations intermédiaires, elles recouvrent tous les biens qui sont complètement renouvelés sur l'exploitation chaque année, à savoir les engrais, les semences ...⁽¹¹⁾.

Nous considérons un cas biproduct ($M = 2$): les céréales et les autres productions. Afin d'éviter des problèmes complexes d'agrégation, nous avons préféré utiliser les valeurs des produits au lieu des quantités. De plus, les coûts de production correspondants sont donnés par le RICA.

Ainsi, 6 variables sont associées à chaque exploitation: le coût total de production C , la valeur de l'output en céréales y_1 , la valeur de l'output y_2 représentant toutes les autres productions de l'exploitation, et les prix des trois inputs p_K , p_L , p_M . Quelques statistiques portant sur ces variables et d'autres caractéristiques de l'échantillon sont présentées dans le tableau 1.

⁽¹⁰⁾ Voir le tour d'horizon récent de la théorie des indices par Diewert et Nakamura (à paraître).

⁽¹¹⁾ De plus amples détails sur l'agrégation des données figurent dans Monier-Dilhan et Ossard (1990) et Perrigne (1991).

Tableau 1.
Caractéristiques de
l'échantillon (n = 329)

Variables	Moyenne	Ecart-type	Maxi	Mini
y_1 (francs)	375 277	243 764	1 451 986	32 991
y_2 (francs)	144 107	128 227	985 161	0
P_K	1,24	0,83	5,62	0,02
P_L	0,99	0,17	3,04	0,82
P_M	1,66	0,33	2,79	0,69
Coût (francs)	548 410	291 868	2 132 751	117 282
Superficie (hectares)	66,15	38,55	296,00	13,30
% céréales	73	14,24	100	16,42
% blé dans y_1	46,21	26,43	100	0
% maïs dans y_1	33	32,37	100	0

Résultats

Sans perte de généralité, nous considérons une fonction de coût satisfaisant la propriété d'homogénéité de degré un dans les prix. Cette hypothèse est usuelle dans les études empiriques des fonctions de coût. En effet, elle est cohérente avec l'hypothèse que le producteur minimise son coût de production. Nous estimons donc une fonction de coût de la forme suivante :

$$C \left(y_1, y_2, \frac{p_K}{p_L}, \frac{p_M}{p_L} \right) \quad (12)$$

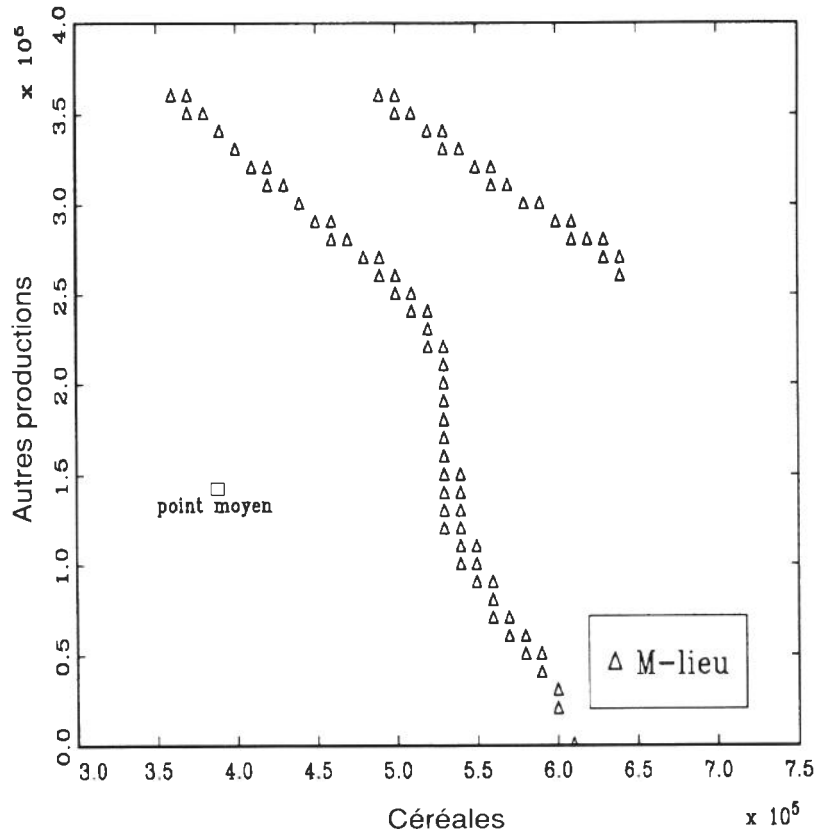
p_K/p_L et p_M/p_L étant respectivement les prix relatifs du capital et des consommations intermédiaires par rapport au travail.

La fonction de coût ainsi que ses dérivées premières par rapport à y_1 et y_2 sont alors estimées en utilisant la méthode du noyau sur une grille de 100×100 valeurs de y_1 et y_2 , respectivement comprises entre [300 000, 700 000] et [0, 400 000]. Quant aux prix relatifs des inputs, ils sont fixés à leurs valeurs moyennes, notées \bar{p}_K/\bar{p}_L et \bar{p}_M/\bar{p}_L , comme il est usuel de le faire dans les approches paramétriques (cf. Akridge et Hertel, 1986). Les estimations des économies d'échelle sont alors données par :

$$\hat{S} = \frac{\hat{C} \left(y_1, y_2, \left(\frac{\bar{p}_K}{\bar{p}_L} \right), \left(\frac{\bar{p}_M}{\bar{p}_L} \right) \right)}{y_1 \frac{\partial \hat{C}}{\partial y_1} \left(y_1, y_2, \left(\frac{\bar{p}_K}{\bar{p}_L} \right), \left(\frac{\bar{p}_M}{\bar{p}_L} \right) \right) + y_2 \frac{\partial \hat{C}}{\partial y_2} \left(y_1, y_2, \left(\frac{\bar{p}_K}{\bar{p}_L} \right), \left(\frac{\bar{p}_M}{\bar{p}_L} \right) \right)} \quad (13)$$

où \hat{C} , représente l'estimateur de la fonction de coût (...), $\partial \hat{C} / \partial y_1$ et $\partial \hat{C} / \partial y_2$ les estimateurs de ses dérivées partielles par rapport à y_1 et y_2 .

Figure 2.
Le M-lieu estimé

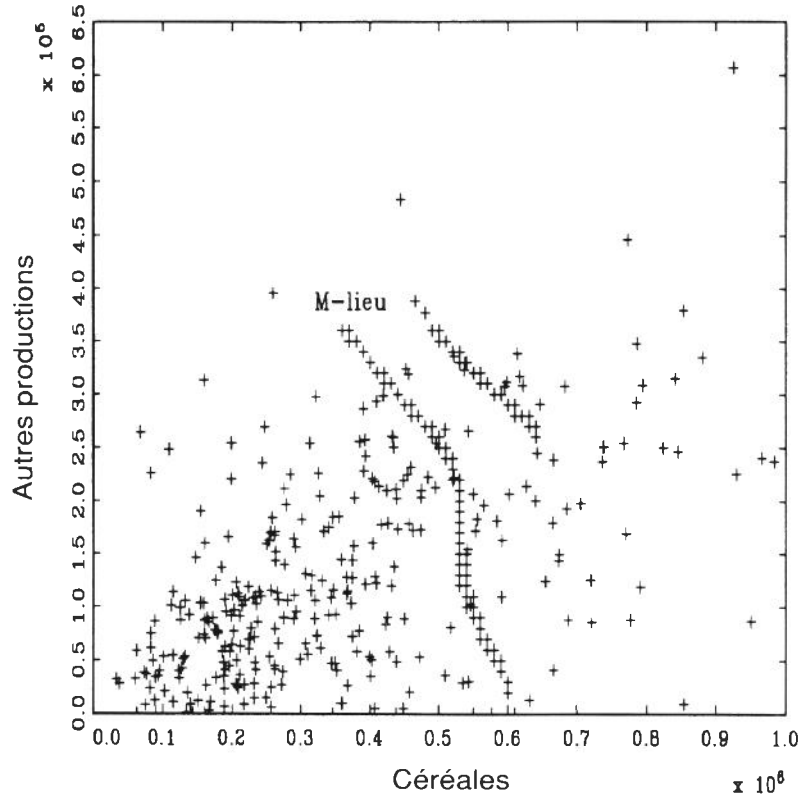


Les estimations des économies d'échelle obtenues nous permettent de construire la frontière où celles-ci sont égales à un, ou *M*-lieu. La figure 2 reporte le lieu des combinaisons des deux outputs telles que les économies d'échelle estimées sont comprises entre 0,99 et 1,01, comme estimation du *M*-lieu. Il est à noter que le *M*-lieu n'atteint pas l'axe y_2 étant donné l'absence d'observations dans notre échantillon correspondant à une quantité nulle de l'output y_1 (production de céréales). La présence de points où les économies d'échelle sont proches de un dans l'angle nord-est de la figure 3 peut ne pas être prise en compte. En effet, elle peut s'interpréter comme un retournement dans les valeurs des économies d'échelle dû à la rareté des observations dans cette région. Il s'agit donc d'un phénomène de bord fréquent en estimation non paramétrique.

Il est intéressant d'étudier comment se situent les exploitations par rapport au *M*-lieu estimé. Rappelons qu'à l'intérieur du *M*-lieu, les économies d'échelle sont supérieures à un, alors qu'à l'extérieur de celui-ci, elles sont inférieures. La figure 3 présente les combinaisons d'outputs observées des 329 exploitations et donne leurs positions par rapport au

M-lieu estimé. Une large majorité de ces exploitations est ainsi caractérisée par des économies d'échelle supérieures à un (environ 90 % de l'échantillon). D'une part, une conclusion possible serait que les exploitations figurant dans notre échantillon ne sont pas viables financièrement dans une situation de long terme d'équilibre concurrentiel où les prix des outputs sont égaux aux coûts marginaux. D'autre part, elles ne semblent pas produire à une échelle efficace. En effet, si nous reportons le point moyen représentant les quantités moyennes produites des outputs, ou (y_1, y_2) , sur la figure 2, nous constatons qu'en moyenne la taille des exploitations est trop petite. Les exploitations devraient produire à une plus grande échelle.

Figure 3.
Le lieu des productions
observées



Nous pouvons également étudier la configuration optimale de cet échantillon au sens de Baumol, Panzar et Willig en calculant le nombre optimal d'entreprises $m(y^1)$. Plus précisément, nous pouvons évaluer \underline{t} et \bar{t} . Des calculs élémentaires donnent $\underline{t} = 0,0038$ et $\bar{t} = 0,0044$. Pour de telles valeurs, la structure optimale du secteur correspondrait à un nombre d'exploitations compris entre 113 et 526. Quoique cet intervalle contienne le nombre d'exploitations de notre échantillon, *i.e.* 329, il permet d'envisager un nombre largement infé-

rieur à 329 comme nombre optimal d'exploitations nécessaires à la production des quantités totales d'output y^1 . Ce résultat irait donc dans le sens des résultats obtenus par simple comparaison de la taille moyenne des exploitations et du M -lieu.

Remarquons que si nous avons considéré le cas d'un seul output et si nous avons agrégé en conséquence la production céréalière et les autres productions, le nombre optimal d'exploitations aurait été 240. Ce nombre peut être calculé de manière très simple par l'intermédiaire du quotient entre l'output total du secteur et l'output qui minimise la fonction de coût uniproduit (voir figure 4 de l'annexe 3). Dans un cadre d'hypothèses fortes permettant l'agrégation des outputs, le résultat énoncé indiquerait que la production totale de notre échantillon pourrait être assurée par un moindre nombre d'exploitations. Remarquons ainsi que l'introduction d'un cadre multiproduit conduit à une conclusion plus nuancée.

CONCLUSION

L'approche développée par Baumol *et al.* (1982) a pour objet l'étude des déterminants technologiques de la structure des entreprises et des marchés. Une telle approche peut être aisément utilisée pour étudier la structure optimale d'une production agricole donnée (ici, la production céréalière) et dégager des éléments d'analyse de l'évolution possible de celle-ci. Ainsi, quoique nous ne puissions pas apporter de conclusion définitive sur la taille optimale du marché céréalière, notre analyse suggère que la taille des exploitations céréalières est en moyenne inférieure à celle qui permettrait d'épuiser les économies d'échelle. Les résultats obtenus pourraient alors expliquer les tendances observées de croissance de la taille des exploitations et de concentration de la production céréalière. Néanmoins, une telle analyse devrait être affinée par la prise en compte des aspects dynamiques des choix de production en agriculture, en particulier les choix en matière de capital (cf. Fernandez-Correjo *et al.*, 1992). Il s'agit là d'une piste pour de futures recherches.

Cet article illustre aussi l'utilité des méthodes d'estimation non paramétrique dans les études empiriques liées à l'estimation de fonction de coût multiproduit. D'une part, les résultats obtenus ne dépendent plus du choix d'une forme particulière de cette fonction. D'autre part, ces méthodes permettent d'estimer non seulement la fonction de coût mais aussi ses dérivées en différents points et d'étudier ainsi l'allure générale de ces fonctions et de fonctions s'en déduisant.

BIBLIOGRAPHIE

- AKRIDGE (J.T.) et HERTEL (T.W.), 1986 — Multiproduct cost relationships for retail fertilizer plants, *American Agricultural Journal of Economics*, 68, pp. 928-938.
- BAUMOL (W.J.), PANZAR (J.-C.) et WILLIG (R.D.), 1982 — *Contestable Markets and the Theory of Industry Structure*, New York, Harcourt, Brace and Jovanovich.
- BIERENS (H.J.), 1987 — Kernel estimators of regression functions, in: BEWLEY (T.F.), *Advances in Econometrics*, Cambridge, Cambridge University Press, pp. 99-144.
- BLOM (J.C.), 1991 — Une évaluation des instruments de la politique céréalière de la CEE, in: CARLES (R.) et VERMERSCH (D.), eds, *Marchés, revenus et politique céréalière*, *Cahiers d'Economie et Sociologie Rurales*, 18-19, pp. 19-50.
- BONNIEUX (F.), 1989 — Quelques orientations de recherche en économie de la production, *Cahiers d'Economie et Sociologie Rurales*, 11, pp. 23-33.
- BOSQ (D.) et LECOUTRE (J.-P.), 1987 — *Théorie de l'estimation fonctionnelle*, Paris, Economica.
- BUREAU (J.-C.) et VERMERSCH (D.), 1988 — Substituabilité entre facteurs de production dans les exploitations céréalières, *Cahiers d'Economie et Sociologie Rurales*, 8, pp. 27-45.
- DELECROIX (M.) et SIMIONI (M.), 1991 — Cost Functions and Optimizing Behavior: a Nonparametric Approach, Toulouse, INRA-ESR, WP 91-13.
- DIEWERT (W.E.) et NAKAMURA (A.O.) — *Essays in Index Number Theory*, Amsterdam, North-Holland, vol.1, à paraître.
- FERNANDEZ-CORNEJO (J.), GEMPESAW II (C.M.), ELTERICH (J.G.) et STEFANOU (S.E.), 1992 — Dynamic measures of scope and scale economies: an application to German agriculture, *American Journal of Agricultural Economics*, 74, pp. 329-342.
- HÄRDLE (W.), 1990 — *Applied Nonparametric Regression*, Cambridge, Cambridge University Press.
- KASERMAN (D.L.) et MAYO (J.W.), 1991 — The measurement of vertical economies and the efficient structure of the electric utility industry, *Journal of Industrial Economics*, 34, pp. 483-502.

- McFADDEN (D.), 1978 — Cost, revenue, and profit functions, in: FUSS, (M.) et McFADDEN (D.), eds, *Production Economics: A Dual Approach to Theory and Applications*, Amsterdam, North-Holland, pp. 3-108.
- MONIER-DILHAN (S.) et OSSARD (H.), 1990 — Méthode de construction d'une base de données de panel pour l'économie de la production, Toulouse, INRA-ESR, WP 90-01.
- MOSCHINI (G.), 1990 — Nonparametric and semiparametric estimation: an analysis of multiproduct returns to scale, *American Journal of Agricultural Economics*, 74, pp. 329-342.
- PANZAR (J.-C.), 1989 — Technological determinants of firm and industry structure, in: SCHMALENSEE (R.) et WILLIG (R.D), eds, *Handbook of Industrial Organization*, Amsterdam, Elsevier Science Publishers B.V., pp. 3-59.
- PERRIGNE (I.), 1991 — Frontière de production et mesure d'efficacité technique: applications aux exploitations céréalières, Thèse de doctorat, Université des sciences sociales, Toulouse.
- SUTTON (J.), 1992 — *Sunk Costs and Market Structure*, Cambridge, MIT Press.
- VARIAN (H.R.), 1984 — *Microeconomic Analysis*, New-York, Norton.
- VINOD (H.D.) et ULLAH (A.), 1988 — Flexible production function estimation by nonparametric kernel estimators, in: RHODES (G.F.) et FOMBY (T.B.), eds, *Advances in Econometrics: Nonparametric and Robust Inference*, Greenwich, JAI Press Inc., pp. 139-160.

ANNEXE 1

MÉTHODE D'ESTIMATION DU NOYAU

Définitions et propriétés^(1,2)

Considérons un vecteur de $p + 1$ variables continues aléatoires noté (x_1, \dots, x_p, z) . Notre objectif est d'établir la relation entre la variable endogène z et le vecteur des p variables exogènes noté x (x_1, \dots, x_p). Nous cherchons ainsi à caractériser la fonction $R(\cdot)$ telle que z soit le plus proche de $R(x)$. Choisissons comme critère de proximité l'erreur intégrée quadratique moyenne définie par $E[(z - R(x))^2]$. La minimisation de ce problème donne comme fonction $R(x)$, l'espérance conditionnelle de z sachant x , notée $E(z|x)$. Cette solution s'écrit sous la forme :

$$R(x) = E(z|x) = \frac{\int z f_{x,z}(x,z) dz}{f(x)} \quad (\text{A.1})$$

où $f_{x,z}(x,z)$ est la densité jointe du vecteur (x,z) et $f(x)$ est la densité marginale du vecteur x .

Un choix simple pour estimer $R(\cdot)$ au point x est l'estimateur du noyau de convolution. Considérons ainsi n observations notées $(X_i, Z_i)_{i=1}^n$ du vecteur (x,z) , cet estimateur peut être défini de la sorte :

Définition 1 : L'estimateur du noyau de $R(\cdot)$ au point x , noté $R_n(x)$, s'écrit :

$$R_n(x) = \frac{\sum_{i=1}^n Z_i K\left(\frac{x-X_i}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right)} \quad (\text{A.2})$$

où la fonction $K(\cdot)$, appelée noyau, est une fonction continue, bornée, symétrique et à valeurs réelles telle que $\int K(u) du = 1$. Quant au scalaire h_n , appelé fenêtre, il détermine la taille des poids $K\left(\frac{x-X_i}{h_n}\right) / \sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right)$ comme une fonction de la taille de l'échantillon n . Plus précisément, la valeur de la fenêtre influe sur les pondérations des observations prises en compte dans le calcul de l'estimateur du noyau de $R(\cdot)$ au point x . Ces

^(1,2) Pour une présentation plus détaillée de cette méthode d'estimation non paramétrique, nous renvoyons le lecteur aux ouvrages de Bosq et Lecoutre (1987) et Härdle (1990).

pondérations décroissent lorsque sont considérées des observations de plus en plus éloignées du point où est réalisée l'estimation.

Les propriétés de l'estimateur du noyau sont détaillées dans les *surveys* de Bierens (1987) et Härdle (1990). Ainsi, cet estimateur est convergent quand n tend vers l'infini. Ce résultat de convergence impose des conditions sur le choix de la fenêtre h_n : quand n tend vers l'infini, celle-ci doit tendre vers zéro et nh_n^p vers l'infini.

Estimateur du noyau de la dérivée

Considérons maintenant l'estimation des dérivées partielles de $R(x)$ par rapport aux composantes du vecteur x . En utilisant les notations de Vinod et Ullah (1988), notons $pd(x)$ la dérivée partielle de $R(x)$ par rapport à x_j , soit :

$$pd(x) = \frac{1}{f(x)} [g'(x) - R(x) f'(x)] \quad (A.3)$$

où $g(x)$ est le numérateur de $R(x)$.

La méthode de l'estimation du noyau peut alors être utilisée pour estimer la dérivée de la fonction de régression $R(x)$. L'estimateur du noyau de la dérivée est obtenu en différenciant $R_n(x)$. Remarquons ainsi que $R_n(x)$ peut être écrit sous la forme suivante :

$$R_n(x) = \frac{g_n(x)}{f_n(x)} \quad (A.4)$$

où $g_n(x)$ est égal à $1/(nh_n^p)$ fois le numérateur de (A.1) et $f_n(x)$ est égal à $1/(nh_n^p)$ fois le dénominateur de (A.1). La dérivée partielle de $R_n(x)$ par rapport à x_j , notée $pd_n(x)$, s'écrit alors :

$$pd_n(x) = \frac{1}{f_n(x)} (g'_n(x) - R_n(x) f'_n(x)) \quad (A.5)$$

avec

$$g'_n(x) = \frac{\partial g_n(x)}{\partial x_j} = \frac{1}{nh_n^{p+1}} \sum_i Y_i K' \left(\frac{x - X_i}{h} \right) \quad (A.6)$$

$$f'_n(x) = \frac{\partial f_n(x)}{\partial x_j} = \frac{1}{nh_n^{p+1}} \sum_i K' \left(\frac{x - X_i}{h} \right) \quad (A.7)$$

Definition 2: L'estimateur du noyau de $pd(x)$ est défini par $pd_n(x)$ donné en (A.5), (A.6) et (A.7).

Vinod et Ullah (1988) ont établi deux propriétés de $pd_n(x)$. D'une part, ils ont montré la convergence de $pd_n(x)$ vers $pd(x)$. D'autre part, ils ont

donné un résultat ponctuel de normalité asymptotique que Delecroix et Simioni (1991) ont généralisé à k points distincts. Ainsi, ils ont montré que, sous certaines hypothèses techniques usuelles, si l'on considère k vecteurs distincts $x_j, j = 1, \dots, k$, alors le vecteur dont les k composantes sont : $(nb_n^{p+2})^{1/2} (pd_n(x_j) - pd(x_j)), j = 1, \dots, k$, converge en distribution vers une loi normale centrée k -variée. Comme pour la convergence de $R_n(x)$, ce résultat de normalité asymptotique implique des conditions sur le choix de la fenêtre: elle doit être telle que nb_n^{p+2} tend vers l'infini et nb_n^{p+4} tend vers zéro, quand n tend vers l'infini.

ANNEXE 2

ASPECTS TECHNIQUES DE LA MISE EN ŒUVRE DE LA MÉTHODE DU NOYAU

La fonction de coût estimée est de la forme suivante:

$$C(y_1, y_2, \frac{p_K}{p_L}, \frac{p_M}{p_L}) \quad (\text{A.8})$$

Pour l'estimer en utilisant la méthode du noyau présentée dans l'annexe précédente, nous avons utilisé un noyau produit du type:

$$K(u_1, u_2, u_3, u_4) = \prod_{j=1}^4 K_0(u_j) \quad (\text{A.9})$$

avec $K_0(u_j)$ la densité de la distribution normale centrée réduite, soit:

$$K_0(u_j) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u_j^2}{2}\right) \quad (\text{A.10})$$

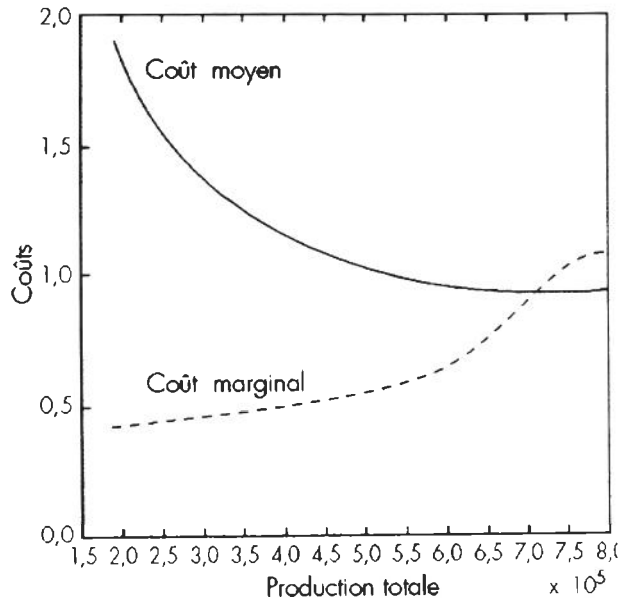
Remarquons qu'à la différence du choix de la fenêtre qui joue un rôle crucial dans la précision locale ou globale d'un estimateur du noyau, le choix du noyau influe peu sur les propriétés de ce type d'estimateurs (Bosq et Lecoutre, 1987).

Comme le suggèrent Vinod et Ullah (1988), nous avons utilisé pour chaque variable exogène $j, j = 1, \dots, 4$, une fenêtre de la forme $h_j = h_n \sigma_j$ où σ_j est l'écart-type de la variable exogène j et $h_n = n^{-\alpha}$ avec α égal à 0,15. Cette valeur permet de respecter les contraintes imposées à la valeur de la fenêtre afin d'obtenir la convergence et la normalité asymptotique des estimateurs (Delecroix et Simioni, 1991). De plus, les résultats obtenus sont peu sensibles à des variations de α dans un voisinage de la valeur choisie.

Les estimations ont été réalisées à l'aide du logiciel GAUSS.

ANNEXE 3

Figure 4.
Coût moyen et coût
marginal dans le cas
d'un seul produit



Le nombre de firmes efficaces est égal à $m(y^I) = \frac{y^I}{y^*}$ où y^* correspond au point d'intersection des fonctions de coût marginal et moyen.