INTERNATIONAL ECONOMIC REVIEW
Vol. 18, No. 2, June, 1977

EFFICIENCY ESTIMATION FROM COBB-DOUGLAS PRODUCTION FUNCTIONS WITH COMPOSED ERROR*

By Wim Meeusen and Julien van den Broeck [1]

1. INTRODUCTION

In the recent economic literature (Aigner and Chu [**2**], Førsund and Hjalmarsson [**5**], Seitz [**11**], and Timmer [**12**]), in continuation of the research initiated by Farrell [**4**], much attention has been paid to the estimation of productive efficiency by means of frontier production functions. In this journal, Richmond [**10**], developing a suggestion made by Afriat [**1**], carried out the estimation of the Cobb-Douglas production frontier by letting

$$y_t = f(x_t)k_t, \qquad t = 1,\ldots, T \tag{1}$$

and applying ordinary least squares to the logarithmic transformation of (1), after putting $k_t = e^{-z_t}$, assigning a Gamma distribution $G_Z(z;\, n)$ to the random variable Z, and introducing the statistical error $v_t = n - z_t$. y_t is the actual output, x_t is the corresponding $(M \times 1)$ vector of inputs, and $f(x_t)$ is the Cobb-Douglas frontier production function.

It follows that the random variable K, with observed values $k_t (t=1,\ldots, T)$, represents the technical level of efficiency at the firm level, i.e., the proportion between actual and potential output.

$\hat{\varepsilon}_R = 2^{-\hat{n}}$ is a consistent although upward biased estimator of the average Richmond-efficiency level of the economy (or industry) given by

$$\varepsilon_R = E(K) = \int_0^\infty e^{-z} dG_Z(z;\, n) = 2^{-n}. \tag{2}$$

$\hat{n} = (\sum_t v_t^2)/(T-M-1)$ is the OLS estimator of the variance n of the random variable V.

Of course, the estimates of the production elasticities are not affected in the foregoing specification: the Afriat-Richmond production frontier comes about by a mere upward shift on the logarithmic scale of the corresponding "average" production function.

* Manuscript received January 16, 1976; revised May 24, 1976.
[1] We should like to express our gratitude to Jacques Mairesse of INSEE at Paris for permitting us to use the data of the 1962 French Census of Manufacturing Industries. We are also grateful to Dr. H. Deckers and G. Peperstraete of the Nuclear Research Centre at Mol (Belgium) for their asistance with the calculations.

It is our purpose to propose an alternative approach to the estimation of the frontier production function, since the Afriat-Richmond measure yields, as we shall show, systematically underestimated values for productive efficiency.

Contrary to Afriat and Richmond, who proceed as if every disturbance from the theoretical value on the frontier results solely from human errors (information deficiencies, adjustment costs, managerial errors on the level of the technical operation of the plant, etc.), we shall introduce in our model of the frontier production function, next to a disturbance due to inefficiency, a statistical disturbance due to randomness in the real sense and to specification and measurement errors.

2. THE EFFICIENCY MODEL WITH COMPOSED ERROR

We propose an efficiency model with a composed multiplicative disturbance term, i.e., the product of a "true" error term and an efficiency measure:

$$y_t = \phi(x_t)k_t u_t, \qquad t = 1,\ldots, T.$$

k_t is an efficiency measure and is distributed in the (0, 1) interval; u_t is a disturbance in the proper sense and is distributed in the $(0, \infty)$ interval. u_t, k_t and x_t are mutually independent.

Considering the Cobb-Douglas production model and putting $k_t = \exp(-z_t)$ and $u_t = \exp(-v_t)$, we get

$$y_t = A \prod_j x_{tj}^{\beta_j} e^{-z_t} e^{-v_t}.$$

We assume that the z_t as well as the v_t have a specified theoretical distribution. The convolution of those distributions and a subsequent straightforward application of the maximum likelihood method yields estimates of the relevant parameters.

It is clear that the two types of errors and not separated with regard to each piece of available information. Once the parameters of the distribution obtained through convolution are estimated, however certain characteristics, such as moments, of the *pdf*'s of respectively K and U can be derived. The two kinds of errors are separated a posteriori in this sense.

It is assumed that the v_t are a random sample of a Gaussian distribution, with zero mean and variance σ^2.

The z_t are of course distributed in the interval $(0, \infty)$.

It turns out that assigning an exponential distribution to z leads to acceptable properties of the model, both from a theoretical and computational point of view.

Other distributions, which from an economic point of view seemed to be more adequate (that is, distributions the *pdf* of which have a modal value between 0 and 1, such as the beta distribution or its right-truncated version) led to insurmountable difficulties in finding a neat expression for the *pdf* of the composed error.

For the pdf of Z we put

$$f_Z(z;\lambda) = \lambda e^{-\lambda z}, \quad z \geq 0 \qquad (\lambda > 0)$$
$$= 0 \quad , \quad z < 0.$$

We get for the pdf of $K = e^{-z}$

$$f_K(k;\lambda) = \lambda k^{\lambda-1}, \quad 0 < k \leq 1$$
$$= 0 \quad , \quad \text{elsewhere.}$$

The *pdf* of k is monotone decreasing or increasing for $0<\lambda<1$, resp. $\lambda>1$. If $\lambda=1$, k is uniformly distributed in the interval.

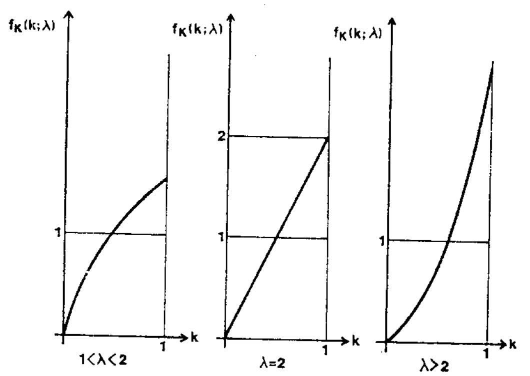

FIG. 1 a

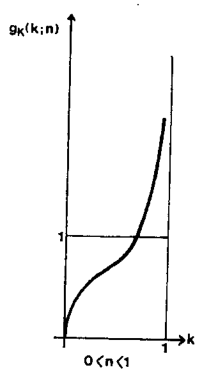

FIG. 1 b

There are three possible shapes for the pdf corresponding to the situation where higher efficiencies have higher probability densities. From an economic point of view only these three cases can be considered as meaningful (Figure 1a).

This is in contrast with the pdf for k, $g_K(k;n)$, generated by a Gamma probability density function (the procedure followed by Afriat and Richmond) where the above mentioned situation only corresponds to one specific interval for the relevant parameter ($0<n<1$; see Figure 1b). The exponential pdf for Z clearly imposes fewer restrictions on the distribution of efficiency than the Gamma pdf.

An additional argument in favor of the exponential pdf consists in thé finite values of the probability density corresponding to a maximal level of efficiency. This is not the case with the pdf for $K=e^{-z}$ resulting from a Gamma-distribution for Z.

Most of all, the choice of an exponential distribution for Z, leads to neat expressions for the average efficiency $\varepsilon_c = E(K)$, and for the pdf of the composed error term $W = Z + V$ and its moments.

For the average efficiency level ε_c, we obtain

$$\varepsilon_c = E(K) = \frac{\lambda}{\lambda+1}. \tag{3}$$

Using the general expression for the pdf of a sum of two independent random variables with respective pdf's $f_Z(z)$ and $f_V(v)$, we get after some transformations

$$\begin{aligned} f_W(w) &= \int_{-\infty}^{+\infty} f_Z(z) f_V(w-z)\,dz \\ &= \frac{\lambda}{2} \exp\left(\frac{\sigma^2\lambda^2}{2} - \lambda w\right) \operatorname{erfc}\left(\frac{\sigma^2\lambda - w}{\sigma\sqrt{2}}\right). \end{aligned} \tag{4}$$

The following moments can easily be computed:

$$\mu_1 = \frac{1}{\lambda}, \qquad \mu_2 = \frac{1}{\lambda^2} + \sigma^2,$$

$$\mu_3 = \frac{2}{\lambda^3}, \qquad \mu_4 = \frac{9}{\lambda^4} + \frac{6\sigma^2}{\lambda^2} + 3\sigma^4.$$

The above expression for μ_2 is useful in determining which part of the total regression error is attributable to respectively inefficient operating and "true" statistical disturbance. The skewness measure is positive and therefore indicates asymmetry to the right. From μ_4 we conclude that the distribution of the composed error is leptocurtic, i.e., the corresponding pdf is sharper around the mode than normal. If λ tends to infinity, i.e., efficiency approaches the 100% level, $f_W(w)$ obviously tends to the normal distribution.

It is clear that the average efficiency like defined in expression (3) can now no longer be interpreted, as do Afriat and Richmond, as the average proportion ρ between the actual and potential production level, since it holds that

$$\rho = \int_{\Omega_X}\int_{\Omega_W} \frac{y(x,w)}{\phi(x)} f_X(x) f_W(w)\,dx\,dw = \int_{\Omega_W} e^{-w} f_W(w)\,dw = \frac{\lambda}{\lambda+1}\, e^{\frac{\sigma^2}{2}}$$

It follows immediately that the average value of the proportion ρ is a thoroughly misleading measure of efficiency.

3. DATA AND ESTIMATION PROCEDURE

The data from the 1962 French Census of Manufacturing Industries have been used for a first verification of the foregoing model. The same data have already been handled by Mairesse **[8]** to estimate production functions at the micro-level for France. He compared them with similar production function results for Norway, estimated by Griliches and Ringstad **[6]**. The results of the latter enabled Richmond to implement his method to measure productive efficiency. Ten industries have been selected; in decreasing order of the number of firms at issue: machine construction and mechanical tools (953), textiles (824), electrical engineering (679), paper (523), industrial chemicals (415), vehicles and cycles (356), footwear (341), milk products (320), sugar-works, distilleries and beverages (292), and glass products (135). In each industry, the sample consists of all corporate firms employing at least 20 workers and other salaried employees. The production variable is measured by the value added at factor prices. The labor variable is the unweighted sum of workers.

For the capital variable only the book value of gross fixed assets was available. This seems justifiable: Griliches and Ringstad used both a capital service flow measure based on fire insurance values of machinery and buildings and given depreciation rates, and capital stock measures; their results were practically unaffected.

We define the loglikelihood function as

$$\log L(w|\theta) = \sum_i \log f_W(w_i/\theta)\,.$$

θ is the shorthand expression for the parameter vector with elements λ, σ, $\log A$, β_1, and β_2. A is the multiplicative constant. β_1 and β_2 are respectively the production elasticities of capital and labor.

The maximum likelihood estimators $\hat{\theta}$ of the parameters of the new model have been estimated by applying a combination of three minimization methods. The first method consists of a Monte Carlo searching technique finding the neighborhood of the global minimum. It leads to reasonable starting values for the next minimization stage: simplices are formed which adapt themselves to the local function landscape, and contract in the neighborhood of a minimum [**9**]. The third and last minimization technique is based on a variable metric method by. Davidon [**3**]. The sequential application of these techniques provides for reliable maximum likelihood estimators and in most cases for their standard errors. The computer program supplies in addition the inverse of the Hessian of the loglikelihood function in $\hat{\theta}$.

Since the sample sizes are large and the Cramér-Rao bound therefore is a close approximation of the real covariance matrix, and since — as it will appear — $f(w)$ is nearly normal in most instances and θ should then be "nearly sufficient", the inverse of the Hessian in $\hat{\theta}$ can reasonably be looked upon as a good estimation of the actual covariance matrix of the parameter θ. In those cases when the loglikelihood function is not approximately parabolic in the neighborhood of the maximum, the program supplies an estimate of the standard error (in the shape of a confidence interval which is obviously no longer symmetrical around the estimated value of the parameter) obtained by means of a scanning procedure.[2]

4. RESULTS OF THE EMPIRICAL ANALYSIS

Comparing the elasticities of production of both the average and the frontier Cobb-Douglas model (Table 1), we see that the elasticities of production remain

[2] These methods have been incorporated in a package of programs called MINUIT, by J. James and M. Roos of CERN at Geneva.

practically the same. In spite of the nonnormality of the distribution of the composed error the regression coefficients have been very little affected. The returns to scale remain constant for the industries under consideration. Only the intercept changes in an upward direction. This is as expected. Although in this case the production frontier turns out to be a neutral shift of the average production function, this shall not necessarily be so in other applications.

From the results in Table 2 it follows that the most important part (at least 75% and in most cases much more) of the variance of the convoluted distribution is due to statistical errors and therefore cannot be attributed to inefficiency. Neglecting this fact may be fatal for efficiency measurement. The efficiency parameter λ ranges from 16.684 (Glass products) to 2.4180 (Sugar-works, distillery, beverages) and is highly significant (at the level of 95%) for all the industries.

When we calculate the average efficiency ε_c of the composed error model and compare it with the efficiencies computed according to Richmond's method, (see Table 3) it is easy to see that the latter are improbably low. This is not the case for the composed error efficiency, which looks economically more plausible. As ε_c is a monotone transformation of λ, the industry with the highest λ has also the highest average efficiency, i.e., glass products with efficiency equal to 0.9435. The sector sugar-works, distillery and beverages has the smallest average efficiency namely 0.7074. Although for sugar-works, distillery and beverages the lower extreme for ε_c corresponds to the lower extreme for the Richmond measure ε_R this does not seem to imply a correspondence of a more general nature. On the contrary, the low correlation coefficient (0.254) between the "composed" efficiency measure and the Richmond measure leads to an opposite conclusion as it is not significantly different from zero at the 90% level.

It is of interest to investigate the relationship between the structural parameters and the efficiency in terms of their correlation (see Table 4). There is a slight negative relation between the labor elasticity of production and the efficiency parameter (-0.031). The correlation between the capital elasticity of production and the efficiency parameter is greater and positive (0.430).

Their difference is not significant at an appreciable level. Consequently there is no evidence on the existence of a relationship between the structural parameters and efficiency. It is however evident that the mere consideration of ten industries is insufficient and that further research on this question is required.

As to the sources of the established deviations from the 100% level of productive efficiency, they might partially overlap with the reasons for Leibenstein's X-inefficiency: the contracts for labor are frequently incomplete, and not all inputs are marketed or, if marketed, are not available on equal terms to all buyers (patents) [**9**, (412)]. Only a complete investigation of the market structure and the internal organization and operation of the firms could give a more conclusive answer on this issue.

It may also be interesting to calculate the proportion of firms with productive efficiency at least equal to some given value κ, where κ is some given constant $(0<\kappa<1)$. We have

TABLE 1

COMPARISON BETWEEN THE REGRESSION COEFFICIENT OF THE STANDARD AND THE COMPOSED ERROR COBB-DOUGLAS MODEL (ARRANGED IN ORDER OF SIZE OF THE EFFICIENCY PARAMETER) (a)

Codification nr	Industry (number firms)	Standard Cobb-Douglas model			Composed error Cobb-Douglas model			
		Constant	Elasticities of production		Constant	Elasticities of production		Efficiency Parameter λ
			Labor	Capital		Labor	Capital	
30	Glass products (135)	5.4882	0.7107	0.2330	5.5500	0.7121	0.2320	16.684 (b)
			(0.0671)	(0.0495)	(0.4986)	(0.0664)	(0.0489)	(4.7970; 0.2153)
43	Milk products (320)	3.0690	0.7237	0.3450	3.1680	0.7236	0.3451	12.600 (b)
			(0.0448)	(0.0371)	(0.2671)	(0.0398)	(0.0221)	(4.8413; 0.4444)
47	Textiles (824)	5.0709	0.8221	0.1473	5.1971	0.8222	0.1457	9.9770 (b)
			(0.0284)	(0.0210)	(0.1698)	(0.0281)	(0.0177)	(2.5210; 4.8970)
21	Machine construction, Mechanical tools (953)	5.7350	0.8160	0.1739	5.8433	0.8159	0.1738	9.3698 (b)
			(0.0241)	(0.0179)	(0.1073)	(0.0240)	(0.0150)	(1.8880; 3.134)
28	Electrical machinery (679)	5.7474	0.8335	0.1585	5.9143	0.8311	0.1597	6.1343 (b)
			(0.0264)	(0.0202)	(0.1265)	(0.0261)	(0.0202)	(1.1399; 3.5864)
26	Vehicles, cycles (356)	4.8353	0.7658	0.2381	5.0549	0.7654	0.2380	4.7109
			(0.0390)	(0.0311)	(0.0526)	(0.0374)	(0.0247)	(0.7128)
35	Industrial chemicals (415)	5.5450	0.8272	0.1887	5.8169	0.8249	0.1897	3.7647
			(0.0329)	(0.0214)	(0.2075)	(0.0320)	(0.0210)	(0.3659)
54	Paper (523)	4.8232	0.7204	0.2616	5.1060	0.7187	0.2618	3.7340
			(0.0356)	(0.0220)	(0.1740)	(0.0347)	(0.0216)	(0.4964)
52	Footwear (341)	4.5163	0.8324	0.1684	4.9115	0.8184	0.1733	3.0981 (b)
			(0.0354)	(0.0290)	(0.1843)	(0.0351)	(0.0250)	(0.4576; 2.177)
42	Sugar-works, distillery beverages (282)	5.7851	0.8242	0.1645	6.1966	0.8221	0.1658	2.4120
			(0.0535)	(0.0363)	(0.3859)	(0.0466)	(0.0353)	(0.2204)

(a): the figures between brackets are standard deviations.
(b): the region in the neighborhood of the estimated value of the parameter is not parabolic; left- and righthand errors corresponding to a 68% confidence interval and therefore comparable with standard deviations in the normal case are given.

TABLE 2

COMPARISON BETWEEN THE VARIANCES OF THE ERRORS OF THE STANDARD AND THE COMPOSED ERROR COBB-DOUGLAS MODEL

Industry	Standard Cobb-Douglas model	Composed error	
	σ_ε^2	σ_E^2	$1/\lambda^2$
1) Glass products	0.8118	0.7898	0.0036
2) Milk products	0.6590	0.6438	0.0063
3) Textiles	0.9482	0.9309	0.0100
4) Machine construction Mechnical tools	0.7206	0.7067	0.0114
5) Electrical machinery	0.8456	0.8154	0.0266
6) Vehicles and cycles	0.7885	0.7365	0.0451
7) Industrial chemicals	1.0039	0.9307	0.0706
8) Paper	0.7701	0.6970	0.0717
9) Footwear	0.4794	0.3714	0.1042
10) Sugar works, distillery and beverages	1.0935	0.9115	0.1719

TABLE 3

COMPARISON BETWEEN THE AFRIAT-RICHMOND AND COMPOSED ERROR PRODUCTIVE EFFICIENCY

Industry	Afriat-Richmond Gamma	Composed error
	ε_R	ε_C
1) Glass products	0.5697	0.9435
2) Milk products	0.6333	0.9265
3) Textiles	0.5183	0.9089
4) Machine construction and Mechanical tools	0.6068	0.9036
5) Electrical Machinery	0.5565	0.8595
6) Vehicles and cycles	0.5789	0.8249
7) Industrial chemicals	0.4987	0.7901
8) Paper	0.5864	0.7888
9) Footwear	0.7173	0.7560
10) Sugar works, distillery and beverages	0.4686	0.7069

TABLE 4

CORRELATION MATRIX OF THE PARAMETERS OF THE COMPOSED ERROR MODEL

Parameters	Constant	β_1	β_2	λ	σ
Constant	1.000				
β_1	−0.808	1.000			
β_2	0.556	0.848	1.000		
λ	−0.586	0.430	−0.031	1.000	
σ	0.465	0.237	0.167	0.150	1.000

$$Pr(K > \kappa) = \int_{\kappa}^{1} \lambda k^{(\lambda-1)} dk = 1 - \kappa^{\lambda}.$$

These proportions are given by way of illustration in Table 5 for various probability and efficiency levels.

A remarkable feature of the results which should finally not be left undiscussed is the often slight, and occasionally important, difference between the standard errors on the production elasticities in the standard Cobb-Douglas model on the one hand, and in the composed error Cobb-Douglas model on the other. The latter are systematically smaller than the former. These differences in our opinion shed some light upon the extent of misspecification involved in the estimation of production elasticities by means of a standard Cobb-Douglas model, that is, a model in which no explicit account has been taken of the phenomenon of productive efficiency.

5. CONCLUSIONS

Taking into account the purely statical definition of efficiency as adopted in the composed error model, we arrive at an average sectoral efficiency for the industries at issue which lies between .70 and .94. There is no significant statistical relationship between our efficiency measure and Richmond's. The latter arrives at results which are systematically lower than ours, which is only natural considering his negligence of the statistical error. The evidence from the limited number of industries which were examined was not conclusive in finding a relationship between the efficiency phenomenon and the other structural characteristics of the production process. The aprioristic choice of the exponential distribution for the efficiency variable and, for that matter, of the Cobb-Douglas specification itself, is, of course, debatable. It could, therefore, be useful to compute the sensitivity of our results with respect to this particular choice.

State University Centre Antwerp, Belgium

REFERENCES

[1] AFRIAT, S., "Efficiency Estimation of Production Functions," *International Economic Review*, XIII (October, 1972), 568–598.

[2] AIGNER, D. AND S. CHU, "On Estimating the Industry Production Function," *American Economic Review*, LVIII (September, 1968), 826–839.

[3] DAVIDON, W., "Variable Metric Method for Minimization," *Computer Journal*, X (1968), 406.

[4] FARRELL, M., "The Measurement of Productive Efficiency", *Journal of the Royal Statistical Society-Series A*, CXX (Part III, 1957), 253–290.

[5] FØRSUND, F. AND L. HJALMARSSON, "On the Measurement of Productive Efficiency," *The Swedish Journal of Economics*, LXXVI (June, 1974), 141–154.

[6] GRILICHES, Z. AND V. RINGSTAD, *Economics of Scale and the Form of the Production Function* (Amsterdam: North Holland Publishing Company, 1971).

[7] LEIBENSTEIN, H., "Allocative Efficiency vs. 'X-Efficiency'," *The American Economic Review*,

LVI (June, 1966), 392–415.

[8] MAIRESSE, J., "Comparison of Production Function Estimates on the French and Norwegian Census of Manufacturing Industries," in F. L. Altmann, O. Kyn and H. I. Wagener, eds, *On the Measurement of Factor Productivities. Theoretical Problems and Empirical Results,* Papers and Proceedings of the 2nd Reisenburg Symposium (Göttingen: Vandenbroeck and Ruprecht, 1976).

[9] NELDER, J. AND R. MEAD, "A Simplex Method for Function Minimization", *Computer Journal,* VII (1967), 308–313.

[10] RICHMOND, J., "Estimating the Efficiency of Production," *International Economic Review,* XV (June, 1974), 515–521.

[11] SEITZ, W., "The Measurement of Efficiency Relative to a Frontier Production Function," *American Journal of Agricultural Economics,* LII (November, 1970), 505–511.

[12] TIMMER, C., "Using a Probabilistic Frontier Production Function to Measure Technical Efficiency," *Journal of Political Economy,* LXXIX (July-August, 1971), 776–794.