Working Paper 01-55
Statistics and Econometrics Series 27
December 2001

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

GMM ESTIMATION OF A PRODUCTION FUNCTION WITH PANEL DATA: AN APPLICATION TO SPANISH MANUFACTURING FIRMS

César Alonso-Borrego and Rocío Sánchez-Mangas*

Abstract

In this paper we consider the estimation of a Cobb-Douglas production function using a panel dataset of Spanish manufacturing firms. As it is stressed in the econometric literature, the use of standard GMM first differences estimators to eliminate the unobserved firm-specific effects may yield imprecise estimates, particularly in the case of the estimation of the production function. The reason is that the high persistence of output and inputs involved in the estimation of production functions make that their lagged levels to be weak instruments for the first differences of these series. The extended GMM estimator proposed by Arellano and Bover (1995) considers further orthogonality conditions based on lagged differences as instruments for the equation in levels. This approach has been applied to the estimation of technological parameters by Blundell and Bond (1999). Our estimation results, based on this approach, confirm the better performance of the extended GMM estimator compared to the standard first-differenced GMM estimator

Keywords: GMM estimation; panel data; production functions.

*Dpto. de Estadística y Econometría, Universidad Carlos III de Madrid. Alonso-Borrego, e-mail: alonso1@elrond.uc3m.es; Sánchez-Mangas, e-mail: rsman@est-econ.uc3m.es.
Acknowledgments: We thank seminar participants at Universidad Carlos III de Madrid, Universitat Autonoma de Barcelona and Universidad de Vigo for helpful comments. Financial support from Spanish DGI, grant BEC2000-0170, is acknowledged.
Corresponding author: Sánchez-Mangas, Rocío. Universidad Carlos III de Madrid. Dept. of Statistics and Econometrics. C/Madrid, 126, 28903 Getafe (Madrid). Spain. Tel: +34 916249826

GMM Estimation of a Production Function with Panel Data: An Application to Spanish Manufacturing Firms*

César Alonso-Borrego   Rocío Sánchez-Mangas†
Universidad Carlos III de Madrid

December 2001

Abstract

In this paper we consider the estimation of a Cobb-Douglas production function using a panel dataset of Spanish manufacturing firms. As it is stressed in the econometric literature, the use of standard GMM first differences estimators to eliminate the unobserved firm-specific effects may yield imprecise estimates, particularly in the case of the estimation of the production function. The reason is that the high persistence of output and inputs involved in the estimation of production functions make that their lagged levels to be weak instruments for the first differences of these series. The extended GMM estimator proposed by Arellano and Bover (1995) considers further orthogonality conditions based on lagged differences as instruments for the equation in levels. This approach has been applied to the estimation of technological parameters by Blundell and Bond (1999). Our estimation results, based on this approach, confirm the better performance of the extended GMM estimator compared to the standard first-differenced GMM estimator.

Keywords: GMM estimation, panel data, production functions

* *Acknowledgments* : We thank seminar participants at Universidad Carlos III de Madrid, Universitat Autonoma de Barcelona and Universidad de Vigo for helpful comments. Financial support from Spanish DGI, grant BEC2000-0170, is acknowledged.

† *Corresponding author* : Sánchez-Mangas, Rocío. Universidad Carlos III de Madrid. Dept. of Statistics and Econometrics. C/Madrid, 126, 28903 Getafe (Madrid). Spain. $e-mail$: rsman@est-econ.uc3m.es Tel : +34 916249826

1 Introduction

The estimation of linear panel data models with predetermined variables is typically done by means of GMM estimators applied to the first differences transformation of the equation of interest, where all the available lags of the predetermined variables are used as instruments. The purpose of this approach is to remove time-invariant unobserved individual heterogeneity. However, this approach yields poorly precise estimates in the case of panel with a small number of time periods with highly persistent data.

In this paper, we focus on the estimation of production functions from panel datasets covering a large number of firms observed for a small number of time periods. In this context, as it has been stressed in Mairesse and Hall (1996), the application of first-differences GMM estimators with lagged levels of the series as instruments has produced insatisfactory results. More specifically, the coefficient of the capital stock is generally low and statistically nonsignificant, and returns to scale appear to be unreasonably low.

Blundell and Bond (1999) suggest that the problem of "weak instruments" is behind the poor performance of standard GMM estimators in this context. The fact that the variables entering the production function modelling, i.e, firm sales, capital, and employment, are highly persistent, induces a weak correlation between the first differences and the lagged levels of these variables. This problem of weak instruments can cause large finite-sample biases and poor precision in the estimators. They also show that these biases could be dramatically reduced by applying the extended GMM estimator proposed by Arellano and Bover (1995). This estimator, labelled as "system GMM", is based on an augmented system which includes level equations with lagged differences as instruments in addition to the differenced equations with lagged levels as instruments. Blundell and Bond (1999) and Blundell, Bond and Windmeijer (2000) apply this "system GMM" to the estimation of production functions from US company panel data. They find that the system GMM greatly improves the performance of the first-differenced GMM estimator. Their estimation results provide a strongly significant capital coefficient, and confirm that the lagged differences are informative instruments for the endogenous variables in levels.

We exploit an unbalanced panel of 1272 Spanish manufacturing firms between 1990 and 1997 from the Encuesta de Estrategias Empresariales to apply a similar approach. We obtain both the first-differenced and the system GMM estimator. Our estimation results resemble the findings in Blundell and Bond (1999) and Blundell, Bond and Windmeijer (2000), that a great improvement in terms of efficiency is obtained with the system GMM estimator.

The rest of the paper is organized as follows. In the next section we describe our specification of the production function. Section 3 briefly describes the econometric issues related to the first-differenced and the system GMM estimator. Section 4 describes the dataset used in this study. Section 5 presents the estimation results and Section 6 concludes.

2 The model

We consider a Cobb-Couglas production function without imposing constant returns to scale:

$$Y_{nt} = a_{nt}\ K_{nt}^{\alpha_K}\ L_{nt}^{\alpha_L} \qquad n = 1, ..., N;\ t = 1, ..., T_n$$

where firms are indexed by n and time is indexed by t, Y_{nt} is production output of firm n at period t, K_{nt} represents capital stock, L_{nt} is the employment and a_{nt} is a productivity shock. Taking logarithms we obtain:.

$$y_{nt} = \alpha_K\ k_{nt} + \alpha_L\ l_{nt} + u_{nt} \tag{1}$$

where $y_{nt} = \ln(Y_{nt})$, $k_{nt} = \ln(K_{nt})$, $l_{nt} = \ln(L_{nt})$ and $u_{nt} = \ln(a_{nt})$. We specify the following structure for the productivity shock:

$$\begin{aligned} u_{nt} &= A_t + \eta_n + v_{nt} \\ v_{nt} &= \rho\ v_{n,t-1} + \xi_{nt} \end{aligned} \tag{2}$$

where A_t is an aggregate effect, η_n is a time invariant firm-specific effect, v_{nt} is an $AR(1)$ idiosyncratic shock and the ξ_{nt} are independently distributed with zero mean and variance σ_ξ^2.

In order to estimate the parameters $(\alpha_K, \alpha_L, \rho)$, we formulate the dynamic representation of (1):

$$\begin{aligned} y_{nt} = \ & \alpha_K \ k_{nt} - \alpha_K \ \rho \ k_{n,t-1} + \alpha_L \ l_{nt} - \alpha_L \ \rho \ l_{n,t-1} + \rho \ y_{i,t-1} \\ & + (A_t - \rho \ A_{t-1}) + (1-\rho) \ \eta_n + \xi_{nt} \end{aligned} \tag{3}$$

or

$$\begin{aligned} y_{nt} = \ & \pi_1 \ k_{nt} + \pi_2 \ k_{n,t-1} + \pi_3 \ l_{nt} + \pi_4 \ l_{n,t-1} + \pi_5 \ y_{i,t-1} \\ & + A_t^* + \eta_n^* + \xi_{nt} \end{aligned} \tag{4}$$

subject to two non-linear restrictions: $\pi_2 = -\pi_1\pi_5$ and $\pi_4 = -\pi_2\pi_5$, and where $A_t^* = A_t - \rho A_{t-1}$ and $\eta_n^* = (1-\rho)\eta_n$.

Given consistent estimates of the unrestricted parameter vector $\pi = (\pi_1, \pi_2, \pi_3, \pi_4, \pi_5)'$ and its variance-covariance matrix, the restrictions can be tested and imposed by minimum distance to obtain estimates for the restricted parameter vector $(\alpha_K, \alpha_L, \rho)'$.

3 First-differenced and system GMM estimators

In order to ilustrate the problem of weak instruments in the first-differenced GMM estimator, Blundell and Bond (1999) consider an $AR(1)$ model in which no additional regressors have been included. Let us consider the following model

$$y_{nt} = \rho y_{n,t-1} + \eta_n + \xi_{nt}$$

with the usual assumptions:

$$E(\eta_n) = 0, \ E(\xi_{nt}) = 0, \ E(\xi_{nt}\eta_n) = 0 \quad \text{for } n = 1, ..., N \text{ and } t = 2, ..., T \tag{5}$$

$$E(\xi_{nt}\xi_{ns}) = 0 \quad \text{for } n = 1, ..., N \text{ and } \forall t \neq s \tag{6}$$

and the standard additional assumption on the initial conditions:

$$E(y_{n1}\xi_{nt}) = 0 \quad \text{for } n = 1, ..., N \text{ and } t = 2, ..., T \tag{7}$$

These assumptions imply the following orthogonality conditions:

$$E(y_{n,t-s}\Delta\xi_{nt}) = 0 \quad \text{for } t = 3, ..., T \text{ and } 2 \leq s \leq t-1 \tag{8}$$

We can construct a GMM estimator based on these moment conditions.

The instruments used in the first-differenced GMM estimator are less informative in two cases: when the autoregressive parameter ρ is close to 1 or when the variance of the individual effect, relative to the variance of ξ_{nt}, is large. Considering the simple case of $T = 3$, we have only one orthogonality condition and the first-differenced GMM estimator is just a simple instrumental variable estimator of π in the equation:

$$\Delta y_{n2} = \pi y_{n1} + r_n, \qquad \text{for } n = 1, ..., N \tag{9}$$

Assuming stationarity and letting $\sigma_\eta^2 = var(\eta_n)$ and $\sigma_\xi^2 = var(\xi_{nt})$, the plim of the OLS estimator of π, $\hat{\pi}$ is given by:

$$\text{p}\lim \hat{\pi} = (\rho - 1)\frac{k}{\sigma_\eta^2 \big/ \sigma_\xi^2 + k} \qquad \text{where } k = \frac{1-\rho}{1+\rho} \tag{10}$$

When π is close to zero in (9), the instrument y_{n1} is only weakly correlated with Δy_{n2}. As we can see in (10), plim $\hat{\pi} \to 0$ as $\rho \to 1$ or as $\left(\sigma_\eta^2 \big/ \sigma_\xi^2\right) \to \infty$.

The performance of the estimation can be improved if we consider the additional assumption on the initial conditions:

$$E\left(\eta_n \Delta y_{n2}\right) = 0 \quad \text{for } n = 1, ..., N \tag{11}$$

If condition (11) holds in addition with (5), (6) and (7), the following $T - 2$ moment conditions are valid:

$$E\left[\left(\eta_n + \xi_{nt}\right)\Delta y_{n,t-1}\right] = 0 \quad \text{for } t = 3, ..., T \tag{12}$$

We can calculate a GMM estimator using the full set of moment conditions given by equations (8) and (12).

The standard approach for testing the validity of the moment conditions in GMM estimation is the Hansen-Sargan test of overidentifying restrictions. Under the null hypothesis that moment conditions are valid, the test statistic given by the GMM criterion multiplied by the sample size is asymptotically distributed as a chi-squared, with degrees of freedom equal to the number of moment conditions minus the number of parameters estimated. Since the system GMM estimator differs from the first-differences GMM estimator by the additional moment conditions that the first one exploits, we can test for the validity of these additional conditions by means of a Sargan difference test. The statistic for such test is just the difference between the Hansen-Sargan statistics for each of these two estimators. Under the null that the level moment conditions are valid (provided that the differenced equations with lagged levels as instruments are valid), the resulting statistic is asymptotically distributed as chi-squared with degrees of freedom equal to the number of additional moment conditions that the system GMM incorporates.

4 The data

The dataset used in this study has been taken from the Encuesta de Estrategias Empresariales (ESEE) conducted by the Spanish Ministry of Industry and Energy. It contains annual information of the balance sheet and other economic variables for a large number of Spanish manufacturing firms. Our sample is an unbalanced panel of 1272 firms between 1990 and 1997.

Table 1 shows the factor shares by industry. For each productive factor, capital and labor, it has been calculated as the proportion that its payment represents over the total payment of factors. As we can see, the average capital share is 0.450 and the average labor share is 0.550 for the whole sample. There are some variations by industry. The capital share (labor share) varies from 0.328 (0.672) in the leather industry to 0.589 (0.411) in the iron, steel and metal industry.

Table 2 shows the distribution of firms by industry and size, measured as the number of employees. We have considered three categories which correspond to the classification that the European Comission establish to define small, medium and large firms in terms of the number of workers. The first category, small firms, is composed by firms with no more than 50 employees. The second one, medium firms, are those with more than 50 and no more than 250 employees. Large firms are those with more than 250 employees. As we can see, more than half of the firms are small firms and only 27.44% are large firms. There are important variations by industry. While large firms are in the majority at industries like iron, steel and metals, motor vehicles or ship building, in industries like leather, garment or wood and furniture, more than two thirds of the firms are small, according to the number of employees.

In Table 3 we can see the distribution of firms by industry and size, where now the size is measured by the annual turnover. Again, we have considered three categories which correspond to the classification that the European Comission establish to define small, medium and large firms in terms of the annual turnover. The first category, small firms, is composed by firms that invoice no more than 7 millions euros per year. The second one, medium firms, are those with a turnover greater than 7 and smaller than 40 millions euros per year. Large firms are those which invoice more than 40 millions euros per year. Only 15% of the firms are in the third category according to this criterion, while almost 60% are considered small firms. By industry, again, large firms are more present in industries like iron, steel and metal and motor vehicles, while small firms are in the majority at industries like leather, garment, wood and furniture and cellulose and paper edition.

Looking at tables 2 and 3 together, it can be seen that in industries in which the majority are small firms, the capital share is around 0.3 or 0.4, while in industries where large firms predominate, this share rises until 0.5.

5 Estimation results

We have obtained estimates of the production function (1). The results are reported in Table 4. We report results for the two-step GMM estimator for both the first-differenced equations and the system. We take as instruments the lagged levels dated $t-2$ and earlier in the first-differenced equations. As additional instruments in the system GMM estimation, we take the lagged differences dated $t-1$. Year dummies have been included in both models. The non-linear restrictions implied by the technological parameters can be then tested and imposed by minimum distance.

Table 4 reports the unrestricted production function estimates, without imposing constant returns to scale. The upper panel provides the estimates for the linear equation (4) using the first-differences GMM and the system GMM estimator in the first and second column, respectively. In the bottom panel, we recover the technological parameters by means of a minimum distance estimator that exploits the constraints associated with the technological parameters. The first differences GMM estimation provides a positive though nonsignificant coefficient for the capital stock, yet the estimated value does not appear to be unreasonably low, as it appeared to be in Mairesse and Hall (1996) or Blundell and Bond (1999) for US company data. Nevertheless, the gains of the system GMM estimation compared with the first-differences GMM estimates, are very apparent: the precision of all the estimates improves considerably, specially in the estimation of the capital coefficient.

The specification tests do not provide evidence against the model. First, the orthogonality conditions in the first-differenced equations with lagged levels dated $t-2$ and earlier as instruments are not rejected. Second, the autocorrelation tests $m1$ and $m2$ (see notes on Table 5) are consistent with the $AR(1)$ structure that we have assumed for the idiosyncratic error term. Third, the low value of the difference Sargan test clearly does not reject the additional orthogonality conditions associated with levels equations with lagged differences as instruments. Additionally, we can see that constant returns to scale and non-linear restrictions are not rejected in none of the models.

We have also estimated the parameters in the production function imposing constant returns to scale. The results are reported in Table 5. Neither the Hansen-Sargan test nor the first and second order autocorrelation tests provide any evidence against the specification. Again, the gains obtained with the system GMM estimation with respect to the first-differences GMM estimation, in terms of the precision of the estimates, are noticiable.

6 Conclusions

In this paper we have estimated a Cobb-Douglas production function using a panel dataset of Spanish manufacturing firms. We have obtained both the standard first-differenced GMM estimator and the extended GMM estimator. Our estimation results confirm the findings of Blundell and Bond (1999) for a panel of US companies. When we apply the first-differenced GMM estimator not imposing constant returns to scale, we obtain poorly precise estimates. The precision of the estimates greatly improves when the extended GMM estimator is used. When we impose constant returns to scale, the results from the first-differenced GMM estimator are better that not imposing that restrictions. Nevertheless, also in this case we obtain a great improvement in terms of efficiency when the system GMM estimator is applied. Both imposing and not imposing constant returns to scale, we find that the additional instruments used in the system GMM estimator are valid and informative.

As it is stressed in Blundell and Bond (1999), similar results have been found when system GMM estimator has been applied in different contexts, such as labor demand equations or investment equations, among others.

References

[1] 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, 277-297.

[2] Arellano, M. and S. Bond (1998): "Dynamic panel data estimation using DPD98 for GAUSS: a guide for users", December 1998.

[3] Arellano, M. and O. Bover (1995): "Another look at the instrumental variable estimation of error-component models", *Journal of Econometrics*, 68, 29-52

[4] Blundell, R. and S. Bond (1999): "GMM estimation with persistent panel data: an application to production functions", *The Institute for Fiscal Studies*, Working Paper Seires No. 99/4.

[5] Blundell, R., S. Bond and F. Windmeijer (2000): "Estimation in dynamic panel data models: Improving the performance os the standard GMM estimators", *The Institute for Fiscal Studies*, Working Paper Series No. 00/12.

[6] Mairesse, J. and B.H. Hall (1996): "Estimating the productivity of research and development in French and US manufacturing firms: an exploration of simultaneity issues with GMM methods", in Wagner, K. and B. Van Ark (eds), *International Productivity Differences and Their Explanations*, Elsevier Science, 285-315.

[7] Mairesse, J., B.H. Hall and B. Mulkay (1999): "Firm-level investment in France and the United States: an exploration of what we have learned in twenty years", *NBER*, Working Paper 7437.

Appendix

A Construction of variables

Employment: Number of employees is calculated at december 31th, as the sum of permanent workers and the average number of temporary workers. The weights to calculate the average number of temporary workers is: 1/4 if the average time in the firm is less than 6 months, 3/4 if the average time is more than 6 months and less than one year and 1 if it is more than one year.

Output: Gross output at retail prices is calculated as total sales.

Capital stock: The dataset contains information on the book value and the average age of the stock of fixed capital and the year of the last regulation. It also includes data on gross nominal investment during the year. Following Alonso-Borrego and Collado (1999), taking period t as reference year, the market value of the stock of fixed capital in period t is calculated as:

$$K_{nt} = (1-\delta_n)^{age_{nt}} KB_{nt} \frac{q_t}{q_{m_n}}$$

where age_{nt} is the average age of the capital stock of firm n at period t, δ_n is the depreciation rate of the sector in which firm n operates, KB_{nt} is the book value of the stock of fixed capital, q_t is the price deflator of the stock of fixed capital and m_n is the year of the last regulation in firm n. The price index is the GDP implicit deflator of investment goods, which is constant over time. The depreciation rate varies across sectors.

Taking t as the reference year, the market value of the stock of fixed capital for any year $s \neq t$ is calculated using a perpetual inventory method:

$$K_{ns} = (1-\delta_n)K_{n,s-1}\frac{q_s}{q_{s-1}} + I_{ns} \quad \text{if} \quad s > t$$

$$K_{ns} = \frac{(K_{n,s+1} - I_{n,s+1})}{(1-\delta_n)}\frac{q_s}{q_{s+1}} \qquad \text{if} \quad s < t$$

where I_{ns} is the investment accounted by the firm n in period t. Using this approach it is possible to obtain negative values of K_{ns} for $s < t$. In that case the market value of the capital stock is set to missing. In an attempt to reduce this problem, the market value of the capital stock for any firm has been calculated using different years as reference. Finally, the reference year was chosen to minimize the number of missing values in the capital stock.

B Data description

Average factor shares by industry		
Industry	Capital	Labor
Iron, steel and metal (22)	0.589	0.411
Bldg. materials, glass, ceramics (24)	0.489	0.511
Chemicals (25)	0.476	0.524
Non-ferrous metal (31)	0.411	0.589
Basic machinery (32)	0.342	0.658
Office machinery (33)	0.360	0.640
Electric materials (34)	0.373	0.627
Electronic (35)	0.373	0.627
Motor vehicles (36)	0.502	0.498
Ship building (37)	0.373	0.627
Other motor vehicles (38)	0.497	0.503
Precision instruments (39)	0.371	0.629
Non-elaborated food (41)	0.523	0.477
Food, tobacco and drinks (42)	0.539	0.461
Basic Textile (43)	0.505	0.495
Leather (44)	0.328	0.672
Garment (45)	0.351	0.648
Wood and furniture (46)	0.424	0.575
Cellulose and paper edition (47)	0.478	0.522
Plastic materials (48)	0.504	0.496
Other non-basic (49)	0.432	0.568
Total	0.450	0.550

Table 1: Factor shares by industry

Distribution of firms by industry and size					
Industry		Number of employees			
		<=50	51-250	>250	Total
Iron, steel	Abs. freq	8	5	19	32
and metal	% by ind.	25.00	15.63	59.38	100.00
(22)	% by size	1.22	1.88	5.44	2.52
Building materials,	Abs. freq.	46	22	21	89
glass, ceramics	% by ind.	51.69	24.72	23.60	100.00
(24)	% by size	7.00	8.27	6.02	7.00
Chemicals	Abs. freq.	36	17	43	96
	% by ind.	37.50	17.71	44.79	100.00
(25)	% by size	5.48	6.39	12.32	7.55
Non-ferrous	Abs. freq.	80	23	25	128
metal	% by ind.	62.50	17.97	19.53	100.00
(31)	% by size	12.18	8.65	7.16	10.06
Basic	Abs. freq.	37	17	19	73
machinery	% by ind.	50.68	23.29	26.03	100.00
(32)	% by size	5.63	6.39	5.44	5.74
Office	Abs. freq.	0	0	1	1
machinery	% by ind.	0.00	0.00	100.00	100.00
(33)	% by size	0.00	0.00	0.29	0.08
Electric	Abs. freq.	38	20	29	87
materials	% by ind.	43.68	22.99	33.33	100.00
(34)	% by size	5.78	7.52	8.31	6.84
Electronic	Abs. freq.	8	8	15	31
	% by ind.	25.81	25.81	48.39	100.00
(35)	% by size	1.22	3.01	4.30	2.44
Motor vehicles	Abs. freq.	7	17	29	53
	% by ind.	13.21	32.08	54.72	100.00
(36)	% by size.	1.07	6.39	8.31	4.17
Ship	Abs. freq.	7	3	8	18
building	% by ind.	38.89	16.67	44.44	100.00
(37)	% by size	1.07	1.13	2.29	1.42
Other	Abs. freq.	0	2	5	7
motor vehicles	% by ind.	0.00	28.57	71.43	100.00
(38)	% by size	0.00	0.75	1.43	0.55

Table 2: Firms distribution by industry and size (number of employees)

Distribution of firms by industry and size (cont.)					
Industry		Number of employees			
		<=50	51-250	>250	Total
Precision instruments (39)	Abs. freq.	2	3	3	8
	% by ind.	25.00	37.50	37.50	100.00
	% by size	0.30	1.13	0.86	0.63
Non-elaborated food (41)	Abs. freq	80	32	33	145
	% by ind.	55.17	22.07	22.76	100.00
	% by size	12.18	12.03	9.46	11.40
Food, tobacco and drinks (42)	Abs. freq.	21	7	26	54
	% by ind.	38.89	12.96	48.15	100.00
	% by size	3.20	2.63	7.45	4.25
Basic Textile (43)	Abs. freq.	22	17	16	55
	% by ind.	40.00	30.91	29.09	100.00
	% by size	3.35	6.39	4.58	4.32
Leather (44)	Abs. freq.	10	3	1	14
	% by ind.	71.43	21.43	7.14	100.00
	% by size	1.52	1.13	0.29	1.10
Garment (45)	Abs. freq.	73	21	15	109
	% by ind.	66.97	19.27	13.76	100.00
	% by size	11.11	7.89	4.30	8.57
Wood and furniture (46)	Abs. freq.	66	5	4	75
	% by ind.	88.00	6.67	5.33	100.00
	% by size	10.05	1.88	1.15	5.90
Cellulose and paper edition (47)	Abs. freq.	63	16	21	100
	% by ind.	63.00	16.00	21.00	100.00
	% by size	9.59	6.02	6.02	7.86
Plastic materials (48)	Abs. freq.	35	23	11	69
	% by ind.	50.72	33.33	15.94	100.00
	% by size	5.33	8.65	3.15	5.42
Other non-basic (49)	Abs. freq.	18	5	5	28
	% by ind.	64.29	17.86	17.86	100.00
	% by size.	2.74	1.88	1.43	2.20
Total	Abs. freq.	657	266	349	1272
	% by ind.	51.65	20.91	27.44	100.00
	% by size	100.00	100.00	100.00	100.00

Distribution of firms by industry and annual turnover					
Industry		Avg. turnover (10^6 euros)			
		<7	7-40	>40	Total
Iron, steel	Abs. freq	12	2	18	32
and metal	% by ind.	37.50	6.25	56.25	100.00
(22)	% by turn.	1.61	0.59	9.42	2.52
Building materials,	Abs. freq.	55	28	6	89
glass, ceramics	% by ind.	61.80	31.46	6.74	100.00
(24)	% by turn.	7.39	8.31	3.14	7.00
Chemicals	Abs. freq.	38	23	35	96
	% by ind.	39.58	23.96	36.46	100.00
(25)	% by turn.	5.11	6.82	18.32	7.55
Non-ferrous	Abs. freq.	91	31	6	128
metal	% by ind.	71.09	24.22	4.69	100.00
(31)	% by turn.	12.23	9.20	3.14	10.06
Basic	Abs. freq.	41	26	6	73
machinery	% by ind.	56.16	35.62	8.22	100.00
(32)	% by turn.	5.51	7.72	3.14	5.74
Office	Abs. freq.	0	0	1	1
machinery	% by ind.	0.00	0.00	100.00	100.00
(33)	% by turn.	0.00	0.00	0.52	0.08
Electric	Abs. freq.	45	29	13	87
materials	% by ind.	51.72	33.33	14.94	100.00
(34)	% by turn.	6.05	8.61	6.81	6.84
Electronic	Abs. freq.	11	14	6	31
	% by ind.	35.48	45.16	19.35	100.00
(35)	% by turn.	1.48	4.15	3.14	2.44
Motor vehicles	Abs. freq.	10	28	15	53
	% by ind.	18.87	52.83	28.30	100.00
(36)	% by turn.	1.34	8.31	7.85	4.17
Ship	Abs. freq.	8	7	3	18
building	% by ind.	44.44	38.89	16.67	100.00
(37)	% by turn.	1.08	2.08	1.57	1.42
Other	Abs. freq.	0	2	5	7
motor vehicles	% by ind.	0.00	28.57	71.43	100.00
(38)	% by turn.	0.00	0.59	2.62	0.55

Table 3: Firms distribution by industry and size (annual turnover)

Distribution of firms by industry and annual turnover (cont.)					
lndustry		Avg. turnover (10^6 euros)			
		<7	7-40	>40	Total
Precision instruments (39)	Abs. freq.	2	5	1	8
	% by ind.	25.00	62.50	12.50	100.00
	% by size	0.27	1.48	0.52	0.63
Non-elaborated food (41)	Abs. freq	86	32	27	143
	% by ind.	59.31	22.07	18.62	100.00
	% by turn.	11.56	9.50	14.14	11.40
Food, tobacco and drinks (42)	Abs. freq.	23	12	19	54
	% by ind.	42.59	22.22	35.19	100.00
	% by turn.	3.09	3.56	9.95	4.25
Basic Textile (43)	Abs. freq.	31	20	4	55
	% by ind.	56.36	36.36	7.27	100.00
	% by turn.	4.17	5.93	2.09	4.32
Leather (44)	Abs. freq.	11	3	0	14
	% by ind.	78.57	21.43	0.00	100.00
	% by turn.	1.48	0.89	0.00	1.10
Garment (45)	Abs. freq.	86	19	4	109
	% by ind.	78.90	17.43	3.67	100.00
	% by turn.	11.56	5.64	2.09	8.57
Wood and furniture (46)	Abs. freq.	68	6	1	75
	% by ind.	90.67	8.00	1.33	100.00
	% by turn.	9.14	1.78	0.52	5.90
Cellulose and paper edition (47)	Abs. freq.	66	20	14	100
	% by ind.	66.00	20.00	14.00	100.00
	% by turn.	8.87	5.93	7.33	7.86
Plastic materials (48)	Abs. freq.	39	24	6	69
	% by ind.	56.52	34.78	8.70	100.00
	% by turn.	5.24	7.12	3.14	5.42
Other non-basic (49)	Abs. freq.	21	6	1	28
	% by ind.	75.00	21.43	3.57	100.00
	% by turn.	2.82	1.78	0.52	2.20
Total	Abs. freq.	744	337	191	1272
	% by ind.	58.49	26.49	15.02	100.00
	% by turn.	100.00	100.00	100.00	100.00

C Estimation results

GMM estimates		
	First differences $t-2$	System $t-2$
k_{nt}	0.332 (0.313)	0.444 (0.154)
$k_{n,t-1}$	-0.348 (0.203)	-0.378 (0.132)
l_{nt}	0.754 (0.310)	0.568 (0.199)
$l_{n,t-1}$	-0.609 (0.212)	-0.255 (0.235)
$y_{n,t-1}$	0.849 (0.136)	0.724 (0.097)
$m1$	-5.695	-7.572
p-value	0.000	0.000
$m2$	0.073	-0.563
p-value	0.942	0.574
Sargan	29.474	31.684
p-value	0.245	0.629
Dif. Sargan	—	2.210
p-value	—	0.980
Minimun distance estimates		
α_k	0.435 (0.225)	0.456 (0.121)
α_L	0.722 (0.231)	0.576 (0.195)
ρ	0.874 (0.082)	0.819 (0.074)
p-value MD test	0.658	0.314
p-value CRS test	0.660	0.880

Table 4: Production function estimates. See notes on Table 5

GMM estimates		
	First differences $t-2$	System $t-2$
k_{nt}	0.289	0.436
	(0.154)	(0.113)
$k_{n,t-1}$	-0.301	-0.368
	(0.117)	(0.092)
$y_{n,t-1}$	0.864	0.844
	(0.090)	(0.061)
$m1$	-7.446	-10.659
p-value	0.000	0.000
$m2$	0.036	-0.245
p-value	0.971	0.807
Sargan	25.811	39.539
p-value	0.529	0.357
Dif. Sargan	—	13.728
p-value	—	0.800
Minimun distance estimates		
α_k	0.357	0.436
	(0.133)	(0.109)
α_L	0.643	0.564
	—	—
ρ	0.878	0.844
	(0.088)	(0.058)
p-value MD test	0.382	0.993

Table 5: Production function estimates imposing constant returns to scale.
Standard errors in parenthesis
m1,m2: tests for first and second order correlation in first differenced residuals
Sargan: Sargan test of overidentifying restrictions
Dif Sargan: test of the validity of the additional instruments in system estimation
MD: minimum distance
CRS: constant returns to scale
Year dummies included in all models