

n° 2003-48

**Econométrie de la production
sur données de panel et dispersion
des prix de production :
Quels biais d'estimation ?***

J. MAIRESSE¹
R. DESPLATZ²

Les documents de travail ne reflètent pas la position de l'INSEE et n'engagent que leurs auteurs.

Working papers do not reflect the position of INSEE but only the views of the authors.

* Les auteurs ont bénéficié notamment des remarques faites par Bruno Crépon, Benoît Mulkay, Patrick Sevestre ainsi que des commentaires de trois référés. Ils remercient particulièrement Benoît Mulkay pour son concours à la constitution de l'échantillon apparié aux enquêtes de conjoncture et le Département de la Conjoncture de l'INSEE pour un accès aux fichiers de ces enquêtes.

¹ CREST-INSEE et Ecole des Hautes Etudes en Sciences Sociales

² Direction de la Prévision et Université Panthéon-Sorbonne Paris 1 au moment de la réalisation de l'étude.

Résumé

Econométrie de la production sur données de panel et dispersion des prix de production :
Quels biais d'estimation ?

L'absence d'informations individuelles sur les prix de production est un problème important de l'économétrie de la production, dont on peut notamment penser qu'il explique en partie les disparités importantes entre les estimations de l'élasticité du capital et des rendements d'échelle sur données de panel, suivant qu'elles s'appuient surtout sur les différences inter-entreprises (estimations de type coupe) ou sur les variations intra-entreprises (estimations de type séries temporelles). Dans cette étude, nous avons pu mobiliser de telles informations sur les prix des entreprises, ainsi que sur les taux d'utilisation de leurs capacités de production, pour un panel d'environ 450 entreprises manufacturières françaises sur quatre années. Que l'on utilise les méthodes d'estimation de panels usuelles ou celles des moments généralisés, les idées que l'on pouvait avoir a priori ne sont pas confirmées, et les différentes estimations restent pratiquement inchangées qu'elles prennent en compte ou non la dispersion des variations de prix et des taux d'utilisation des capacités. Sans doute les biais liés aux autres risques d'erreurs de spécification restent-ils largement prédominants.

Abstract

Panel Data Estimates of the Production Function and Output Price Dispersion :
What Are the Biases?

The lack of individual information on output prices is a major problem of the econometrics of production, which in particular could account for a significant part of the large discrepancies found in panel data studies between the cross-sectional and time series estimates of capital and scale elasticities. In this work, we have been able to obtain such information on firm output prices, as well as information on firm rates of capacity utilization, for a panel of about 450 French manufacturing firms over four years. Whether we rely on the usual panel data estimation methods or use generalized methods of moments (GMM), our expectations are not confirmed, and the various estimations remain practically unaffected by taking into account or not the dispersion of changes in output prices and in capacity utilization. It is likely that the biases arising from other sources of specification errors are more important.

Mots clés :

fonction de production ; données de panel ; biais d'estimation ; erreurs de spécification ; dispersion des prix ; taux d'utilisation des capacités.

Key words :

production function; panel data; estimation bias; specification errors; price dispersion; capacity utilisation.

JEL classification: C23; D24.

INTRODUCTION

L'absence, ou l'indisponibilité, de données individuelles sur les prix de production des entreprises est un des problèmes majeurs des analyses micro-économétriques des comportements d'entreprises, évidemment en matière de prix mais aussi de production et de demande de facteurs. Dans le cas de l'économétrie de la production, qui nous intéresse particulièrement ici, la pratique habituelle qui consiste à déflater la production nominale (ou la valeur ajoutée) des entreprises en remplaçant les indices de prix individuels inconnus par les indices de prix des secteurs (ou branches) d'activité de ces entreprises, est une solution a priori fort imparfaite, même quand ces derniers sont disponibles à un niveau fin de nomenclature. Reprenant une analyse qui avait déjà été faite par Marschak et Andrews (1944) dans leur article sur les problèmes d'identification et d'estimation des fonctions de production, Klette et Griliches (1996) rappellent ainsi qu'une telle pratique ne saurait exclure des biais dans l'estimation des paramètres des fonctions de production lorsque les évolutions des prix des entreprises à l'intérieur des secteurs (même étroitement définis) sont notablement dispersées et qu'elles sont corrélées avec celles des facteurs de production. Ils suggèrent notamment que ces biais pourraient être une des raisons des disparités observées sur données de panels d'entreprises entre les différents types d'estimation des élasticités de facteurs. Dans cet article, nous cherchons à savoir si tel est le cas de fait, sachant que d'autres sources de biais sont elles aussi potentiellement importantes. Dans quelle mesure, l'absence des prix individuels de production explique-t-elle que les estimations de panel privilégiant la dimension temporelle (longitudinale) des données soient souvent assez peu vraisemblables, conduisant notamment à de très faibles valeurs des élasticités du capital et à des rendements d'échelle plus ou moins fortement décroissants ? Autrement dit, dans quelle mesure le fait qu'elles ne requièrent pas la connaissance des prix individuels pourrait-il expliquer que les estimations s'appuyant (au moins pour une part) sur la dimension individuelle (en coupe transversale) des données soient en général plus raisonnables ? La connaissance des prix individuels peut elle conduire ainsi à ce que les deux grands types d'estimations soient moins disparates ?

Nous cherchons à répondre à ces questions en vérifiant directement ce qu'il en est pour les principaux types d'estimations de panel que nous avons pu faire sur un panel cylindré d'entreprises, pour lequel nous disposons d'indices de prix de production individuels. Nous avons pu construire un tel panel pour 468 entreprises manufacturières françaises sur la période 1994-

1997. Nous avons pu en effet mobiliser les informations sur les variations de prix demandées directement aux entreprises dans les enquêtes de conjoncture de l'INSEE, et les apparier aux données comptables habituelles sur la production et ses facteurs disponibles par ailleurs. Jusqu'à présent, à notre connaissance, seule l'étude de Abbott (1991) a pu disposer d'informations sur les prix et aborder la question des disparités d'estimation de la fonction de production sur données de panels suivant une démarche assez comparable, mais sans disposer in fine d'un échantillon satisfaisant.¹

Il nous était possible de mobiliser dans les enquêtes de conjoncture des informations sur les taux d'utilisation des capacités de production des entreprises en même temps que celles sur les prix, et il nous a paru souhaitable de les considérer elles aussi. Ces informations sont en effet le plus souvent manquantes dans les études sur panels d'entreprises, et le fait de devoir ignorer les variations des taux d'utilisation des capacités (de même que celles de heures de travail auxquelles elles sont corrélées) est également considéré comme étant une source de biais des estimations s'appuyant sur la dimension temporelle des panels, et de leurs disparités avec celles s'appuyant surtout sur leur dimension individuelle. Bien qu'elles correspondent à des comportements économiques différents, les deux sources de biais sont assez parallèles. Résultant d'une mesure imparfaite à double titre de la productivité (ou en d'autres termes de l'omission dans la fonction de production de deux variables), elles sont l'une et l'autre susceptibles d'affecter surtout les estimations de type temporel. Pouvoir comparer leur incidence effective dans le cadre d'une même étude est de ce fait intéressant.

Dans le même ordre d'idées, nous cherchons à jauger ces deux sources spécifiques de biais en les associant aux trois sources générales de biais potentiels dits d'hétérogénéité, d'endogénéité et d'erreurs (aléatoires) sur les variables. La différence est bien sûr que pour corriger ces biais nous pouvons dans le premier cas espérer tirer parti d'informations qui sont directement pertinentes, même si elles restent imparfaites, ce qui est une situation a priori bien préférable.² Nous devons surtout dans le second cas nous appuyer sur des méthodes d'estimation alternatives, fondées sur des hypothèses d'exogénéité des variables (« conditions d'orthogonalité ») de moins en moins

¹ Cette étude concerne un échantillon de 40 établissements fabriquant du ciment hydraulique aux Etats-Unis pour lesquels l'auteur a pu calculer un prix de production moyen à partir de données directes sur les prix des différentes variétés de ce type de ciment et des livraisons correspondantes pour les trois années 1972, 1977 et 1981 (du Census of Manufacturing). Ne disposant des prix que pour ces trois années, il ne peut néanmoins calculer comme nous le faisons des estimations sur la base des différences premières.

² Voir sur ce point notamment Dagenais, 1992, qui insiste sur l'intérêt de disposer d'informations directes plutôt que de devoir (essayer de) palier à leur absence aux prix d'hypothèses et de méthodes plus ou moins satisfaisantes.

exigeantes, mais qui conduisent souvent à des estimations de plus en plus fragiles à d'autres formes d'erreurs de spécification, ainsi que de moins en moins précises sur des échantillons qui comme le nôtre ne sont pas de très grande taille.

De façon précise, nous sommes amenés ainsi à comparer les estimations d'une fonction de production simple de type Cobb-Douglas dans laquelle la production est déflatée par un indice de prix sectoriel suivant la pratique courante, à celles, en principe préférables, où elle est déflatée par un indice de prix individuel et où elle est corrigée ou non par le taux d'utilisation des capacités de production. Nous faisons ces comparaisons en considérant d'abord les estimations obtenues par les méthodes usuelles (fondées sur des transformations simples des variables et les moindres carrés ordinaires), et ensuite celles obtenues par les méthodes de moments généralisés (GMM).

Nos résultats ne confirment pas les idées qu'on pouvait avoir a priori. Que l'on utilise les méthodes d'estimation plus usuelles ou celles des moments généralisés, ni l'introduction des prix individuels ni celle des taux d'utilisation des capacités dans la fonction de production ne modifient notablement les estimations des élasticités des facteurs. Les disparités entre les estimations des fonctions de production sur données de panels, plus vraisemblables lorsqu'elles s'appuient sur la dimension individuelle de ces données que si elles sont fondées sur leur seule dimension temporelle, ne semblent donc pas avoir là leur origine première.³ Celle-ci resterait surtout liée au fait que les erreurs sur les variables et d'autres formes d'erreurs de spécification, plus complexes que celles prises en compte par les méthodes d'estimation, se trouvent exacerbées dans la dimension temporelle des données, entraînant des biais plus sérieux dans cette dimension que dans leur dimension individuelle.⁴

Dans ce qui suit, nous illustrons d'abord, dans le cadre d'une modélisation simple d'une fonction de production associée à une fonction de demande de l'entreprise, ce que peuvent être les risques de biais liés à l'absence de prix individuels dans l'estimation de la fonction de production (première section). Nous documentons et commentons ensuite les résultats obtenus

³ Notons que cette conclusion n'est pas celle de l'étude de Abbot, 1991. Bien qu'il ne puisse faire effectivement de régressions en différences premières sur les données dont il dispose (cf. note de bas page supra), il suggère néanmoins sur la base d'un raisonnement indirect que s'il avait pu les faire elles auraient sans doute conduits à des estimations comparables à celles faites en coupe. Son argumentation ne nous paraît pas convaincante, d'autant qu'il obtient par ailleurs des estimations en différences longues (sur cinq ans), très comparables aux nôtres, montrant qu'il n'en est rien. Ces dernières ne sont pas présentées dans l'étude publiée mais uniquement dans un des chapitres de la thèse, qui en est à l'origine.

⁴ Sur ces points, voir notamment Mairesse 1990 et Griliches-Mairesse 1998.

sur notre échantillon suivant les différentes méthodes d'estimation de panels mises en œuvre, traditionnelles et plus récentes (deuxième section). Nous concluons brièvement. Le calcul des indices de prix de production individuels des entreprises, ainsi que celui des taux d'utilisation des capacités de production, à partir des enquêtes de conjoncture, sont détaillés en annexe.

1. Les implications de l'absence de prix individuels

Les implications de l'absence de prix individuels de production sur les estimations de fonction de production s'analysent simplement si on admet que les entreprises fonctionnent sur des marchés de concurrence imparfaite où les différences de prix observées traduisent la différenciation de leurs produits.

Supposons pour simplifier que la fonction de production de l'entreprise i à l'année t est de type Cobb-Douglas, et s'écrit :

$$Q_{it} = A_t K_{it}^{\alpha} L_{it}^{\beta} e^{u_{it}} \quad \text{avec } i = 1, \dots, N \text{ et } t = 1, \dots, T$$

où Q , K et L sont respectivement les volumes de la valeur ajoutée (mesurant la production), et ceux du capital et du travail, où u est le terme d'erreurs ou de perturbations, où A_t est le terme de progrès technique autonome ou de productivité totale des facteurs, et où α et β désignent les élasticités par rapport au capital et au travail. En prenant les logarithmes des variables et en les notant par les minuscules correspondantes, cette fonction s'écrit aussi sous la forme de la régression (log) linéaire habituelle:

$$q_{it} = a_t + \alpha k_{it} + \beta l_{it} + u_{it}$$

Ainsi formulée, cette régression suppose que nous utilisons effectivement une mesure en volume de la valeur ajoutée au niveau individuel de l'entreprise. Comme on ne dispose pas en général d'informations sur les prix de la valeur ajoutée (ou de production) à ce niveau, le volume de la valeur ajoutée de l'entreprise est inconnu et est habituellement approximé par la valeur ajoutée nominale déflatée par l'indice de prix du secteur auquel l'entreprise appartient. Au lieu de $q_{it} = (y_{it} - p_{it})$, on mesure donc $(y_{it} - p_{St})$, où y_{it} est la (log) valeur ajoutée nominale, et p_{it} et p_{St} sont respectivement le (log) indice de prix individuel et (log) indice de prix sectoriel (du secteur S de l'entreprise i). L'équation que l'on estime est en fait :

$$(y_{it} - p_{St}) = a_t + \alpha k_{it} + \beta l_{it} + v_{it} ,$$

alors que l'équation que l'on devrait estimer est en réalité :

$$(y_{it} - p_{St}) = a_t + \alpha k_{it} + \beta l_{it} + ((p_{it} - p_{St}) + u_{it}) .$$

La variable de prix est donc omise de l'équation estimée et se retrouve dans le terme de

$$\text{perturbations } v_{it} = (p_{it} - p_{St}) + u_{it} .$$

Il est alors assez simple de voir que les estimations des élasticités du capital et du travail α et β , et donc celle de l'élasticité d'échelle $\mu = \alpha + \beta$, seront biaisées si les prix pratiqués par les entreprises sont, d'une part, notablement dispersés au sein d'un même secteur d'activité et, d'autre part, corrélés avec les facteurs de production.

Ce problème de biais est d'autant plus gênant a priori qu'il est difficile d'espérer pouvoir le traiter avec une méthode d'estimation par variables instrumentales. Il paraît en effet tout particulièrement ardu de trouver des instruments valides dans ce cas, car toute variable corrélée avec les facteurs de production risque également d'être corrélée avec les prix, via la fonction de production de l'entreprise et sa fonction de demande.

Supposons ainsi que la fonction de demande de l'entreprise, résultant de l'imparfaite substituabilité de ses produits avec ceux des entreprises concurrentes du même secteur, est de type Spence-Dixit-Stiglitz, et qu'elle s'écrit simplement en logarithme :

$$(q_{it} - q_{St}) = \eta(p_{it} - p_{St}) + w_{it}$$

où $(q_{it} - q_{St})$ est l'écart entre les logarithmes de la valeur ajoutée nominale (ou production) de l'entreprise et de la valeur ajoutée sectorielle, $(p_{it} - p_{St})$ est l'écart entre les logarithmes des indices de prix de l'entreprise et de prix sectoriel, w_{it} est le terme de perturbations, et où η désigne l'élasticité prix de la demande, en principe négative ($\eta < 0$). Cette fonction signifie qu'en raison de la différenciation de ses produits l'entreprise est capable de capter une fraction supplémentaire de la demande sectorielle en abaissant son prix relativement au prix sectoriel, une baisse de 1% de son prix relatif lui permettant d'accroître sa part de marché de $-\eta\%$. Le terme de perturbations w_{it} est supposé prendre en compte les autres facteurs et les « chocs » susceptibles d'affecter la demande pour les produits de l'entreprise, indépendamment des variations de prix (investissements de recherche et innovation de produits, dépenses de publicité et efforts commerciaux, évolution des revenus des consommateurs, changement de leurs goûts, ...). La fonction de demande inverse exprimant le prix relatif comme une fonction de la part de marché de l'entreprise et des chocs de demande, s'écrit donc de façon équivalente:

$$p_{it} - p_{St} = \eta^{-1}(q_{it} - q_{St} - w_{it})$$

Si on admet que la quantité demandée est égale à la quantité produite, la variable de prix dépend entre autres des volumes de facteurs de production, et on comprend qu'il puisse être particulièrement difficile d'instrumenter la fonction de production en l'absence d'information sur les prix de l'entreprise et d'une mesure satisfaisante de sa production ou valeur ajoutée réelle.

De façon plus précise, on peut exprimer la variable de prix omise ($p_{it}-p_{St}$) en fonction de la valeur ajoutée déflatée sectoriellement ($y_{it}-p_{St}$) au lieu de la valeur ajoutée en volume q_{it} , en réécrivant la fonction de demande inverse sous la forme :

$$p_{it} - p_{St} = (1 + \eta)^{-1}(y_{it} - p_{St} - q_{St} - w_{it})$$

et par substitution dans la fonction de production estimée on peut formuler celle-ci sous la forme :

$$(y_{it} - p_{St}) = \omega^{-1} (a_t + \alpha k_{it} + \beta l_{it}) + \varepsilon_{it}$$

où $\varepsilon_{it} = (\eta^{-1}q_{St} + \omega^{-1}u_{it} + \eta^{-1}w_{it})$ et $\omega = \eta(1+\eta)^{-1}$ est le ratio de marge (ou de pouvoir de marché ou encore mark-up).

Cette nouvelle écriture de la fonction de production telle qu'elle est estimée en pratique en termes de production en valeur (ou en valeur seulement déflatée par un indice de prix sectoriel), et non en volume, nous suggère quelle peut être l'ampleur des biais susceptibles d'affecter les estimations des élasticités de facteurs.⁵ Il apparaît que les estimations de la fonction de production en « valeur » (ou encore « fonction de revenu ») pourraient être biaisées vers le bas par rapport à celles de la fonction de production au sens propre, en proportion inverse au ratio de marge ω , celui-ci étant d'autant plus élevé que l'élasticité prix de la demande est faible (en valeur absolue).

⁵ Elle montre aussi que le terme de perturbations ε est affecté à la fois par les chocs d'offre u et par ceux de demande w , et que donc seules des variables indépendantes de ces deux sortes de chocs peuvent être des instruments valides pour la fonction de production en « valeur ». On remarque aussi que si l'on introduit dans cette fonction de production explicitement la variable q_{St} de production sectorielle, il devient en principe possible d'identifier et estimer le paramètre η d'élasticité de la demande, et donc celui de pouvoir de marché ω , et partant d'en inférer des estimations non biaisées des élasticités α et β du capital et du travail, même en l'absence d'informations sur les prix individuels. C'est là en fait la solution suggérée par Klette et Griliches, 1996, et mise aussi en œuvre par exemple dans Crépon-Desplatz-Mairesse, 1999.

Elles seraient ainsi sous estimées de l'ordre de 25% pour une élasticité prix de l'ordre de -4 et un ratio de marge de l'ordre de 1,25.⁶

Ces conclusions ne sont évidemment valables que si le cadre d'analyse que nous venons de rappeler est approprié, et notamment l'hypothèse d'une fonction de demande de l'entreprise telle celle spécifiée est satisfaisante. Trois observations importantes doivent être faites à cet égard qui correspondent à trois cas où les biais disparaissent, car la corrélation entre prix et facteurs de production est nulle ou négligeable.

La première de ces observations s'applique au cas extrême de concurrence parfaite où l'élasticité prix de la demande tend vers l'infini, et le taux de marge vers zéro (le ratio de marge vers 1). En ce cas, les biais seraient nuls et la dispersion des variations de prix individuels observées entre entreprises refléterait pour l'essentiel des erreurs (aléatoires) de mesure. Ce cas paraît cependant peu réaliste, sauf pour certains secteurs particulièrement concurrentiels et peut être sur longue période.

La deuxième observation correspond au cas où le lien entre volumes et prix qu'exprime la fonction de demande est faible relativement aux autres déterminants et aux chocs de demande (le terme w_{it} l'emportant largement sur la dispersion des prix). En ce cas la corrélation simple (non conditionnelle aux autres déterminants de la demande) entre prix et volumes de production sera faible, et la corrélation transmise aux facteurs de production sera elle aussi faible. Cette dernière pourra donc être plus ou moins négligeable relativement aux corrélations entre les chocs d'offre u_{it} et les facteurs de production, qui peuvent résulter d'autres erreurs de spécification de la fonction de production. C'est là sans doute la raison principale qui explique que nos estimations dans la dimension temporelle des données ne soient en fait que peu modifiées, ainsi que nous le constatons, lorsqu'on tient compte de la dispersion des variations des prix de production individuels.

La troisième observation est particulièrement importante dans le contexte des données de panel puisqu'elle apporte une justification à l'estimation de la fonction de production sur les niveaux des variables, que la production soit d'ailleurs en volume ou en valeur. En coupe, dans la dimension individuelle des données, l'estimation d'une fonction de production de l'entreprise en « volume » *stricto sensu*, c'est-à-dire en termes d'unités ou de quantités physiques, n'a en effet

⁶ Ces ordres de grandeur sont ceux retenus de préférence dans Crépon-Desplatz-Mairesse, 1999, pour un échantillon d'entreprises industrielles françaises comparable à celui de cette étude.

pas de signification, à moins de supposer que l'on se restreint à un secteur où les biens produits sont suffisamment homogènes pour que le volume de leur production puisse être mesuré ainsi et non pas seulement « en volume aux prix d'une année de référence ».⁷ Dans le cas général où les biens produits sont hétérogènes, seule n'acquiert de sens que l'estimation de la fonction de production (de revenu) en « valeur » ou « en volume », où la production est mesurée en unités de valeur (en euros aux prix courants) ou de « volume » (en euros aux prix d'une année de référence). En coupe, à moins de considérer là encore un secteur de biens homogènes, on ne saurait par ailleurs s'attendre à une corrélation significative entre différences de prix et différences de production (en valeur, volume ou quantité). L'estimation en coupe d'une fonction de demande de l'entreprise, contrairement à celle de la fonction de production, ne peut donc quant à elle acquérir une signification pratique. On ne doit pas s'étonner non plus que le fait de déflater ou non par des prix sectoriels ou individuels conduise à des estimations pratiquement inchangées, lorsqu'elles sont faites en coupe, ou plus généralement lorsqu'elles privilégient la dimension individuelle des données.

2. Fonction de production estimée avec et sans prise en compte de la dispersion des prix individuels

2.1. Echantillon, variables et spécification économétrique

En appariant les données comptables habituelles sur la production et ses facteurs, obtenues à partir de la base SUSE (Système Unifié de Statistiques d'Entreprises de l'INSEE), et les informations sur les prix, ainsi que celles sur les taux d'utilisation, que nous avons pu mobiliser dans les enquêtes trimestrielles de conjoncture de l'INSEE, nous avons construit un échantillon cylindré sur quatre années 1994 à 1997 de 468 entreprises industrielles de taille moyenne et grande (20 salariés et plus). Compte tenu des limites de cet appariement puisque les enquêtes de conjoncture ne sont pas exhaustives, et des problèmes liés notamment à l'absence fréquente de réponses dans ces enquêtes, ou encore à l'élimination d'observations extrêmes et vraisemblablement erronées (« outliers »), cet échantillon représente le meilleur compromis que

⁷ Cette critique formulée dès les premiers travaux de Douglas estimant la fonction de production (de Cobb-Douglas) sur données d'entreprises en coupe, et non plus sur séries agrégées, a été reprise avec force notamment par Phelps-Brown, 1957.

nous pouvions obtenir entre le nombre d'entreprises et la longueur de la période d'études : quatre ans pour les variables en niveaux et trois ans (1995-1997) pour celles en différences premières (taux de croissance logarithmiques). On trouvera en annexe des précisions sur la mesure des différentes variables, et tout particulièrement sur la façon dont les données trimestrielles sur les prix et les taux d'utilisation des capacités ont été annualisées, et aussi dont elles ont été interpolées dans les cas relativement fréquents où ces données étaient renseignées pour trois trimestres seulement d'une même année.⁸

Sur un tel échantillon d'étude, il nous est donc possible de mesurer une valeur ajoutée des entreprises en volume déflatée par les indices de prix de production propre à chacune d'elle ($y_{it} - p_{it}$), et non plus seulement la valeur ajoutée nominale déflatée sectoriellement ($y_{it} - p_{St}$), et espérer obtenir des estimations qui ne seraient pas affectées de biais de ce fait. Nous pouvons aussi introduire la variable de prix individuel relatif ($p_{it} - p_{St}$) comme régresseur aux côtés des variables de travail et de capital dans la fonction de production estimée en pratique sur la valeur ajoutée déflatée sectoriellement. L'intérêt est alors de voir directement si le coefficient estimé de cette variable de prix individuel est statistiquement positif, confirmant ainsi le bien fondé de la déflation par les indices de prix individuels (au-delà de celle par les indices de prix sectoriels). L'intérêt est également de voir si ce coefficient, tout en étant inférieur à 1, est assez proche de 1, permettant de juger dans quelle mesure la variable de prix individuel est affectée par des erreurs de mesure de type aléatoire. L'existence de telles erreurs paraît assez vraisemblable, ne serait-ce que parce qu'il s'agit d'une variable fondée à la base sur les appréciations des entreprises (indépendamment d'une comptabilité précise).⁹ Dans ce cadre d'analyse d'erreurs de mesure aléatoires, la valeur du coefficient de la variable de prix s'interprète en effet comme une estimation du ratio de la variance vraie de la variable de prix individuel à sa variance observée (égale dans ce cadre d'interprétation à la somme de la variance vraie et de la variance des erreurs de mesure).¹⁰

⁸ Les essais que nous avons pu faire en considérant des échantillons de moindre taille en nombre d'entreprises mais une période d'études un peu plus longue conduisent à des résultats peu différents de ceux présentés ici. Il en va de même pour ceux faits sur des échantillons non cylindrés (mais fortement « troués »), comportant plus d'observations sans permettre pour autant d'obtenir des estimations usuelles en différences et en GMMdif. beaucoup plus précises que celles présentées (tableaux 3 et 4).

⁹ La mesure des variations de prix est bien évidemment affectée d'autres erreurs de mesure que celles inhérentes à l'appréciation des chefs d'entreprises. Elle devrait être idéalement corrigée des variations liées aux améliorations de qualité, celles-ci contribuant en principe aux évolutions de volume. C'est là un point important pour les analyses de productivité, notamment dans les secteurs fortement innovants, mais qui est clairement hors du champ de cette étude.

¹⁰ Soulignons que le fait que les indices de prix individuels soient entachés d'erreurs aléatoires de mesure, même importantes, ne constitue pas une cause de biais (mais seulement une source de moindre précision) pour nos

Nous pouvons procéder avec les taux d'utilisation des capacités de production exactement de la même façon qu'avec les indices de prix individuels, en corrigeant simplement la valeur ajoutée des entreprises ou en introduisant la variable correspondante (tc_{it}) comme un autre régresseur dans la fonction de production.¹¹

En résumé, nous nous intéressons à estimer sur l'échantillon considéré, par différentes méthodes et sous diverses contraintes, une fonction de production de Cobb-Douglas qui peut s'écrire sous forme de la régression générale suivante :

$$(y_{it} - p_{St} - l_{it}) = a_t + \alpha(k_{it} - l_{it}) + (\mu - 1)l_{it} + \gamma(p_{it} - p_{St}) + \delta tc_{it} + u_{it}$$

où $(y_{it} - p_{St} - l_{it})$ est le logarithme de la productivité du travail de l'entreprise mesurée par la valeur ajoutée par personne employée déflatée par un indice de prix sectoriel, $(k_{it} - l_{it})$ est le logarithme du capital en début d'année par personne employée, l_{it} le logarithme du travail en nombre de personnes employées, p_{it} le logarithme de l'indice de prix individuel de l'entreprise, tc_{it} le logarithme du taux d'utilisation des capacités de production, et le terme a_t désigne des indicatrices temporelles de progrès technique (et autres mouvements d'ensemble, par exemple conjoncturels). Le coefficient α est le paramètre d'élasticité du capital, μ est celui de rendements d'échelle, et donc $\beta = (\mu - \alpha)$ celui du travail. Nous supposons aussi, comme il est d'usage dans la mise en œuvre des méthodes économétriques de panel, que le terme d'erreurs u_{it} est à erreurs composées, soit $u_{it} = u_i + u_{it}$, où nous faisons ou non l'hypothèse que le terme d'effet individuel u_i peut être corrélé avec le travail et les autres variables explicatives (effets corrélés), et que le terme d'erreur individuel temporel (idiosyncrasique) u_{it} peut être lui-même corrélé avec ces variables (erreurs sur les variables, endogénéité,...).

Nous considérons précisément la fonction de production pour les quatre spécifications emboîtées suivantes :

(1) sans faire intervenir de corrections ($\gamma = \delta = 0$) ;

estimations dans la mesure où ils sont utilisés pour corriger la production, c'est-à-dire la variable dépendante des régressions faites (et non les régresseurs). De ce point de vue il est préférable de corriger directement la production par les prix individuels (et pareillement par les taux d'utilisation), plutôt que d'introduire ces derniers comme variables supplémentaires de contrôle dans les régressions, ce que nous faisons ici à titre de complément d'analyse.

¹¹ Pour ne pas multiplier les estimations nous introduisons la correction d'utilisation des capacités de production après celle des prix individuels. Les observations que l'on peut faire lorsqu'on permute l'ordre de ces deux corrections en introduisant d'abord seule celle des taux d'utilisation ne sont pratiquement pas différentes. Dans les enquêtes de conjoncture, on dispose d'une information sur les marges de production avec et sans embauche, permettant de calculer des taux d'utilisation avec et sans embauche. Nous avons utilisé ici les taux d'utilisation sans embauche préférable a priori dans le cadre de la fonction de production de Cobb-Douglas (à facteurs substituables).

- (2) en déflatant la production par les prix individuels ($\gamma = 1$ et $\delta = 0$) ;
- (3) en corrigeant la production également des taux d'utilisation des capacités ($\gamma = 1$ et $\delta = 1$) ;
- (4) enfin, en introduisant les variables de prix individuels et de taux d'utilisation directement comme variables de contrôle (γ et δ non contraints à 0 ou 1).

Nous estimons ces quatre spécifications par un ensemble de méthodes d'estimation, qui ont suivant les cas l'intérêt de pallier aux risques de biais d'effets corrélés, d'erreurs sur les variables et d'endogénéité. Nous considérons ainsi successivement les principales méthodes usuelles de panels, et les méthodes de moments généralisées, avec les variables en niveaux pour variables instrumentales appliquées à la fonction de production réécrite en différences premières (GMM-dif), ou avec les variables en différences premières pour variables instrumentales appliquées à la fonction de production laissée en niveau (GMM-niv). Nous donnons les estimations obtenues suivant un même format (colonnes 1 à 4) dans les tableaux 3 et 4, après avoir présenté dans les tableaux 1 et 2 quelques statistiques descriptives simples de moyenne, dispersion et corrélation qui les complètent utilement.¹²

2.2. Statistiques descriptives

Les statistiques du tableau 1 montrent notamment que notre échantillon est surtout constitué d'entreprises de taille moyenne, avec un premier et troisième quartiles de 74 et 340 personnes respectivement, et que ces entreprises ont connu des progrès de productivité fortement dispersés, avec un premier quartile de -8,2%, et un troisième de 11,2%. Elles font apparaître que la dispersion des variations de prix individuels, sur les trois années 1995-1997 où nous pouvons les calculer, est notable sans être très forte, avec des quartiles de 0% et 1,7% et des vingtiles de -

Nous avons aussi observé que les taux d'utilisation avec embauche jouaient en fait de façon nettement moins significative que celles sans embauche dans nos régressions, sans améliorer la qualité de l'ajustement.

¹² Nous avons également considéré en détail parallèlement aux spécifications de la fonction de production en termes de valeur ajoutée les spécifications correspondantes en termes de production, en ajoutant comme facteur de production les consommations intermédiaires. Nous ne documentons pas non plus ces résultats pour ne pas alourdir la présentation. Ils ne modifient pas en effet les conclusions auxquelles conduisent ceux obtenus avec la valeur ajoutée. Ils sont très comparables à ces derniers, à ceci près, comme il convient de s'y attendre, que les estimations des élasticités du capital et du travail sont grosso modo diminués en proportion de $(1 - \varphi)$, où φ est l'élasticité des consommations intermédiaires estimée à des valeurs allant de 0.45 à 0.65. Tout au plus peut-on remarquer que les estimations des coefficients γ et δ (lorsqu'ils ne sont contraints à 0 ou à 1) tendent à être elles aussi diminuées mais dans une proportion moindre, ce qui est plutôt satisfaisant. On peut aussi noter que les valeurs estimées pour φ , l'élasticité des consommations intermédiaires, sont très sensiblement moins élevées (par exemple 0.45 au lieu de 0.55 en différences premières avec un écart-type de 0.02), quand on corrige la production du taux d'utilisation des

5,2% et 4,5%. Cette dispersion est cependant nettement plus élevée que celle des prix sectoriels comme on pouvait s'y attendre. Elle n'est sans doute pas plus forte, car elle concerne une période où l'inflation moyenne des prix a été très faible : 0,3% sur la base des prix individuels et 0,4% sur celle des prix sectoriels. On peut faire une remarque analogue sur les variations des taux d'utilisation des capacités dont la dispersion reste assez modérée, avec des quartiles de -2,7% et 3,4% et des vingtiles de -9,5% et 9,5%, sur une période où la croissance a été peu chahutée.

TABLEAUX 1 et 2

Les corrélations documentées dans le tableau 2 montrent notamment que les variations de prix individuels telles que nous les mesurons dans les enquêtes de conjoncture sont corrélées positivement et de façon significative avec celles de valeur ajoutée, et cela de manière significativement plus marquée que pour les variations de prix sectoriels (0,13*** au lieu de 0,09***). Elles sont également corrélées significativement et positivement avec les variations d'effectifs (0,10***). Ces corrélations ont le signe attendu, et quoique modérées confirment que nos mesures apportent une véritable information sur les variations effectives des prix de production, au delà des erreurs de mesure très vraisemblables (dont on s'attend à ce qu'elles jouent dans le sens d'une atténuation des corrélations).¹³ On note cependant que la corrélation avec les variations de capital est presque nulle (0,01), et que celles des variations de valeur ajoutée ou encore des variations d'effectifs avec les variations de capital apparaissent aussi assez faibles (0,05* pour l'une et l'autre). Cette observation va dans le sens de fortes erreurs de mesure sur les variations de capital, ce qui n'est pas pour étonner.¹⁴ Des remarques analogues peuvent être faites pour ce qui est des corrélations avec les variations des taux d'utilisation des capacités de production ; on note en particulier que leur corrélation avec les variations de capital est comme on doit s'y attendre significativement négative (-0,09***), ce qui est assez rassurant à la fois sur

capacités ($\delta = 1$), ce qui n'est pas étonnant dans la mesure où production, capacités de production et consommations intermédiaires des entreprises sont directement variables en fonction de leur conjoncture propre.

¹³ L'étude que nous avons pu faire en confrontant les informations de prix tirées des enquêtes de conjoncture et celles construites à partir des données mensuelles ou trimestrielles à un niveau très fin de la nomenclature des produits collectés par l'INSEE pour établir les indices des prix de ventes industrielles (IPVI) confirme ce jugement. Elle montre en effet une assez bonne cohérence entre ces deux sources d'information, bien qu'elles soient évidemment très différentes tant dans leur conception que dans leur réalisation (cf. chapitre 1, Desplatz, 2000).

¹⁴ Lorsque le capital est mesuré en fin d'année plutôt qu'en début d'année (ce qui suppose un délai moyen de mise en fonctionnement normal des équipements de 6 mois qui paraît raisonnable), on trouve des corrélations assez nettement plus élevées entre variations de capital et celles de valeur ajoutée et d'effectifs (0,09 et 0,16) mais inchangée avec celles des prix de production.

notre mesure des taux d'utilisation et sur celle du capital.¹⁵ On observe aussi avec une certaine satisfaction que la corrélation entre les variations des prix et celles des taux d'utilisation est significativement positive (0,07**).

2.3. Estimations de panel usuelles

Les résultats présentés dans le tableau 3 sont dans l'ensemble assez représentatifs de ce que l'on obtient généralement sur données de panel lorsqu'on estime une fonction de production suivant les différentes méthodes usuelles d'économétrie des panels. Ainsi les estimations en Total (sur les variables en niveaux x_{it}) sont a priori vraisemblables avec une élasticité du capital α de l'ordre de 0,25 et des rendements d'échelle μ quasi-constants. De façon plus étonnante, mais satisfaisante par rapport aux résultats habituels qui le sont moins, les estimations en différences longues (sur $x_{i(97)} - x_{i(94)}$), qui ont l'avantage d'éviter les biais d'effets corrélés (entre u_i et x_{it}) tout en étant moins sensible aux erreurs (aléatoires) sur les variables que les différences premières, apparaissent également fort raisonnables ($\alpha = 0,20$ et $\mu = 1$).¹⁶ Les estimations obtenues en Intra (sur les variables en déviation à leurs moyennes individuelles $x_{it} - x_i$), et surtout en différences premières (sur $x_{it} - x_{i(t-1)}$) sont quant à elles comparables à ce que l'on trouve habituellement, avec une élasticité du capital faible (et même statistiquement non significative en différences premières) et des rendements d'échelle fortement décroissants.¹⁷

Au-delà de cette configuration d'estimations assez bien établie, les résultats obtenus apportent une information certes nouvelle et au total rassurante, néanmoins essentiellement négative. Ils montrent que la prise en compte de la dispersion des prix de production ainsi que celle des taux d'utilisation des capacités jouent de façon extrêmement modeste sur les estimations de l'élasticité

¹⁵ Il n'est pas sans intérêt de constater que la corrélation avec les variations de capital mesuré en fin d'année est quant à elle nulle, ce qui tend à conforter le choix fait de préférence d'une mesure du capital en début d'année dans la fonction de production.

¹⁶ La période d'étude est trop courte pour que l'existence d'erreurs sur les variables, plus fortement exacerbées en différences premières qu'en différences longues, puisse entièrement rendre compte des différences entre les deux types d'estimations.

¹⁷ Pour des estimations présentant des disparités grosso modo comparables, obtenues pour des panels d'entreprises américaines, françaises et japonaises portant sur une période beaucoup plus longue (13 années), voir Mairesse, 1988 et 1990, ou Griliches-Mairesse, 1998, où elles sont rappelées.

du capital et des rendements d'échelle, et qu'elles n'atténuent pas, sinon de façon négligeable, les disparités observées entre les différents types d'estimation.^{18, 19}

TABLEAU 3

Si l'on considère cependant les estimations des coefficients des variables de prix et de taux d'utilisation γ et δ dans la spécification générale (en colonne 4), ceux-ci sont bien significativement positifs et aussi (significativement) inférieurs à 1, quelque soit la méthode d'estimation.²⁰ C'est là l'indication tout à la fois que ces variables apportent effectivement une information réelle, même si elles sont fortement affectées par des erreurs de mesure importantes, ce qui est bien sûr nullement étonnant.²¹ Cette information reste toutefois largement « orthogonale » aux variations des facteurs de production (capital et travail). L'explication la plus vraisemblable de cet apparent paradoxe est celle que nous avons déjà suggérée (section 1). Les corrélations entre les variations de prix et de taux d'utilisation et les variations de facteurs, même si elles sont statistiquement significatives (cf. tableau 2), restent négligeables relativement aux autres sources de corrélation entre ces facteurs et les perturbations affectant la fonction de production. Elles ne sont pas suffisantes pour modifier sensiblement les estimations usuelles sur panel de la fonction de production, et les disparités observées entre estimations de types « coupe » (Total et Inter à peine différentes) et « temporel » (Intra et différences premières et longues), et entre ces dernières entre elles doivent trouver ailleurs leur origine. Ce n'est donc qu'à la condition

¹⁸ Tout au plus peut-on remarquer en comparant les estimations des trois premières colonnes du tableau 3 que les rendements d'échelle cessent d'apparaître très légèrement croissants en Total, et un peu moins décroissants en Intra et différences premières. Cela va dans le sens attendu mais de façon tout à fait marginale.

¹⁹ On peut noter par ailleurs que les estimations, non présentées ici, dans lesquelles on utilise la valeur ajoutée comme variable de production y_{it} sans même la déflater par un indice de prix sectoriel, ne sont pas sensiblement différentes non plus de celles où elle est déflatée par un tel indice. La raison pour laquelle il reste préférable d'utiliser une valeur ajoutée déflatée par un indice de prix sectoriel concerne l'estimation des termes a_t de progrès technique ou de productivité totale des facteurs, dont on souhaite qu'ils ne soient pas affectés par une inflation sectorielle (et globale) des prix.

²⁰ Contrairement aux autres variables, la variabilité Intra (de $x_{it} - x_i$) des deux variables de prix et de taux d'utilisation n'est pas faible par rapport à la variabilité Inter (de x_i), ce qui explique que les estimations en Total de leurs coefficients apparaissent statistiquement significatives et non déraisonnables.

²¹ Si on admet que le seul problème affectant l'estimation du coefficient de la variable de prix est celui d'erreurs de mesure aléatoires, on peut inférer d'un coefficient de 0,55 en différences premières que la variance vraie des variations de prix est de moitié environ de celle observée (et égale à celle des erreurs). On s'attendrait toutefois à ce que ce coefficient soit plus élevé en différences longues (sur trois ans), ce qui n'est pas le cas avec un coefficient estimé de 0,35 statistiquement peu différent, néanmoins plus faible. Une autre raison de biais vers le bas du coefficient de la variable de prix individuel, allant dans le même sens que celle d'erreurs de mesure aléatoires, pourrait être une corrélation entre une hausse (ou une baisse) de prix et des chocs négatifs (ou positifs) de

de pouvoir tenir compte, au moins en partie, des autres raisons susceptibles de rendre compte effectivement de ces disparités que nous pouvions a priori garder un espoir (minime) de mettre en évidence un impact notable de la dispersion des prix et des taux d'utilisation des capacités. C'est ce que nous avons tenté en mettant en œuvre la méthode des moments généralisés.

2.4. Estimations par la méthode des moments généralisés

Il est souvent (trop rapidement) admis en économétrie des panels que les disparités entre estimations de type coupe et temporel s'expliquent surtout par une hétérogénéité prenant la forme d'effets individuels corrélés, et que les secondes doivent être acceptées et les premières remises en cause, alors même qu'elles peuvent paraître beaucoup plus vraisemblables, comme c'est ici le cas pour les estimations des rendements d'échelle et des élasticités du capital et du travail dans la fonction de production. En outre, comme dans notre cas aussi, les estimations de type temporel peuvent elles mêmes différer entre elles de façon significative, indiquant la présence de corrélations entre les variables explicatives observées et les termes d'erreurs u_{it} habituels (variables entre individus et temporellement). Les trois grandes raisons dont on considère généralement qu'elles peuvent être à l'origine de telles corrélations et des biais d'estimation correspondants sont l'endogénéité des variables explicatives, l'existence d'erreurs de mesure sur ces variables et l'omission de variables explicatives corrélées avec celles présentes (et non prises en compte par les effets individuels pour les variables omises constantes, ou presque, sur la période d'étude). On peut dans les deux premiers cas espérer trouver une solution dans les méthodes de moments généralisés (GMM). Ces méthodes, appliquées notamment aux données de panel, peuvent l'être suivant un assez grand nombre de modalités distinctes. Dans le tableau 4, nous présentons les estimations obtenues suivant cinq groupes de modalités qui paraissent a priori utiles d'explorer, sans comme il est l'usage se contenter de montrer les seules estimations jugées les plus satisfaisantes, en particulier parce qu'elles passent les tests de spécification du χ^2 (ou tests de suridentification de Sargan).

Les deux premiers groupes d'estimation correspondent à l'instrumentation de la fonction de production en différences premières (à l'année t) suivant les deux modalités notées GMMdif-CEEV et GMMdif-CEWS : c'est à dire pour la première à une instrumentation par les variables

productivité. Dans le cas de la variable de taux d'utilisation, l'hypothèse d'erreurs de mesure paraît suffisante pour expliquer les coefficients estimés de 0.25 et 0.65 en différences premières et longues respectivement.

explicatives en niveaux futures (à t+1 et au-delà) et passées (à t-2 et avant), et pour la seconde par les seules variables en niveaux passées. Ces modalités permettent de tenir compte pour l'une et l'autre des risques de biais d'effets corrélés et d'erreurs (non auto-corrélés) sur les variables explicatives, et pour la seconde également des risques d'endogénéité faible de ces variables (ou endogénéité sans rétroaction).²² Les estimations ainsi obtenues sont très imprécises, tout particulièrement celles de GMMdif-CEWS. Les écart-types sont en fait tellement importants que ces estimations ne sont pas incompatibles avec des valeurs jugées raisonnables comme avec des valeurs jugées déraisonnables. La contrepartie de la robustesse souhaitée aux erreurs de spécification est la très forte imprécision des estimations et leur grande fragilité, au moins sur des échantillons de la taille du nôtre (et des périodes aussi courtes).²³ En tout état de cause, ces estimations, même essayant de faire abstraction de leur imprécision, ne font pas apparaître de différences sensibles entre celles faites avec et sans déflation de la production par les prix individuels et avec et sans correction par les taux d'utilisation des capacités.

TABLEAU 4

Les trois derniers groupes d'estimation correspondent à l'instrumentation de la fonction de production en niveaux (à l'année t) suivant les trois modalités notées GMMniv-CE, GMMniv-CEEV et GMMniv-CEWS : c'est à dire pour la première à une instrumentation par les variables en différences premières futures, présentes et passées, pour la seconde par les seules différences premières futures (à t+2 et au-delà) et passées (à t-1 et avant), et pour la troisième par les seules différences premières passées. Ces modalités permettent de tenir compte pour toutes trois des risques de biais d'effets corrélés, pour les deux dernières des risques également d'erreurs sur les variables explicatives, et enfin pour la dernière de ceux aussi d'endogénéité faible de ces

²² Sur les estimations de type GMMdif. voir l'article originel de Arellano et Bond, 1991, ainsi que Sevestre, 2002 et Arellano, 2003. Nous ne présentons pas les résultats de GMMdif-CE, où les variables en différences premières sont instrumentées par les variables en niveaux futures, présentes et passées, ces estimations n'étant autres que celles obtenues en Intra déjà donnés dans le tableau 3 (avec des écart-types robustes à l'hétéroscédasticité de l'ordre de 40% plus élevés).

²³ Les « mauvaises » performances dans de tels cas des estimateurs de type GMMdif. ne sont pas nouvelles ni très étonnantes. Par exemple, les simulations réalisées par Crépon et Mairesse, 1995, montrent que, sur un échantillon de taille comparable (T=3 et N=400) et dans une configuration choisie a priori comme proche de celle de l'estimation d'une fonction de production, les écart-types sont, par rapport à ce qu'ils seraient si on connaissait la vraie spécification, de l'ordre de quatorze fois plus élevés si la vraie spécification est « seulement » de type effets corrélés et erreurs de mesure, ou encore de l'ordre de cinquante cinq fois plus élevés (!) si la vraie spécification est « seulement » de type effets corrélés, alors qu'on suppose une spécification plus robuste combinant effets corrélés, erreurs de mesure et endogénéité faible.

variables.²⁴ Les estimations obtenues gagnent en considérablement en précision par rapport aux précédentes.²⁵ Les estimations de l'élasticité du capital et les rendements d'échelle sont satisfaisantes (et assez proches de celles obtenues en coupe) pour ce qui au moins des modalités d'instrumentation CE et CEEV, les estimations de l'élasticité du capital paraissant étonnamment élevées pour la troisième modalité CEWS. On observe en revanche que les tests de suridentification de Sargan sont fortement rejetés pour les deux premières modalités, mais acceptés (au seuil de confiance 5%) pour la troisième que nous devrions donc au regard de ce critère préférer. Pour ce qui est l'objet premier de notre analyse, le résultat essentiel est cependant, comme précédemment, l'absence de différence significative entre les estimations faites avec les mesures de la production corrigées ou non par les prix individuels et les taux d'utilisation des capacités.²⁶ Le très faible espoir qu'il en soit autrement en considérant les estimations robustes aux différents autres risques majeurs de biais obtenus par les méthodes moments généralisés se révèle illusoire.

Conclusion

Dans leur survol sur l'identification et de l'estimation des fonctions de production et l'apport de l'économétrie des panels, Griliches et Mairesse (1998) insistent sur la nécessité d'approfondir notre compréhension des comportements d'entreprises et de leurs sources d'hétérogénéité, mais aussi sur les besoins d'améliorer et enrichir la mesure de la production et de ses facteurs, et notamment l'utilité qu'il y aurait à disposer d'informations sur les prix de production des entreprises. Dans cette étude, nous cherchons à savoir si les estimations des fonctions de production sur données de panel peuvent être effectivement améliorées quand on mobilise les

²⁴ Sur les estimations de type GMMniv. voir l'article originel de Arellano et Bover, 1995, ainsi que Sevestre, 2002 et Arellano, 2003. Celles-ci par rapport aux estimations de type GMMdif. font l'hypothèse supplémentaire de stationnarité de la covariance entre effets corrélés et variables explicatives. Notons que aurions pu faire également des estimations de type GMMsys. qui s'appuient sur l'ensemble des conditions d'orthogonalité de GMMdif. et GMMniv. ; cependant dans la mesure même où GMdif. et GMMniv. conduisent à des estimations différentes, cela ne paraît pas justifié. La même raison explique que nous ne présentons pas les estimations usuelles de type MCQG combinant les estimations Total et Intra, dont il apparaît clairement qu'elles sont statistiquement très différentes.

²⁵ Ce résultat se comprend d'autant mieux que les variables explicatives sont plus fortement auto-corrélées. En effet, en ce cas, les variables en différences premières passées sont des instruments d'autant plus puissants pour les niveaux présents. A l'inverse, et c'est une raison de la faible précision des estimations en GMMdif., les variables en niveaux passés sont des instruments d'autant plus faibles pour les différences premières présentes. A la limite les variables en niveaux, sous l'hypothèse où elles satisfont à un processus de marche aléatoire, sont sans corrélation avec les différences premières passées.

informations individuelles relatives aux prix de production, ainsi que sur les taux d'utilisation des capacités, qui sont demandées dans les enquêtes de conjoncture de l'INSEE auprès des entreprises, et donc à l'inverse à savoir si ces estimations sont biaisées en l'absence de telles informations, ce qui est la situation commune.

Contrairement à ce que l'on pouvait penser sur la base d'une analyse a priori, nous devons conclure que l'amélioration des mesures de production et de productivité permise par la disponibilité d'informations sur les prix individuels des entreprises ne joue pas sensiblement sur les estimations ; il en est ainsi que l'on ait simplement recours aux méthodes d'estimation usuelles de l'économétrie des panels ou à celles plus récentes des moments généralisés ; et il en est de même pour l'utilisation des informations sur les taux d'utilisation des capacités. Ces informations ne contribuent pas notamment à l'explication des disparités importantes entre estimations de type coupe (priviliégiant la dimension individuelle des données) et de type série (fondées sur leur dimension temporelle).

Ces résultats sont rassurants puisqu'ils confortent la pratique habituelle qui consiste à déflater simplement les mesures comptables de production en valeur par des indices de prix sectoriels. Ils sont néanmoins décevants par rapport aux espoirs que l'on pouvait avoir d'améliorer les estimations des élasticités du capital et des rendements d'échelle peu vraisemblables souvent trouvées, dès lors que l'on cherche à s'assurer contre les risques de biais d'hétérogénéité et d'endogénéité. Ils sont surtout déconcertants car ils semblent impliquer qu'il ne serait pas possible de distinguer en pratique fonction de production et fonction de revenu sur données microéconomiques. Il est probable, comme nous l'avons suggéré, que l'erreur de spécification qui peut résulter d'une absence de prise en compte de la dispersion des variations des prix individuel et des taux d'utilisation des capacités est faible relativement aux autres erreurs de spécification, notamment celles liées aux erreurs sur les variables. En pratique, les méthodes d'estimation des moments généralisées ne réussissent à pallier ces erreurs de spécification qu'assez imparfaitement, et font courir le risque en traitant certaines d'entre elles d'aggraver fortement les conséquences des autres.

Il n'en reste pas moins que des efforts importants doivent (devraient) être poursuivis en matière d'informations sur les prix de production à des niveaux fins (des transactions

²⁶ Ce résultat négatif ne doit paraître vraiment surprenant dans la mesure où on peut assimiler les estimations de type GMMniv. s'apparentant avant tout, comme les estimations en Total, aux estimations en coupe, pour lesquelles seul fait vraiment sens l'estimation d'une fonction de production en valeur (ou fonction de revenu).

élémentaires, des produits, des établissements, des entreprises,...) de sorte qu'elles puissent être disponibles dans les études microéconomiques. Les problèmes posés par la définition et la mesure des prix, par leur comparaison entre « mêmes » produits de caractéristiques et qualité différentes, sont particulièrement complexes. Le champ des études sur les comportements d'entreprises et le fonctionnement des marchés qui serait ouvert grâce à la disponibilité de données individuelles de prix est évidemment considérable. Pour ce qui est de l'objet de notre étude sur l'amélioration des mesures de productivité et des estimations des fonctions de production sur données de panel, ses résultats restent encore exploratoires ; ils devront naturellement être confirmés sur des échantillons différents, de plus grande taille, sur une période plus longue, avec des mesures de prix fondées sur des sources d'information différentes, et de préférence aussi pour des périodes et des pays ou des secteurs, où l'inflation aurait été notable. Sur de tels échantillons, et dans un cadre d'analyse plus large, ils pourront également être approfondis et mieux compris, du moins peut-on l'espérer.

REFERENCES

- Abbott T. A., 1991, "Producer Price Dispersion, Real Output and the Analysis of Production". *The Journal of Productivity Analysis*, 2, 179-195.
- Arellano, M., 2003, "*Panel Data Econometrics*". Oxford University Press, Oxford.
- Arellano, M. and S. Bond, 1991, "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations". *Review of Economic Studies*, 58, 277-97.
- Arellano, M. and Bover, O., 1995, "Another look at the instrumental-variable estimation of error-components models". *Journal of Econometrics*, 68, 29-52.
- Crépon B. and J. Mairesse, 1996, "The Chamberlain Approach: An Overview and some Simulation Experiments". In L. Matyas and P. Sevestre (eds.), *The Econometrics of Panel Data: A Handbook of the Theory with Applications* (2nd edition, chapter 14), Kluwer Academic Publishers, 323-391.
- Crépon B., R. Desplatz and J. Mairesse, 1999, "Estimating Price-Cost Margins, Scale Economies and Workers' Bargaining Power at the Firm Level". Working Paper CREST, G9917.
- Dagenais M., 1992, "Pièges et limitations de l'analyse micro-économétrique", *Economie et Prévision*, 102-103, 1-9.
- Desplatz R., 2000, "Hétérogénéité des prix et des salaires, pouvoir de marché et emploi : quatre analyses économétriques sur données individuelles d'entreprises", Thèse, Université de Paris 1- Panthéon-Sorbonne.
- Griliches Z. and J. Mairesse, 1998, "Production Function: The Search for Identification". In S. Ström (ed.), *The Ragnar Frisch Centennial Symposium*, Cambridge University Press, Cambridge.
- Klette J. and Z. Griliches, 1996, "The Inconsistency of Common Scale Estimators when Output Prices are Unobserved and Endogenous". *Journal of Applied Econometrics*, 11, 343-361.
- Mairesse J., 1990, "Time-Series and Cross-Sectional Estimates on Panel Data: Why Are they Different and why Should they Be Equal?" In J. Hartog and al. (eds.), *Panel Data and Labor Market Studies*, Amsterdam : North-Holland Publishing Co., 81-95.
- Mairesse J., 1988, "Les lois de la production ne sont plus ce qu'elles étaient : une introduction à l'économétrie des panels". *Revue Economique*, n° spécial sur Economie, Sociologie, 39(1), 225-271.
- Marschak, J. and W. Andrews, 1944, "Random Simultaneous Equations and the Theory of Production". *Econometrica*, 46, 49-85.
- Phelps-Brown, E.H., 1957, "The Meaning of the Fitted Cobb-Douglas Function". *The Quarterly Journal of Economics*, 71, 546-560.
- Sevestre P., 2002, "*Econométrie des données de panels*". Dunod, Paris.

ANNEXE

Sources de données utilisées

Pour construire notre échantillon nous avons apparié principalement deux sources de données de l'INSEE. La première, les enquêtes *trimestrielles* de conjoncture, comprend les informations sur les prix et les taux d'utilisation des facteurs au niveau de l'entreprise. La seconde, la base SUSE (Système Unifié de Statistiques d'Entreprises), fondée sur les informations principalement comptables des enquêtes annuelles d'entreprises et des déclarations fiscales BIC (« Bénéfices Industriels et Commerciaux »), permet de mesurer la production et la valeur ajoutée, le capital et l'emploi des entreprises.

Après appariement de ces deux bases de données, nous disposons d'un échantillon d'un millier d'entreprises sur la période 1991-1997, où les observations relatives aux prix et aux taux d'utilisation de l'enquête de conjoncture sont souvent manquantes (échantillon « troué »). Après cylindrage de l'échantillon sur les années récentes 1994-1997, et après élimination des d'observations extrêmes (« outliers ») sur les variables principales de l'étude, notamment en taux de croissance de la productivité et du capital, qui sont vraisemblablement erronées, nous obtenons le fichier de 468 entreprises, qui a été finalement retenu.

1. Enquêtes trimestrielles de conjoncture

Les enquêtes *trimestrielles* de conjoncture de l'INSEE auprès des entreprises incluent des questions sur l'évolution de leurs prix de vente mais également sur ses marges de production, à partir desquelles nous avons pu calculer au niveau même des entreprises des indices de prix annuels et des taux annuels d'utilisation des capacités de production, et donc des volumes de valeur ajoutée déflatés par ces indices de prix et corrigés par ces taux d'utilisation.²⁷

Si on considère qu'un quart de la valeur ajoutée annuelle en valeur est réalisée chaque trimestre et plus précisément en milieu de trimestre, on peut calculer l'indice de prix annuel de l'entreprise i à l'année t , en base 100 au 1er Janvier de cette année, ou plus précisément son inverse IP^{-1}_{it} , par la formule :

$$IP^{-1}_{it} = \frac{1}{4} \left(\frac{1}{(1+VP_{i,t,1}/2)} + \frac{1}{(1+VP_{i,t,1})(1+VP_{i,t,2}/2)} + \frac{1}{(1+VP_{i,t,1})(1+VP_{i,t,2})(1+VP_{i,t,3}/2)} + \frac{1}{(1+VP_{i,t,1})(1+VP_{i,t,2})(1+VP_{i,t,3})(1+VP_{i,t,4}/2)} \right)$$

où $VP_{i,t,q}$ est la variation de prix trimestrielle de l'entreprise i au cours du trimestre q ($=1,2,3$ ou 4) de l'année t .

²⁷ Notons que dans les cas assez peu fréquents où les entreprises déclarent aux enquêtes de conjoncture des réponses séparées pour leurs principales branches d'activités (ce qui est assez rare), nous avons au préalable calculé leurs

La formule ci-dessus de calcul de l'indice annuel de prix nécessite en principe la connaissance des variations de prix trimestrielles pour les quatre trimestres de l'année considérée pour l'entreprise considérée. En raison de réponses assez souvent manquantes aux enquêtes de conjoncture (que ce soit pour un questionnaire entier à un trimestre ou seulement pour la question sur les prix), et afin d'éviter de perdre trop d'entreprises dans la construction de notre échantillon, nous avons aussi calculé les variations de prix annuelles à partir de la connaissance des variations de prix trimestrielles pour trois trimestres, en remplaçant la variation de prix trimestrielle manquante sur l'année par la moyenne trimestrielle calculée à partir des trois variations de prix trimestrielles connues. Si on suppose que la variation de prix manquante est celle du quatrième trimestre, nous avons ainsi appliqué la formule :

$$1 + VP_{i,t,4} = [(1 + VP_{i,t,1})(1 + VP_{i,t,2})(1 + VP_{i,t,3})]^{1/3}$$

Deux types de taux d'utilisation des facteurs de production ou capacités de production peuvent être calculés à partir des réponses aux questions suivantes des enquêtes trimestrielles de conjoncture: « Quel pourrait-être l'accroissement de votre production avec le matériel dont vous disposez et sans que vous ayez besoin d'embaucher du personnel supplémentaire ? » et « Pourriez-vous produire davantage encore en embauchant du personnel supplémentaire ? Si oui, de combien pourriez-vous augmenter au total votre production avec le matériel dont vous disposez ? ». Les réponses à la première question renseignent sur les *marges de production trimestrielles sans embauche* supplémentaire (accroissement de production possible avec les moyens actuels de l'entreprise) alors que les réponses à la seconde question renseignent sur les *marges de production trimestrielles avec embauche* (accroissement de production possible avec les équipements existants mais avec du personnel supplémentaire). Les *marges annuelles* avec et sans embauche, soit M_a et M_s , sont obtenues en principe en faisant la moyenne simple des marges trimestrielles. Pour ne pas diminuer la taille de notre échantillon, nous avons été cependant jusqu'à les calculer à partir d'une seule réponse trimestrielle. Les taux d'utilisation annuels avec et sans embauche TU_a et TU_s s'expriment alors directement en fonction des marges (exprimées en pourcentage) par les formules:

$$TU_s = 1/(1 + M_s / 100) \quad \text{et} \quad TU_a = 1/(1 + M_a / 100).$$

Dans l'étude nous avons retenu de préférence le taux d'utilisation sans embauche, le taux avec embauche n'apportant apparemment pas d'information supplémentaire.²⁸

variations de prix trimestrielles comme les moyennes pondérées des variations de prix déclarées au niveau des branches.

²⁸ En principe, on pourrait interpréter le taux d'utilisation sans embauche comme un taux d'utilisation du travail ou de rétention de la main d'oeuvre dans l'entreprise (labour hoarding) : $TU_L = TU_s$. On peut effectivement penser que la production ne peut s'accroître que par une hausse de l'effort productif des salariés, à stock de capital et effectifs de travailleurs fixés et à durées d'utilisation des équipements et heures de travail également constantes. On peut aussi évaluer le taux d'utilisation du capital par le rapport des taux avec et sans embauche : $TU_K = TU_a / TU_s$.

2. Données annuelles comptables

Les données utilisées pour mesurer la production, la valeur ajoutée, le capital et l'emploi sont issues de la base SUSE de l'INSEE. La production est égale aux ventes corrigées des variations de stocks ; la valeur ajoutée (en valeur) est évaluée au coût des facteurs. L'emploi est mesuré par l'effectif moyen du personnel dans l'entreprise. La valeur du stock de capital est mesurée par les immobilisations corporelles brutes (terrains, construction, matériel, autres immobilisations corporelles, immobilisations en cours et avances et acomptes), soit IMB_t en début d'année t . Sachant que cette grandeur est évaluée en général à son coût historique, on en déduit le stock en volume K_t en début d'année t , à partir d'une estimation de son âge moyen AGC_t et de l'indice de prix $PINV(t)$ des investissements de l'industrie manufacturière dans les comptes nationaux, aux prix d'une année de référence (soit 1980) suivant l'approximation suivante : $K_t = IMB_t / PINV(t-AGC_t)$. Cette approximation suppose que les immobilisations ont été acquises globalement à l'année $(t-AGC_t)$.²⁹ L'âge moyen du capital est lui-même estimé approximativement à partir du rapport des immobilisations amorties IMA_t et des immobilisations brutes IMB_t en supposant que la durée de vie économique moyenne des immobilisations est de 16 années et leur durée de vie fiscale moyenne de 8 années:

$$AGC_t = 8 * (IMA_t / IMB_t), \text{ si } (IMA_t / IMB_t) \leq 0.5, \text{ ou } AGC_t = 16 * [(IMA_t / IMB_t) - 0.25] \text{ si } (IMA_t / IMB_t) > 0.5.^{30}$$

Cette méthode a l'avantage d'utiliser l'information sur les immobilisations brutes au bilan (et sur la proportion de ces immobilisations qui sont amorties fiscalement) et d'éviter d'utiliser la méthode chronologique dite « d'inventaire permanent », qui demande de disposer de séries d'investissements longues et de faire des hypothèses de déclassement (en général à taux constant).

²⁹ Remarquons que $(t-AGC_t)$ n'étant pas une valeur entière, l'indice de prix $PINV$, supposé mesuré en milieu d'année t , est simplement interpolé entre milieux d'années consécutives. On a pu vérifier de façon assez générale dans d'autres études que les estimations de fonction de production sont peu sensibles aux modalités précises d'une telle correction des immobilisations brutes pour comptabilisation aux coûts historiques.

³⁰ AG serait l'âge du capital si la durée de vie fiscale était égale à la durée de vie économique (16 ans), sous les deux hypothèses de durée de vie fixe et d'amortissement linéaire. AGC est corrigé de façon approximative pour tenir compte du fait que la durée de vie fiscale est plus courte (8ans) et que le capital peut donc être amorti au bilan à 100% (après 8 ans), alors qu'il est toujours productif (jusqu'à 16 ans).

Table 1: Statistiques descriptives : Moyennes et indicateurs de dispersion.

Variables:		Moyenne	Ecart- Type	Q1	Q3	P5	P95
Effectif moyen (en milliers)	L	275	365	74	340	33	869
<u>Taux de croissance moyen en %</u>							
Effectif	DLL	0.1	8.6	-3.3	3.4	-11.9	12.1
Productivité (non déflatée)	DLVAL	1.7	18.5	-8.2	11.2	-27.2	33.4
Productivité déflatée par les prix sectoriels	DLVADSL	1.3	18.5	-9.0	11.0	-27.9	32.6
Productivité déflatée par les prix individuels	DLVADEL	1.4	18.5	-8.7	10.9	-28.3	33.2
Productivité déflatée par les prix individuels et corrigée du taux d'utilisation	DLVADECL	1.3	19.3	-9.5	11.8	-29.7	35.9
Capital par personne en début d'année	DLKIL	1.5	11.7	-4.7	6.8	-12.8	21.4
Prix sectoriel	DLPPIS	0.4	2.6	-0.2	1.2	-3.8	2.2
Prix individuel	DLPPIE	0.3	3.5	0.0	1.7	-5.2	4.5
Ecart entre prix individuel et sectoriel	DLPPIES	-0.1	3.7	-1.2	1.4	-6.0	5.1
Taux d'utilisation des capacités de production	DLTUS	0.2	6.4	-2.7	3.4	-9.5	9.5

Echantillon cylindré 1995-1997, 468 entreprises.

Q1 et Q3 désignent respectivement le premier et troisième quartiles, et P5 et 95 le premier et dix-neuvième vingtiles.

Les taux de croissance sont calculés comme les différences premières des logarithmes.

La « productivité » est mesurée par la valeur ajoutée par personne ; le « capital » par les immobilisations corporelles au bilan avec correction de prix (cf. annexe). Les taux de croissance sont calculés comme les différences premières des logarithmes.

Table 2: Statistiques descriptives : Corrélations entre variables principales en différences premières.

Corrélations	DLVA	DLL	DLK1	DLPPIS	DLPIIE	DLPIES	DLTUS
Valeur ajoutée (en valeur) DLVA	1						
Effectif DLL	0.22***	1					
Capital (en début d'année) DLK1	0.05*	0.05*	1				
Prix sectoriel DLPPIS	0.09***	0.08***	-0.05**	1			
Prix individuel DLPIIE	0.13***	0.10***	0.01	0.28***	1		
Ecart de prix DLPIES	0.06**	0.03	0.05*	-0.43***	0.74***	1	
Taux d'utilisation DLTUS	0.09***	0.04	-0.09***	0.05*	0.07**	0.03	1
Ecart- type (en %)	18.3	8.6	8.3	2.5	3.5	3.7	6.4

Echantillon cylindré 1995-1997, 468 entreprises. Les astérisques ***, ** et * indiquent respectivement que les corrélations sont statistiquement significatives au seuil de confiance de 1%, 5% et 10%. Définitions et notations des variables : voir tableau 1.

Tableau 3 : Estimations de panel usuelles

	Avec prix sectoriels	Avec prix individuels	Avec prix et taux d'utilisation individuels, en imposant des coefficients unitaires ou non	
	LVADSL	LVADEL	LVADECL	LVADSL
Total				
LK1L	0.28 (0.01)	0.26 (0.01)	0.26 (0.01)	0.26 (0.01)
LL	0.04 (0.01)	0.03 (0.01)	0.01 (0.01)	0.03 (0.01)
LPPIES	-	-	-	0.23 (0.10)
LTUS	-	-	-	0.25 (0.07)
R ² (m.s.e.)	0.998 (0.274)	0.998 (0.278)	0.998 (0.284)	0.998 (0.273)
Intra				
LK1L	0.15 (0.04)	0.13 (0.04)	0.18 (0.04)	0.16 (0.04)
LL	-0.21 (0.05)	-0.25 (0.05)	-0.22 (0.05)	-0.22 (0.05)
LPPIES	-	-	-	0.45 (0.09)
LTUS	-	-	-	0.32 (0.07)
R ² (m.s.e.)	0.068 (0.119)	0.072 (0.119)	0.086 (0.122)	0.094 (0.117)
Différences premières				
LK1L	0.09 (0.06)	0.07 (0.06)	0.13 (0.06)	0.10 (0.06)
LL	-0.46 (0.08)	-0.50 (0.08)	-0.45 (0.08)	-0.47 (0.08)
LPPIES	-	-	-	0.55 (0.13)
LTUS	-	-	-	0.25 (0.08)
R ² (mse.)	0.078 (0.179)	0.081 (0.178)	0.09 (0.184)	0.097 (0.177)
Différences Longues				
LK1L	0.21 (0.06)	0.19 (0.06)	0.23 (0.06)	0.23 (0.06)
LL	0.02 (0.08)	-0.02 (0.08)	-0.01 (0.08)	0.02 (0.08)
LPPIES	-	-	-	0.34 (0.13)
LTUS	-	-	-	0.63 (0.14)
R ² (mse.)	0.035 (0.073)	0.033 (0.075)	0.047 (0.073)	0.088 (0.071)

Les écart-types sont indiqués entre parenthèses. Des indicatrices sectorielles et temporelles ont été ajoutées dans les régressions « Total », et des indicatrices temporelles dans les régressions « Intra », « Différences premières », et « Différences longues ».

Echantillon cylindré, 468 entreprises : 1994-1997 en Total, et 1995-1997 en Intra, Différences premières et Différences longues. Les variables sont toutes en logarithmes. LVADSL, LVADEL et LVADECL désignent respectivement la productivité corrigée par les prix sectoriels, par les prix individuels, et à la fois par les prix individuels et les taux d'utilisation des capacités, où la productivité est mesurée en termes de valeur ajoutée par personne ; LK1L désigne le capital par personne en début d'année mesurée par les immobilisations corporelles brutes au bilan avec correction de prix (cf. annexe) par personne; LL désigne les effectifs ; LPPIES désigne l'écart (le log du rapport) entre prix individuels et sectoriels ; et LTUS désigne le taux d'utilisation des capacités.

Tableau 4 : Estimations par la méthode des moments généralisées (en différences premières et niveaux)

	Avec prix sectoriels	Avec prix individuels	Avec prix et taux d'utilisation individuels, en imposant des coefficients unitaires ou non	
	LVADSL	LVADEL	LVADECL	LVADSL
GMMdif-CEEV				
LK1L	0.18 (0.09)	0.16 (0.09)	0.22 (0.09)	0.18 (0.10)
LL	0.14 (0.16)	0.14 (0.17)	0.16 (0.16)	0.11 (0.16)
LPPIES	-	-	-	0.18 (0.22)
LTUS	-	-	-	0.29 (0.42)
CHI2(dl) / P-valeur	18.6 (10) / 0.04	20.5 (10) / 0.03	21.4 (10) / 0.02	36.7 (20) / 0.01
GMMdif-CEWS				
LK1L	0.57 (0.89)	0.45 (0.93)	0.54 (0.88)	0.10 (0.66)
LL	1.15 (0.75)	1.28 (0.78)	1.14 (0.74)	0.09 (0.50)
LPPIES	-	-	-	-0.94 (0.83)
LTUS	-	-	-	0.10 (0.99)
CHI2(dl) / P-valeur	5.9 (4) / 0.20	4.1 (4) / 0.39	4.7 (4) / 0.32	18.9 (8) / 0.02
GMMniv-CE				
LK1L	0.27 (0.05)	0.26 (0.04)	0.30 (0.04)	0.26 (0.04)
LL	-0.05 (0.06)	-0.07 (0.07)	-0.07 (0.07)	-0.08 (0.05)
LPPIES	-	-	-	0.48 (0.13)
LTUS	-	-	-	0.36 (0.10)
CHI2(dl) / P-valeur	98.6 (22) / 0.00	100.3 (22) / 0.00	88.5 (22) / 0.00	122.9 (44) / 0.00
GMMniv-CEEV				
LK1L	0.35 (0.06)	0.32 (0.06)	0.34 (0.06)	0.29 (0.05)
LL	0.07 (0.08)	0.05 (0.09)	0.03 (0.09)	-0.01 (0.06)
LPPIES	-	-	-	0.45 (0.17)
LTUS	-	-	-	0.30 (0.36)
CHI2(dl) / P-valeur	49.2 (10) / 0.00	51.7 (10) / 0.00	45.2 (10) / 0.00	76.6 (20) / 0.00
GMMniv-CEWS				
LK1L	0.65 (0.13)	0.63 (0.14)	0.65 (0.14)	0.45 (0.09)
LL	0.14 (0.19)	0.17 (0.20)	0.11 (0.20)	-0.19 (0.11)
LPPIES	-	-	-	0.66 (0.27)
LTUS	-	-	-	1.23 (0.85)
CHI2(dl) / P-valeur	7.6 (4) / 0.11	7.6 (4) / 0.11	7.5 (4) / 0.11	28.3 (8) / 0.00

Les estimations reportées sont celles de première étape. Les écart-types sont indiqués entre parenthèses. Des indicatrices temporelles sont ajoutées dans toutes les régressions. GMMdif-CEEV : Instrumentation des différences premières par les niveaux futurs (t+1 et après) et passés (t-2 et avant). GMMdif-CEWS : Instrumentation des différences premières par les niveaux passés (t-2 et avant). GMMniv-CE : Instrumentation des niveaux par les différences premières futures (t+2 et après), présentes (t+1 et t) et passées (t-1 et avant). GMMniv-CEEV : Instrumentation des niveaux par les différences premières futures (t+2 et après) et passées (t-1 et avant). GMMniv-CEWS : Instrumentation des niveaux par les différences premières passées (t-1 et avant).

Echantillon cylindré 1995-1997, 468 entreprises. Définitions et notations des variables : voir tableau 3.