Nº 88
Noviembre 1983

Documento de Trabajo

ISSN (edición impresa) **0716-7334**
ISSN (edición electrónica) **0717-7593**

Measuring Technical Efficiency: A Comparison of Alternative Methodologies with Census Data

Vittorio Corbo
Jaime de Melo

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
INSTITUTO DE ECONOMIA

Oficina de Publicaciones
Casilla 274-V, Correo 21

MEASURING TECHNICAL EFFICIENCY : A COMPARISON OF ALTERNATIVE METHODOLOGIES WITH CENSUS DATA

*Vittorio Corbo
**Jaime de Melo

Documento de Trabajo N° 88
Santiago, Noviembre 1983

This paper is part of the World Bank Research Project "Liberalization with Stabilization in the Southern Cone" (RPO 672-85). It is a revision of Corbo and de Melo (1983) which was presented at the Latin American Meetings of the Econometric Society in Santiago, Chile in July 1983. We would like to thank Richard Caves, Zvi Griliches, and Yair Mundlak for helpful discussions and Michael Carter, John Page, David Reifschneider and Eric Swanson for helpful comments on a previous draft. We are also grateful to Debbie Bateman, Timothy Condon, Narayana Poduval, Kalayarat Suriyasat, and José Miguel Sánchez for research assistance and Raúl Galleguillos and Eric Swanson for computational support.

* Pontificia Universidad Católica de Chile.
** Banco Mundial.

I. INTRODUCTION

Following Farrell's pioneering work (1957), the development and refinements of the economic and statistical foundations of production and cost frontiers has progressed rapidly in recent years. Some of the more recent developments, in particular the formulation of stochastic (or "composed error") frontiers, attempt to recognize that the measurement of technical and allocative efficiency is fraught with difficulties: (a) the firm operates in a world where its performance is likely to be affected by events outside its control; (b) the empirical relationship contains a good deal of statistical noise in the form of measurement error, omitted variables and so on. Yet the illustration of new techniques, including the comparison of different approaches, has generally taken place on a limited number of data sets: the Swedish milk industry [Van der Broeck et al (1980)] and the U.S. steam power generating plants [Kopp and Smith (1981)]. The usefulness of these techniques would be enhanced as tools for policy analysis if they yielded plausible (and hopefully similar) results on a wider range of data sets including manufacturing census data where there is great interest in measuring and understanding firm efficiency within and across sectors.

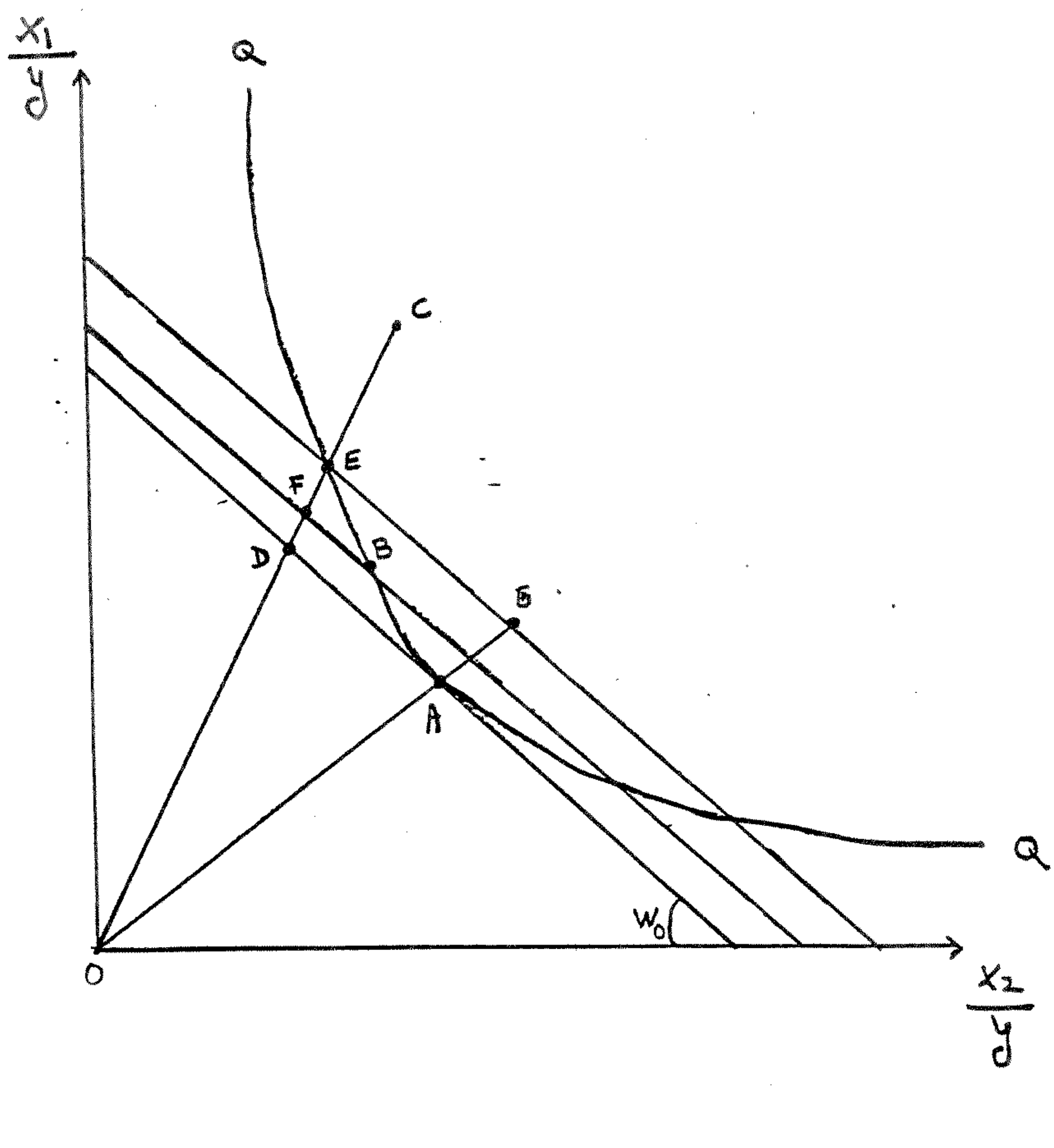
The purpose of this paper is to provide a comparison of these alternative approaches on such a data set. To our knowledge with the exception of the two above-mentioned studies, no one has yet undertaken such a comparative performance of alternative methods. The comparison is undertaken with the 1967 Chilean manufacturing census where data is gathered at the four-digit ISIC level for all establishments employing more than five workers.

Comparison takes place over the range of production frontiers (deterministic parametric, deterministic statistical and stochastic) which have appeared in the literature since Farrell's (1957) pioneering work. These recent developments allow both the estimation of statistically more robust frontiers and the retrieval of indices of efficiency at the firm level (Section II and Appendix I). We start in Section II with definitions and interpretations of measures of efficiency upon which the analysis is based. Section III discusses various ways to estimate production fronters and presents the alternative methods that are subjected to comparison. Section IV discusses the choice of functional form and estimation procedures for the production frontiers; section V presents efficiency estimates under different models and their interpretation. Section VI reports the results of a comparison of estimates based on the different models. Conclusions follow in Section VII.

II. DEFINITIONS AND INTERPRETATIONS OF EFFICIENCY

Consider Figure 1 where a sample of firms is depicted for an industry producing a single output y, with two inputs $X \equiv (X_1, X_2)$ available at fixed prices $W = (W_1, W_2)$. The output can be sold at a fixed price, P. The frontier production function can be characterized by the unit isoquant QQ, provided that technology can be described by a linear homogeneous production function so that we can express the production function as $1 = f\left(\frac{X_1}{y}, \frac{X_2}{y}\right)$. Farrell distinguished between technical and allocative efficiency. These concepts are discussed with the help of Figure 1. For the constant returns to scale case, a firm is technically efficient if it chooses an input mix on the unit isoquant. A firm is allocatively efficient if the marginal rate of substitution between the two inputs is equal to the factor price ratio. Given

Figure 1



the factor prices W_0, firm A is the only firm in the industry which is both technically and allocatively efficient; firm B is technically efficient and is allocatively inefficient, for it is using the wrong factor proportions and a relative index of its inefficiency, derived from unit cost comparisons, is given by E_B = OD/OF. Finally, firm C is both technically and allocatively inefficient and its overall inefficiency OD/OC can be decomposed into an allocative and a technical component.

$$E_C = \frac{OD}{OC} = \frac{OD}{OE} \; \frac{OE}{OC}$$

Total inefficiency = allocative inefficiency x technical inefficiency.

The above analysis does not consider the optimality of the level of production, since the scale of production is indeterminate in the case of constant returns to scale. However if the technology is non constant returns to scale (homogenous or not), then the scale of production will be optimal if and only if at the chosen levl of output, price is equal to marginal cost. 1/ A firm is on the cost frontier if it is both technically and allocatively efficient. Finally, a firm is said to be scale efficient if it chooses a profit maximizing level of production.

Since in our comparison we only rely on information on output and input quantities we cannot distinguish allocative from technical or scale

---

1/ For a general (contant returns or not) well-behaved production function $y = f(x)$, a firm is technically efficient if the observed production and input combination (x_o, y_o) satisfies $y_o = f(x_o)$. Technical inefficiency arises when $y_o < f(x_o)$. The definition of allocative efficiency is not altered.

inefficiency. Thus the figures below should be interpreted as measures of technical inefficiency with respect to the industry production frontier. To illustrate, refer back to figure 1: both firms A and E will appear as efficient, and firm G will appear as inefficient even though it is allocatively efficient and it achieves the same degree of overall efficiency as firm E.

There are several interpretations to the scatter of points in figure 1. One explanation is that firms do not have access to the same technology, in which case there is no reason to investigate differences in efficiency. If observations could be grouped by technology class, then efficiency could be studied within classes. Along the same lines one can attribute the scatter to the sample containing firms with equipment of different vintage. In this case the relevant efficiency frontier is different for observations belonging to different vintages. Observations should then be grouped by vintages and comparisons should be made within a vintage. This is clearly an important consideration in a world where the technological structure of manufacturing industries has different substitution possibilities before and after investments in new techniques. A third interpretation is that all firms face the same technology but that some are more successful in using it than others. This corresponds to the full-frontier (or deterministic frontier) approach discussed below.

A fourth interpretation is that all firms face the same technology up to a random factor that takes into account the effects on production of measurement errors in the output variable and other random shocks outside the firm's control. Thus, the resulting production frontier is stochastic and departure from this frontier reflects technical inefficiency. 1/ This corresponds to the stochastic frontier concept discussed below.

The measures of efficiency reported in this paper refer to a single point in time and are therefore static. Given our selection of inputs in the production function, what is being measured is the technical efficiency of: physical plant and equipment, unskilled, and skilled labor in producing output. Therefore the resulting measure of efficiency is a multiple-factor index. 2/

III. ESTIMATION OF PRODUCTION FRONTIERS 3/

The estimation of production frontiers has proceeded along two general paths: full-frontiers which force all observations to be on or below the frontier and hence where all deviation from the frontier is attributed to inefficiency; and stochastic frontiers where deviation from the frontier is decomposed into a random component reflecting measurement error and statistical noise, and a component reflecting inefficiency. The advantage of the stochastic frontier approach is that it incorporates the traditional

---

1/ See Forsund et al (1980, pp. 21-23) and Stigler (1976) for further discussion on the interpretation of inefficiency.

2/ See Kopp (1981) who introduces single-factor Farrell efficiency measures for full frontiers. This corresponds to the case where the technical efficiency of a subset of factor inputs is fixed by *ex-ante* decisions. Then single-factor measures of efficiency may be more appropriate measures of a plant's *ex-post* efficiency since they do not penalize a production organization for *ex-ante* mistakes. And, in the case where data on factor prices are available, the firm is not penalized by its inability to adjust.

3/ This section draws on Forsund et al (1980) who also discuss cost and profit frontiers.

random error of the regression. In this case the random error -- besides capturing the effect of unimportant left out variables and errors of measurement in the dependent variable -- would also capture the effect of random breakdown on input supply channels not correlated with the error of the regression. The measures reported in this paper are based on both full frontiers and stochastic frontiers. Below we describe briefly both approaches; models and estimation techniques are presented in the appendix.

A. FULL-FRONTIERS

In Farrell's work, the basic procedure was to construct the efficient unit isoquant from the observed input-output ratios by linear programming techniques. Although constant returns to scale (CRTS) was assumed, the major advantage of this approach is that it imposes no functional form on the data. Furthermore, Farrell's approach has been lately extended to allow for non-homothetic and inhomogenous functions. 1/ Thus one approximates with a minimum of restrictions the unknown frontier without particular functional form restrictions. However, only in the case of CRTS does this procedure provide enough information to determine a production function. The estimation is termed non-parametric in the literature in the sense that the model is not based on any explicit model of the frontier. This is the approach followed by Meller (1976) on the same data set used in this paper.

The next step in the estimation of production frontiers was to move to a parametric full-frontier where a functional form is imposed on the production function and the elements of the parameter vector describing the

---

1/ See Kopp (1981) for the weak restrictions on the functional form and on the derivation of efficiency indexes.

production function are estimated by programming [Aigner and Chu (1968)] or by statistical [Richmond (1974), Greene (1980a)] techniques. This is one of the approaches used in this paper. The drawback of these techniques is that, like the Farrell technique, they are extremely sensitive to outliers; and hence if the outliers reflect measurement errors they will heavily distort the estimated frontier and the efficiency measures derived from it.

The advantage of estimating full frontiers by statistical rather than programming techniques is that if the distribution of technical inefficiency is properly specified, then under certain regularity conditions [see Greene (1980a)] one can derive maximum likelihood estimates with their usual desirable statistical properties. This gain in confidence about the statistical properties of the parameters is made at the cost of imposing a particular distribution of technical inefficiency which, as discussed below, introduces another form of sensitivity to the results. 1/

Next consider the relation between the "average" function and the "frontier" function. In the standard estimation of production models it is usually assumed that $y = f(x)e^{\varepsilon}$, where y is output, x is the input vector and ε is a random variable distributed in the interval $(-\infty, \infty)$. The estimated model in this case is an "average" production model. The "frontier" function is given by: $y = f(x)\ e^{-u}$, where $u \geq 0$ is a random variable, which is generated by independently identically distributed (iid) statistical drawings from some distribution. The vector u represents inefficiency.

1/ Another difference is that in the non-statistical case, maximality describing the frontier is over all the points in the sample. Hence one obtains a best-practice frontier, whereas in the statistical case, maximality takes place over all possible sample points given technology so one obtains an absolute frontier [(Forsund, et al (p. 20)].

Under the conditions specified above the average function is conceptually identical to the frontier except for the realized value of the multiplicative efficiency term. 1/ In contrast, in the non-statistical case, the unknown frontier is estimated directly rather than in relation to the average. Finally, until recently, the major advantage of full frontier models over the stochastic model presented below was that they provided efficiency indexes for each firm. However, lately Jondrow et al (1982) have derived estimates of expected efficiency at the firm level for the stochastic frontier model.

B. STOCHASTIC FRONTIERS

The stochastic frontiers models is given by:

$y = f(x)e^{v}$.	e^{-u}	$u \geq 0$,
stochastic frontier	inefficiency term	v is a random variable that takes values in the range $(-\infty, +\infty)$

From the estimation of this model one gets a set of efficiency values such as: (1) an average efficiency index for the sector; (2) an expected efficiency index for each observation relative to the stochastic frontier; and (3) a measure, $\lambda = \sigma_u/\sigma_v$, indicating whether most of the variance from the frontier is due to randomness or to inefficiency. 2/

1/ As explained below, it is this relationship which allows us to obtain an estimate of the frontier by correcting the constant term from the OLS estimate of the average frontier (hence the name COLS).

2/ See the appendix for derivations.

As mentioned above, we estimate both full frontiers and stochastic frontiers. The great advantage of the stochastic frontier for our data set is that it allows for randomness and measurement errors in the dependent variable. One apparent disadvantage of the stochastic frontier is that the correction factor required to obtain a consistent estimate of the efficiency term draws on an estimate of the third central moment of the composite error $v-u$. If the model is correct, then the population value of this third moment is negative. However, if the sample value of this third moment is positive then the estimation procedure breaks down (this corresponds to the type I error in Olson et al (1980))

In effect, as shown in Appendix I, a consistent estimate of the constant requires the use of a consistent estimate of $E[u]$ which can be obtained by the method of moments. If we assume that v is $N(0,\sigma_v^2)$ then the second and third central moments of the distribution of the composite error $\varepsilon=v-u$ provide the information to estimate $E(u)$ under alternative assumptions about the distribution of u. Regardless of the distributional assumption about u, the population value of the third central moment of $\varepsilon(\mu_3')$ is always negative. However, there is nothing guaranteeing that the sample estimation, $\hat{\mu}_3'$ which is a consistent estimation of μ_3', will be negative in which case we have what Olson et al (1980) refer to as a Type I error. Likewise, if the sample estimate $\hat{\sigma}_v^2$ is negative the estimation also breaks down in what Olson et al (1980) refer to as a Type II error. In that case λ is meaningless.

If the wrong sign of this moment is due to a few outliers arising from measurement errors then the correctly specified model cannot be estimated. This drawback is not apparent with the deterministic statistical full-frontiers, although outliers would create a bias of unknown consequences. We also present measures based on the programming estimation technique since, unlike the other methods, it provides a direct estimate of the frontier and does not force the unknown frontier to be approximated by the average frontier.

IV. DATA, SPECIFICATION OF FUNCTIONAL FORM, AND ESTIMATION

As shown in Appendix I, for both the statistical full frontier case and the stochastic frontier case, OLSQ estimation of a model linear in the parameters provides BLUE estimators of all the coefficients except the constant. This justifies drawing on previous work on "average frontiers" for the selection of the appropriate functional form. Of course, OLSQ is BLUE only if the model is correct and only if the left-hand side variables are exogenous. Exogenous right hand side variables can be obtained in models where the firm maximizes expected profits [Zellner, Kmenta and Dreze (1966)].

The specification of the production frontier draws on previous work by Corbo and Meller (1979b and 1982), where the technology of Chilean manufacturing sectors was studied in detail using the same data set. The data set provides cross-section data for individual establishments within each sector. The data covers 44 four-digit ISIC manufacturing sectors. The output variable is value-added (V) and the inputs are: number of man days (L), skill units (S) and the value of fixed assets (K), all variables used by Corbo and Meller (1982) 1/

1/ In another paper [Corbo and Meller (1979a)] inputs were defined as blue collar workers, white collar workers, and value of fixed assets.

The precise definitions of the input variables used are as follows:

L = average annual number of man days. It is measured by the sum of production workers, blue-collar workers in auxiliary activities, white-collar workers, and entrepreneurs times the number of days worked by the establishment. 1/

S = skill-days units. It is measured by the average annual number of blue-collar days equivalent minus 1. 2/ The number of blue collar days equivalent is measured by the ratio of the total wage payments, plus an imputation for entrepreneurs, to the minimum wage rate of the whole industrial sector. 3/

K = Book value of machinery at 1967 prices less accumulated depreciation. 4/

1/ As reported in Corbo and Meller (1979b) preliminary statistical tests and regressions were performed using the number of annual man-hours worked by production workers; however, this variable turned out to be highly unreliable. The only other available measurement of a flow variable for labor is the number of workers times the number of days worked by an establishment during the year. The use of this variable implies the following for all establishments of the same industry: workers work the same number of hours; absenteeism and part-time workers are equally distributed (part-time workers are negligible in Chilean manufacturing); and the number of shifts worked is the same (most Chilean manufacturing establishments work only one shift).

2/ The implicit assumption here is that each worker is composed of two parts: body and skills and that wage differentials are due to quality differences. See Griliches (1967).

3/ The wage rate of entrepreneurs is assumed to be two and a half times the average wage rate of white-collar workers within a given establishment. To minimize the possibility of measurement error, the minimum wage rate of the whole industrial sector is computed as the simple average of the ten lowest wage rates of blue-collar workers observed in the census.

4/ In a persistently inflationary economy like Chile's, the use of book values to measure the capital service factor, (besides the traditional limitations of ignoring differences in capacity utilization, accounting procedures, and depreciation rates) leads to an underestimation of the capital factor of the older establishments, thereby exaggerating their technical efficiency. However, the use of an available alternative measure would not greatly affect our results. Meller (1976) used a flow measure of capital services instead of the value of the stock. The

Y = Gross value added at 1967 prices.

The units of L, S, K, and Y are chosen in such a way that for a given industry i, the mean of each one of them is equal to one.

In that paper the technology was represented by a translog production function. The translog function was estimated directly and then more restricted nested models were estimated testing for CRTS and input separability. When testing for CRTS, in only three cases out of forty-four was the null hypothesis rejected at the 1 percent level. These sectors were bakery products (ISIC 3112), wearing apparel except footwear (ISIC 3220), and cement for construction (ISIC 3693).

For the forty-one CRTS sectors, further tests for global separability (a Cobb-Douglas technology) were performed. For thirty-five out of the forty-one sectors the Cobb-Douglas technology could not be rejected [See Corbo and Meller (1982, Table 5.2)]. Thus, there were six CRTS sectors for which the Cobb-Douglas technology was rejected: spinning, weaving, and finishing textiles (3211); sawmills, planing, and other wood mills (3311); printing, publishing, and allied industies (3420); furniture and fixtures primarily of

---

capital service variable was defined as $K = .10\ K_M + .03\ K_B + .20\ K_V + .10\ (K_M + K_B + K_V + K_I)$, where K_M, K_B, K_V, and K_I are the book values of machinery, buildings, vehicles, and inventory goods. Geometric depreciation rates of .10, .03, and .20 were used for machinery, buildings, and vehicles, and a 10 percent real interest rate was used as the cost of capital. The simple correlation between the capital service measure and the book value of machinery measure was above .95 in sixteen out of the twenty-one industrial sectors considered in that study, with the smallest correlation coefficient being .823. Similar high correlation coefficients were obtained with standard alternative capital measures like electricity consumed by the establishment measured in kilowatt hours and installed capacity of the production machinery measured in horsepower.

metal (3812); special industrial machinery (3824); and machinery and equipment not elsewhere classified (3829). For these six sectors the data indicate that a CRTS translog function is appropriate. Finally, for the three sectors for which the CRTS hypothesis was rejected, tests were performed for a Cobb-Douglas non CRTS technology. In all three cases the null hypothesis could not be rejected at the 1% significance level.

For the results reported here, we have relied on a CRTS Cobb-Douglas functional form. Thus for our data set, with the few exceptions mentioned above, the Cobb-Douglas technology provides an appropriate representation of the "average" sectoral production function.

A final issue is the proper measurement of the output variable in the production function. If it is believed that the proper definition should be at world prices and if there is a systematic variation in protection across firms within a sector, then the estimated coefficients will be biased. Appendix II derives the bias for the Cobb-Douglas case.

We briefly illustrate our estimation procedure with the more general case of the stochastic frontier model. The model is linear in parameters and given by:

$$Y = \beta_0 + X\beta_1 + \varepsilon \tag{1}$$

where $\varepsilon = v - u$

v iid $N(0, \sigma^2 v)$
u iid half-normal or exponential
$u \geq 0$

and u and v are independently distributed. If the columns of X are exogenous, then the OLSQ estimator $\hat{\beta}_1$ of β_1 is BLUE. However, the OLS estimator $\hat{\beta}_0$ of the constant term is biased. The biases arise from the fact that $E(\varepsilon) \neq 0$. By adding and subtracting $E(\varepsilon)$ to the right hand side of (1), the transformed model has a new random error given by $\varepsilon - E(\varepsilon)$ and a constant equal to $\beta_0 + E(\varepsilon)$. In this transformed model, the expected value of the error term is zero. Therefore OLSQ provides BLU estimates of all coefficients including the new constant $\beta_0 + E(\varepsilon)$. A consistent estimate of β_0 is given by subtracting a consistent estimates of $E(\varepsilon)$ from the OLSQ estimator of $\beta_0 + E(\varepsilon)$. (See Appendix I for details). This estimation procedure is known as corrected least squares or COLS.

The above model could also be estimated by maximum likelihood (ML) which yields different results from OLS. As is well-known ML only has asymptotic properties. Montecarlo experiments for a similar model (Olson et al 1980) indicate that for sample sizes below 400, COLS is superior or equal to MLE for all parameters. 1/ Estimation performed by ML, for a few sectors, not reported in this paper, yielded results very similar to the COLS estimates provided below. Based on these findings and cost considerations we estimate all sectors by COLS.

V. EFFICIENCY ESTIMATES AND INTERPRETATION

This section presents efficiency results obtained using both deterministic and stochastic models. Results will be interpreted both at the

---

1/ As with all Montecarlo experiments, these results could be sensitive to the particular values selected for the inputs and the parameters of the distribution of the random errors.

sector and within each 4-digit ISIC sector.

Table 1 lays out the main characteristics of individual sectors. We present variables related to size (gross value of production), technology (capital/labor ratio), trade orientation (exports/gross production: imports/(imports + gross production)); industrial concentration (Hirschman-Herfindahl index, HD). Sectors have been grouped into four categories: exportables, import-competing, non-import competing and non-tradables. Exportables are those sectors for which apparent consumption (production + imports - exports), C, is less than production (X). Import competing sectors are sectors where $0.01 < (\frac{C-X}{C}) < 0.75$ non-import-competing sectors are sectors where $(\frac{C-X}{C}) > 0.75$. Finally non tradables are sectors where

$$0 < (\frac{C-X}{C}) < 0.01. \ \underline{1/}$$

As can be seen from the resulting distribution of sectors in the above categories, there is a clustering of sectors in the import competing group. This is not surprising given the bias against exporting activities of the then existing trade regime. 2/ As expected, import-competing sectors did not appear to be exposed to much foreign competition as indicated by the low import penetration ratios in column 5.

Altogether, there are 44 4-digit sectors which had enough observations (i.e., number of firms) for statistical estimation. As can be

---

1/ This classification is inspired by Krueger, et al. (1981, ch. 1) and Corbo and Meller (1981).

2/ The biases of this trade regime are enalyzed in Behrman (1976) and Corbo and Meller (1981).

Table 1

Industry (ISIC) / Trade Classification	N° OBS.	Gross Value of Production (Thousands of Pesos)	Capital/Labor	Exports/Production	IMP/IMP+PROD	HO
Exportables						
3111 Slaughtering, preparing and preserving meats	100	1,521.841	7.239	0.094	0.062	0.0204
3113 Canning and preserving of fruits and vegetables	32	144.071	18.850	0.063	0.037	0.0548
3114 Canning, preserving, and processing of fish, crustaceans, and similar foods.	37	147.008	21.606	0.060	0.0003	0.0578
3132 Wine Industries	70	472.323	30.770	0.017	0.0002	0.0106
3311 Sawmills, planing, and other wood mills	252	525.386	7.076	0.023	0.003	0.0059
3411 Pulp, paper, and paperboard	19	334.716	74.205	0.422	0.102	0.1066
Import Competing						
3112 Dairy Products	46	393.208	29.949	0.0	0.158	0.0482
3115 Vegetable and animal oils and fats	34	313.541	45.671	0.018	0.172	0.0482
3119 Cocoa, chocolate, and sugar confectionery	26	131.132	16.495	0.003	0.051	0.0132
3121 Manufacture of food products n.e.c.	39	339.731	26.985	0.008	0.136	0.3787
3131 Distilling, rectifying, and blending of spirits	25	167.459	16.187	0.0002	0.016	0.0637
3211 Spinning, weaving, and finishing of textiles	232	1.331.919	21.878	0.037	0.114	0.0243
3212 Made-up textile goods except wearing apparel	22	27.623	4.933	0.005	0.715	0.0391
3220 Wearing apparel, except footwear	239	626.913	4.879	0.00006	0.044	0.0137
3233 Products of leather and leather substitutes except footwear	30	33.223	3.520	0.0009	0.119	0.0391
3240 Footwear, except vulcanized or molded rubber or plastic footwear	138	389.540	4.772	0.001	0.25	0.0360
3312 Wooden and cane containers and small caneware	27	20.294	5.679	0.0007	0.145	0.0379
3320 Furniture and fixtures, except primarely of metal	132	176.327	3.756	0.002	0.248	0.0248
3420 Printing, publishing, and allied industries	149	426.721	17.381	0.031	0.179	0.0294
3511 Basic industrial inorganic chemicals except fertilizers	32	177.217	44.512	0.227	0.317	0.0437
3521 Paints, varnishes, and lacquers	25	134.480	14.318	0.0	0.056	0.0685
3522 Drugs and medicines	45	260.230	21.049	0.0005	0.242	0.0435
3523 Soap and cleaning products, perfumes, cosmectics and other toilet preparations	52	283.449	20.386	0.00005	0.029	0.0162
3529 Chemical Products n.e.c.	37	144.332	14.288	0.021	0.261	0.0899
3559 Rubber Products n.e.c.	24	87.254	9.834	0.034	0.453	0.2142
3560 Plastic Products n.e.c.	77	222.107	14.806	0.058	0.304	0.0456
3620 Glass and glass Products	32	137.317	12.450	0.004	0.163	0.0992
3710 Iron and steel basic industries	42	703.711	54.584	0.074	0.293	0.0916
3812 Furniture and Fixtures primeraly of metal	47	105.503	8.169	0.0	0.027	0.0785
3813 Structural metal products	76	165.352	10.156	0.009	0.549	0.0366
3814 Metal containers and metal housewares	56	280.966	11.386	0.00007	0.039	0.0608
3815 Cable, wire, and their products	31	100.530	13.122	0.0	0.031	0.1279
3819 Fabricated metal products except machinery and equipment n.e.c.	86	120.526	7.499	0.001	0.066	0.0286
3822 Agricultural Machinery and Equipment	30	45.850	11.811	0.0006	0.346	0.0538
3829 Machinery and equipment except electrical n.e.c.	89	436.586	14.218	0.011	0.361	0.0753
3839 Electrical apparatus and supplies n.e.c.	19	148.450	18.564	0.076	0.336	0.1037
3843 Motor vehicles	73	619.223	10.926	0.0006	0.354	0.0733
Non Import Competing						
3811 Cutlery, hand tools, and general hardware	26	41.754	7.653	0.019	0.798	0.0738
3824 Special industrial machinery and equipment except metal and woodworking machinery	19	20.595	8.854	0.021	0.918	0.0683
Non Tradables						
3117 Manufacture of Bakery Products	293	602.864	6.426	0.00001	0.001	0.0115
3213 Knitting mills	145	430.431	13.346	0.00003	0.004	0.0262
3231 Tanneries and leather finishing	57	210.875	14.350	0.0	0.00006	0.0437
3841 Shipbuilding and Repairing	19	93.168	6.912	0.0	0.007	0.4751

HO: Is the Herfindahl Index of Industrial Concentration taken from Heller and Swinburn (1975) seen from column 1 the number of firms per sector ranges from a minimum of 19 establishments up to a maximum of 252 establishments. Only establishments with a minimum of five employees are included in the census. Since Meller (1976) found substantial differences in technology between establishments employing less than ten employes and establishment employing more than ten employ, we exclude establishments with less than ten employees. Table 2 gives the measures of efficiency according to the trade classification discussed above.

There are six measures corresponding to four deterministic and two stochastic models. Comparing the measures for a given sector (across columns) one notes several systematic differences. Among the statistical deterministic models, the lowest levels of average efficiency are necessarily those from the "distribution free" model (see Appendix I). In all cases the estimates under the assumption of a Gamma distribution for the error structure yield higher values than those obtained with an exponential distribution. This is so because the differences in the estimated expected efficiency is only a function of $\hat{\sigma}$ which is estimated by the standard error of the regression). It is easy to show that for $0 < \hat{\sigma} < 1$ ($\hat{\sigma} > 1$) the expected efficiency from the Gamma distribution is higher (lower) than the expected efficiency computed from the exponential distribution. For our data set, in all cases the standard error of the regression (which is a consistent estimate of the parameter (ϕ) of the Gamma distribution (see Appendix I) is less than unity. As shown by Richmond (1974) $\phi < 1$ corresponds to the case where the mode of the distribution is at $u = 0$ (u is the non-negative error corresponding to inefficiency) which in turn implies a distribution of inefficiency across firms such that most firms are efficient, a result similar to the one found by

Table 2

EFFICIENCY MEASURES

INDUSTRY (ISIC) TRADE CLASSIFICATION	DETERMINISTIC MODELS				STOCHASTIC MODELS			
	LP	E Distribut. Free Av. Eff.	E(e^u) Gamma	E(e^u) EXP	E (e^u) Half Normal	E(e^u) EXP	λ^HN	λ^EXP
EXPORTABLES								
3111 Slaughtering, preparing and preserving meats	0.294	0.226	0.737	0.601	0.680	0.794	0.956	0.425
3113 Canning and preserving of fruits and vegetables	0.529	0.513	0.809	0.643	nc	nc	nc	nc
3114 Canning, preserving and processing of fish, crustaceans, and similar foods	0.504	0.427	0.840	0.660	0.692	0.802	1.315	0.566
3132 Wine Industries	0.532	0.318	0.724	0.594	0.541	0.691	3.431	0.865
3311 Sawmills, planing and other wood mills	0.370	0.310	0.800	0.638	0.652	0.774	1.413	0.599
3411 Pulp, paper and paperboard	0.668	0.548	0.932	0.758	nc	nc	nc	nc
IMPORT COMPETING								
3112 Dairy Products	0.379	0.197	0.705	0.585	0.742	0.837	0.613	0.286
3115 Vegetable and animal oils and fats	0.322	0.215	0.737	0.601	nc	nc	nc	nc
3119 Cocoa, chocolate and sugar confectionery	0.393	0.513	0.911	0.732	0.892	0.934	0.419	0.198
3121 Manufacture of food products n.e.c.	0.302	0.125	0.743	0.604	nc	nc	nc	nc
3131 Distilling, rectifying, and blending of spirits	0.367	0.327	0.696	0.580	0.732	0.830	0.637	0.296
3211 Spinning, weaving and finishing of textiles	0.225	0.117	0.850	0.674	nc	nc	nc	nc
3212 Made-up textile goods except wearing apparel	0.563	0.455	0.917	0.738	nc	nc	nc	nc
3220 Wearing apparel, except footwear	0.378	0.346	0.847	0.672	0.649	0.772	1.963	0.761
3233 Products of leather and leather substitutes except footwear	0.528	0.390	0.880	0.700	0.898	0.938	0.330	0.157
3240 Footwear, except vulcanized or molded rubber or plastic footwear	0.367	0.210	0.880	0.699	nc	nc	nc	nc
3312 Wooden and cane containers and small caneware	0.514	0.339	0.868	0.688	nc	nc	nc	nc
3320 Furniture and fixtures, except primarely of metal	nc	0.268	0.853	0.675	nc	nc	nc	nc
3420 Printing, publishing and allied industries	0.334	0.246	0.856	0.678	nc	nc	nc	nc
3511 Basic industrial inorganic chemicals except fertilisers	0.390	0.113	0.649	0.559	nc	nc	nc	nc
3521 Paints, varnishes and lacquers	0.508	0.505	0.861	0.683	0.754	0.844	0.959	0.432
3522 Drugs and medicines	0.443	0.321	0.896	0.715	nc	nc	nc	nc
3523 Soap and cleaning products, perfumes, cosmetics	0.344	0.268	0.815	0.648	0.627	0.756	1.87	0.736
3529 Chemical products n.e.c.	0.350	0.266	0.807	0.642	nc	nc	nc	nc
3559 Rubber products n.e.c. and other toilet preparations	0.637	0.460	0.896	0.715	0.676	0.791	2.543	0.887
3560 Plastic products n.e.c.	0.332	0.283	0.828	0.657	nc	nc	nc	nc
3620 Glass and glass products	0.517	0.408	0.859	0.681	0.799	0.874	0.697	0.322
3710 Iron and steel basic industries	0.435	0.331	0.840	0.666	0.716	0.818	1.117	0.493
3812 Furniture and fixtures primarely of metal	0.453	0.238	0.840	0.666	nc	nc	nc	nc
3813 Structural metal products	0.311	0.223	0.820	0.652	nc	nc	nc	nc
3814 Metal containers and metal housewares	0.430	0.404	0.874	0.694	0.775	0.859	0.885	0.402
3818 Cable, wire and their products	0.543	0.468	0.841	0.664	nc	0.722	nc	1.20
3819 Fabricated metal products except machinery and equipment n.e.c.	0.470	0.357	0.881	0.701	nc	nc	nc	nc
3822 Agricultural Machinery and Equipment	0.445	0.468	0.874	0.694	0.751	0.843	1.052	0.469
3829 Machinery and equipment except electrical n.e.c.	0.372	0.1867	0.824	0.655	nc	nc	nc	nc
3839 Electrical apparatus and supplies n.e.c.	0.474	0.614	0.886	0.705	nc	nc	nc	nc
3843 Motor vehicles	0.312	0.229	0.779	0.625	nc	nc	nc	nc
NON IMPORT COMPETING								
3811 Cutlery, hand tools, and general harware	0.598	0.562	0.859	0.681	nc	0.730	nc	1.273
3824 Special industrial machinery and equipment except metal and woodworking machinery	0.571	0.418	0.806	0.642	0.680	0.793	1.21	0.528
NON TRADABLES								
3117 Manufacture of Bakery Products	0.272	0.068	0.848	0.672	nc	nc	nc	nc
3213 Knitting mills	0.420	0.306	0.883	0.702	0.682	0.795	1.983	0.765
3231 Tanneries and leather finishing	0.286	0.124	0.838	0.664	nc	nc	nc	nc
3841 Shipbuilding and Repairing	0.474	0.530	0.929	0.754	nc	nc	nc	nc

NOTE: For the definitions of the efficiency measures see Appendix 1.

NC = Not computed because under the assumed distribution of u the estimated variance of u was negative.

Richmond for Norwegian manufacturing but surprising for our sample. However, the value of $\hat{\sigma}$ is not independent of the unit of measurement of the independent variable so that one cannot draw any conclusions from this comparison. 1/

As shown by Schmidt (1976), the LP procedure is equivalent to the deterministic statistical model with an exponential error structure. Thus, the difference between the two measures (cols 1 and 4 in Table 2) can be viewed as a comparison of two consistent estimation methods for the statistical deterministic frontier (COLS and ML).

Finally, in the stochastic case, expected efficiency under a half-normal assumption for the error structure is always lower than under the assumption of an exponential distribution for the inefficiency error. As can be seen from Appendix I this result is also related to the value of the standard error of the regression which is the only parameter that enters in the calculation of the expected efficiency in both cases.

A general pattern of the results under all models is a low level of measured efficiency. Table 3 compares our results with those obtained by other authors. That table presents under alternative Cobb-Douglas model specifications, the minimum and maximum sectoral values of efficiency. In general, sectors with high efficiency are those producing the more homogenous products and the opposite is the case for sectors with the lowest level of efficiency. However, though functional forms are the same, direct comparison with other studies is difficult. First we have different definitions for the

---

1/ This point seems to have been overlooked in the literature. See Richmond (p. 519) and Forsund et al (pp. 12-13).

input and output variables. Second our results are at a much more disaggregated level and thus we are able to control much better for product heterogeneity.

In interpreting the low values of efficiency for the deterministic statistical stochastic models, one should keep in mind possible biases in our estimates. As shown in Appendix II, if variations in effectives rates of protection across firms within a sector are positively (negatively) correlated with capital (labor) uses then not only will the elasticity estimates be biased but the estimated variances of the error will be upward biased. In the case of the exponential, this will result in a downward-biased estimate for the expected efficiency.

Another source of concern is the specification of the error structure. As can be seen from Table 2, we could not compute the stochastic half-normal frontier in 25 out of 43 cases and the stochastic exponential in 23 out of 43 cases. 1/ As explained in the appendix, this result arose because, for the remaining sectors, the sample estimate of the third moment of the composite error had the wrong sign. In that case the estimated variance of u is negative and the estimation procedure collapses. This corresponds to the "type I error" (see Olson et al., 1980, p. 70) and is likely to occur when $\lambda \to 0$ i.e., when most of the variance on the frontier is due to randomness rather than inefficiency. Since the mode of the half-normal (and exponential) distributions describing the inefficiency structure is also at $u = 0$,

---

1/ In the half normal case, for two sectors (3811 and 3815), the estimated variance of the two sided error turned out negative. (This corresponds to the type II error in Olson et al., 1980, p. 70.)

Table 3

COMPARISON OF EFFICIENCY MEASURES WITH OTHER STUDIES

Authors	N° of Sectors	MODEL: Deterministic Statistical	Deterministic Non Parametric	Stochastic
Menseen Van der Broeck (Mfg: 10 sectors at 3 digit level of French Census 1962)	10	b/ min = 0,468 (sugar works, distillery and beverage) Max = 0,717 (Footwear)		e/ min = 0,708 (sugar works, distillery and beverage) max = 0,944 (Glass products) Estimation: ML
Richmond (Norwegian 1963 census of mining and manufacturing)	27	b/ min = 0,748 (industrial chemicals) max = 0,957 (Steel foundries)		
Tyler and Lee (2-digit Colombian manufacturing 1974: small and medium size)	5	-		d/ min = 0,554 (apparel) max = 0,984 (Furniture) Estimation: ML
Tyler: Brazilian Plastic and Steel balance Sheet data (1971)	2		Plastic = 0,48 Steel = 0,62	
Corbo/de Melo 1967 Chilean Manufacturing Census 4-digit ISIC	43 deterministic 18 (stochastic Half normal) 20 stochastic exponential	a/ min = 0,117 (spinning, wearing and finishing of textiles) max = 0,562 (cutlery, hand tools and general hardware). b,c/ min = 0,649 (Inorganic chemicals except fertilizers) max = 0,932 (pulp, paper and paperboard)	min = 0,225 (spinning, weawing and finishing of textiles) max = 0,668 (pulp, paper and paperboard)	d/ min = 0,541 (mine industries) max = 0,898 (leather products except footwear) e/ min = 0,722 (cable, wike and their products) max=0,938 (leather products except footwear)

a/ = Distribution free, b/ = Gamma deterministic, c/ = Exponential deterministic.

d/ = Stochastic Normal/Halfnormal, e/ Stochastic = Normal exponential.

(implying that most firms are forced to have close to zero inefficiency), our results suggest caution regarding the appropriateness of this error structure.

Turning to the pattern of sectoral results, we find them in general to be plausible. For instance, spinning weaving and finishing of textiles has the lowest efficiency for both the statistical and non-parametic deterministic cases (the efficiency for this sector could not be estimated for the stochastic case). This sector, with an effective rate of protection of 492%, 1/ was one of the most highly protected sectors in the whole economy.

In turn, the most efficient sector, pulp, paper and paperboard, has a fairly homogenous output and was the leading export sector, by far, in manufacturing (42% of output was exported). For the stochastic measures of efficiency, the most efficient sector was leather products except footwear which had one of the lowest effective protection rates in the import competing sector (18%). The most inefficient sectors were wine industries (half normal) and cable wire and their products (exponential case). This result for the wine sector is surprising as this is an exportable sector. However, it could be due to a overly restrictive measure of capital that excludes inventories (aging wine) which is an essential input in production. Finally, cable wire and their products have an effective protective rate slightly below the median (64%) but is the fifth most concentrated sector in manufacturing (see column 6, Table 1).

VI. COMPARISON OF ALTERNATIVE MEASURES

As mentioned in the introduction, one of the problems faced by the practitioner is the sensitivity of results to model selection including the

---

1/ Corbo and Meller (1981 p. 96).

selection of functional form for the average production frontier. Yet, for a given functional form, an empirical issue remains: how sensitive are the resulting measures of technical efficiency to the selection of error structure and to the specific characteristics of the distribution of the error term? This is an empirical issue over which there is so far little evidence to draw upon. Two comparative studies are available. One by Van den Broeck et al (1980) is a comparison for a panel of 28 Swedish dairy plants of the programming, statistical and stochastic approaches. The other by Kopp and Smith (1982) compares alternative formulations of the production function (Cobb-Douglas, CES and translog) along with the same three approaches on a cross-section of 43 steam electric generating plants. Although these papers are a contribution towards a better understanding of technology and efficiency in the specific sectors studied, they are too limited in coverage to be useful for assessing whether the measures of inefficiency are sensitive to the selection of computational method. Our census data set is the most appropriate for such an evaluation.

We investigate the correlation among the different measures, using both Pearson and Spearman correlation coefficients. The results are reported in Table 4 which also gives the significance level for the test that the population value of the respective coefficient is equal to zero as well as the number of observations used in the computation of the correlation coefficients. Consider first the two stochastic measures (EFCHN, EFCEX). Both correlation coefficients are numerically very close to one. This leads to the first conclusion: there is little to be gained in a cross-sector comparison by choosing between a half-normal and an exponential error structure for the inefficiency component of the composite error. The same

Table 4

Correlation Among Efficiency Estimates 1/

PEARSON CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

	EFLP	EFSFR	EFSGA	EFSEX	EFCHN	EFCEX
EFLP	1.00000	0.81391	0.40982	0.44589	0.05891	-0.14242
	0.0000	0.0001	0.0070	0.0031	0.8164	0.5492
	42	42	42	42	18	20
EFSFR	0.81391	1.00000	0.53542	0.56309	0.45547	0.16280
	0.0001	0.0000	0.0002	0.0001	0.0575	0.4929
	42	43	43	43	18	20
EFSGA	0.40982	0.53542	1.00000	0.98511	0.45912	0.35208
	0.0070	0.0002	0.0000	0.0001	0.0553	0.1279
	42	43	43	43	18	20
EFSEX	0.44589	0.56309	0.98511	1.00000	0.50905	0.41248
	0.0031	0.0001	0.0001	0.0000	0.0310	0.0707
	42	43	43	43	18	20
EFCHN	0.05891	0.45547	0.45912	0.50905	1.00000	0.99925
	0.8164	0.0575	0.0553	0.0310	0.0000	0.0001
	18	18	18	18	18	18
EFCEX	-0.14242	0.16280	0.35208	0.41248	0.99925	1.00000
	0.5492	0.4929	0.1279	0.0707	0.0001	0.0000
	20	20	20	20	18	20

S T A T I S T I C A L   A N A L Y S I S   S Y S T E M

SPEARMAN CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

	EFLP	EFSFR	EFSGA	EFSEX	EFCHN	EFCEX
EFLP	1.00000	0.81778	0.45247	0.44553	0.20754	-0.09474
	0.0000	0.0001	0.0026	0.0031	0.4086	0.6912
	42	42	42	42	18	20
EFSFR	0.81778	1.00000	0.55981	0.55095	0.46051	0.17143
	0.0001	0.0000	0.0001	0.0001	0.0545	0.4699
	42	43	43	43	18	20
EFSGA	0.45247	0.55981	1.00000	0.99860	0.44444	0.40196
	0.0026	0.0001	0.0000	0.0001	0.0646	0.0789
	42	43	43	43	18	20
EFSEX	0.44553	0.55095	0.99860	1.00000	0.44525	0.41761
	0.0031	0.0001	0.0001	0.0000	0.0641	0.0669
	42	43	43	43	18	20
EFCHN	0.20754	0.46051	0.44444	0.44525	1.00000	0.99948
	0.4086	0.0545	0.0646	0.0641	0.0000	0.0001
	18	18	18	18	18	18
EFCEX	-0.09474	0.17143	0.40196	0.41761	0.99948	1.00000
	0.6912	0.4699	0.0789	0.0669	0.0001	0.0000
	20	20	20	20	18	20

1/ the measures appear in the same order as in Table 2 columns 1-6.

conclusion carries over to the comparison between the Gamma (EFGA) and exponential (EFX) error structures for the statistical frontiers.

Comparing the statistical and stochastic models indicates that although the correlation coefficients are statistically different from zero (approximately 5% significance level), they are only around 0.5. Therefore, a second conclusion is that in a cross-sector comparison of efficiency levels, the results are sensitive to the choice between a statistical deterministic and a stochastic frontier.

Finally, within the deterministic frontiers, there is a high correlation between the LP and the "distribution-free" statistical measures. Although both measures are the only ones in the set that force all observations to be below the frontier, one should note that while the "distribution free" frontier differs from the other statistical and stochastic frontiers by the value of the constant term, the LP frontier also allows for different slopes vis-a-vis the other measures. It therefore appears that there is little to be gained in choosing between alternative full frontier models. Thus, for the error structures usually considered in the literature, the main choice to be made is between a full frontier and a stochastic frontier approach.

VII. CONCLUSIONS

The purpose of this paper was to provide some guidance on the effect of alternative frontier models specifications on the measurement of technical efficiency. As stated in the introduction, the only systematic comparisons available so far were undertaken with extremely small data sets. The alternatives considered in the paper included parametric full frontier models (linear programming and statistical deterministic) and stochastic models. For the statistical deterministic and the stochastic models, we examined the influence of the various error structures proposed in the literature.

The models and error structures were evaluated on all establishments employing more than ten workers in the 1967 Chilean manufacturing census giving rise to estimation over 43 manufacturing sectors classified at the four digit ISIC level. The model comparisons indicated that the choice of error structures proposed in the literature has a very small impact on the measurement of inefficiency. However, in a cross-sector comparison of efficiency, results are sensitive to the selection between statistical and stochastic formulations. Finally, within the full frontier models, the linear programming and statistical models yield highly correlated measures of technical efficiency.

Another important finding is that, in contrast to other studies, we found that approximately half of the sectors considered could not support the estimation of a stochastic frontier because the skewness of the distribution of the overall residual was of the wrong sign. This result suggests at least two explanations: the error structures considered in the literature are not appropriate or the purged data might still include observations with measurement errors.

APPENDIX I: ESTIMATION OF PRODUCTION FRONTIERS

This appendix presents the three models estimated in the main body of the paper. It also gives the formulas for the efficiency measures presented in the text.

Model 1: Deterministic Full Frontier

The model is given by: $y = f(x)e^{-u}$; $u \geq 0$ where $f(x)$ is Cobb-Douglas in n factor inputs.

After taking logarithms, we have for the j^{th} observation $\ln f(x) = \beta_0 + \Sigma\beta_i X_{ij}$; $Y_j = \ln y_j$, $X_{ij} = \ln x_{ij}$.

The programming method for estimating the frontier consists of:

$$\text{Minimize} \quad \sum_{j=1}^{m} |\hat{u}_j|$$

$$\text{Subject to:} \quad \hat{\beta}_0 + \hat{\beta}_1 X_{11} + \dots \hat{\beta}_n X_{n1} \geq Y_1$$

$$\hat{\beta}_0 + \hat{\beta}_1 X_{1m} + \dots \hat{\beta}_n X_{nm} \geq Y_m$$

$$\hat{\beta}_0 \dots \hat{\beta}_n \geq 0$$

$$\text{and where:} \quad \hat{u}_j = \hat{\beta}_0 + \hat{\beta}_1 X_{1j} + \dots \hat{\beta}_n X_{nj} - Y_j$$

The estimation yields an estimate $[\hat{\beta}_0 \dots \hat{\beta}_n]$ for $[\beta_0 \dots \beta_n]$. See Aigner and Chu (1968) and Timmer (1971).

Efficiency Indexes

The efficiency index of firm j, E_j, is given by:

$$E_j = \frac{y_j}{\exp(\hat{Y}_j)}$$

where $\hat{Y}_j = \hat{\beta}_0 + \sum_i \hat{\beta}_i X_{ij}$

and the sector's weighted average efficiency index is given by:

$$AE = \Sigma W_j E_j$$

where $W_j = \frac{y_j}{\Sigma y_j}$

Model 2: Statistical Frontier

Model: $y = f(x)e^{-u}$ $u \geq 0$

The model to be estimated is linear in parameters and is given by:

$$Y = \beta_0 + X\beta_{-1} - u, \text{ where } Y = \ell ny,\ X = [1\ \ \ell nx_1, \ldots \ell nx_n]$$

and $\beta_{-1} = (\beta_1 \ldots \beta_n)$

and X is independent of u

Case (a) u is iid from the one parameter Gamma distribution

$$g(u;\ \phi) = \frac{1}{\Gamma(\phi)} u^{(\phi-1)} \exp(-u)$$

for which $E(u) = \phi$, $\mathrm{var}(u) = \phi$

Case (b): u is iid from the exponential distribution of

$$g(u, \phi) = 1/\phi \exp(-u/\phi)$$

for which $E(u) = \phi; \text{ var }(u) = \phi^2$

Efficiency Indexes

$$E_j = y_j / \exp(\hat{\hat{Y}}_j) = \exp(-\hat{\hat{u}}_j)$$

where $\hat{\hat{Y}}_j$ is obtained using the COLS estimator described below and $\hat{\hat{u}}_j$ is the residual from the COLS estimator. The COLS unbiased estimator of β_0 is given by:

$$\hat{\hat{\beta}}_0 = \hat{\beta}_0 + E(u)$$

where $\hat{\beta}_0$ is the OLS estimator of β_0.

A consistent estimate for E(u) is derived from the choice of the distribution function for u. Two efficiency indexes can be defined. The first measure is the average efficiency index computed at the point of means:

$$AE = \frac{\exp \bar{Y}}{\exp[\bar{Y} + E(u)]} = e^{-E(u)}$$

where $\bar{Y} = \ln f(\bar{x})$

Figures obtained from this measure are not reported in the paper since they are very close to the values obtained from the expected efficiency measure presented below.

The second measure is the expected efficiency of the sector. It is obtained by aggregating over firms and is given by:

$$E\,(e^{-u}) = \begin{cases} 2^{-\hat{\sigma}2} & \text{for the Gamma case} \\ (1 + \hat{\sigma})^{-1} & \text{for the exponential case} \end{cases}$$

Where $\hat{\sigma}$ is the standard error of the OLS regression.

It should be noted that some observations may lie above the frontier which is troublesome when computing efficiency at the firm level. The only way to guarantee that all observations lie under the frontier is to apply the consistent estimate for β_0 proposed by Greene (1980, pp. 31-34).

$$\hat{\hat{\beta}}_0 = \hat{\beta}_0 + \max \mid (\hat{u}_i) \mid$$

The use of this correction factor, although it affects the measure of efficiency for individual establishments, it does not affect the computation of expected efficiency since it does alter the value of $\hat{\sigma}$. Although consistent, this estimator is different from the COLS estimator of β_0 discussed above.

In this case, a "distribution free" measure of average efficiency computed at the point of means is given by:

$$AE = \frac{\exp \overline{Y}}{\exp\,(\overline{Y} + \max \mid (\hat{u}_i) \mid)}$$

This is the measure reported in Table 2, col. 2.

Model 3: Stochastic Frontier

Model: $y = f(x)\, e^{v-u}$

As before the model is linear in parameters and given by:

$$Y = \beta_0 + X\beta_{-1} + \varepsilon$$

where $\varepsilon = v - u$ $v \sim$ iid $N(0, \sigma^2 v)$

$u \geq 0$ $u \sim$ iid half-normal, exponential (or Gamma), u and v independently distributed.

As before, $\hat{\beta}_{-1}$ is unbiased and efficient and an unbiased estimator of β_0 is given by:

$$\hat{\hat{\beta}}_0 = \hat{\beta}_0 + E(u)$$

A measure of the relative variability of the two sources of error is given by:

$$\lambda = \frac{\sigma_u}{\sigma_v}$$

Estimation of $E(u)$, σ_u, σ_v relies on the observation that the moments of the distribution of $\varepsilon = v - u$ can be expressed in terms of the moment of the distribution describing u and v and of the property that the moments of ε can be estimated consistently from the moments of the OLS residuals. 1/ However

---

1/ Since $V \sim N(0, \sigma_v^2)$, $(0, \sigma_v^2)$, the second and third central moments of the distribution of ε are given by:

$$\mu_2' = E(u^2) + E(v^2) - (E(u))^2$$

$$\mu_3' = 3E(u^2)\, E(u) - 2E(u)^3 - E(u^3)$$

Replacing μ_2' and μ_3' by their estimates from the OLS residuals and replacing moments by the relevant parameters yields the desired estimates.

it may either turn out that $\hat{\mu}'_3$ the estimate of μ'_3 (which is always negative) has the wrong sign or that $\hat{\sigma}^2 v = \hat{\sigma}^2 \varepsilon - \hat{\sigma}^2 u \leq 0$. 1/ Either ocurrence, raises questions about the sample and/or the appropriateness of the selected error structure.

The COLS consistent estimators and formulas for the models used in the text are:

Distribution of u

	Half Normal	Exponential
$E(u)$	$\sqrt{\frac{2}{\pi}}\ \hat{\sigma}_u$	$[\frac{\hat{\mu}'_3}{2}]^{1/3}$
$\hat{\sigma}^2 u$	$[\sqrt{\frac{2}{\pi}}\ (\frac{\pi}{\pi-4})\ \hat{\mu}'_3\ ]^{2/3}$	$[\ (\frac{\hat{\mu}'_3}{2})\ ]^{2/3}$
$\hat{\sigma}^2 v$	$\hat{\mu}'_2 - (\frac{\pi-2}{\pi})\ \hat{\sigma}^2_u$	$\hat{\mu}'_2 - \hat{\sigma}^2_u$
$E(e^u)$	$2e^{\hat{\sigma}^2 u/2}\ (1 - F^*\ (\hat{\sigma}_u))$	$\frac{1}{1 + \hat{\sigma}_u}$

Finally, to compare levels of efficiency across observations, one forms the conditional distribution of u_i given e_i, $f(u_i/\varepsilon_i)$:

$$f(u_i/\varepsilon_i) = \frac{f(u_i,\ \varepsilon_i)}{f(\varepsilon_i)}$$

1/ See Olson et al (1980, p. 70), Schmidt and Lovell (1976, p. 351).

and use the mean of this distribution as a point estimate of u.

As derived by Jondrow et al (1982), the measure of firm efficiency based on the means are given by:

$$E(u_i/\varepsilon_i)_{HN} = (\hat{\sigma}^2_u \; \hat{\sigma}^2_v \; \hat{\sigma}^2) \left[\frac{f(\hat{\varepsilon}_i \hat{\lambda}/\hat{\sigma})}{1-F(\hat{\varepsilon}_i \hat{\lambda}/\hat{\sigma})} - \hat{\varepsilon}_i \frac{\hat{\lambda}}{\hat{\sigma}} \right]$$

$$E(u_i/\varepsilon_i) \text{ expon.} = \hat{\sigma}_v \left[\frac{f(\hat{\varepsilon}_i/\hat{\sigma}_v + \hat{\lambda}^{-1})}{1-F(\hat{\varepsilon}_i/\hat{\sigma}_v + \hat{\lambda}^{-1})} - (\hat{\varepsilon}_i/\hat{\sigma}_v + \hat{\lambda}^{-1}) \right]$$

where $\hat{\varepsilon}_i$ are the residuals of the COLS regression $\hat{\sigma}^2 = \hat{\sigma}^2_u + \hat{\sigma}^2_v$ and the other parameters take the values given by the estimates discussed above.

APPENDIX II:

THE MEASUREMENT OF EFFICIENCY UNDER A PROTECTED TRADE REGIME

One of the difficulties with the estimation of efficiency within a sector is the existence of a differentiated structure of effective rates of protection across firms within a specific ISIC sector.

This problem can be studied as a special case of specification error. We take the special case of a Cobb-Douglas function. 1/
The estimated function is:

$$(1) \quad \ln Q^D_i = \alpha^D + \beta^D \ln K_i + \gamma^D \ln L_i + v^D_i - u^D_i$$

whereas the correct model is:

$$(2) \quad \ln Q^I_i = \alpha^I + \beta^I \ln K_i + \gamma^I l_n L_i + v^I_i - u^I_i$$

Where Q^I_i is value added at international prices.

Under protection, value added at domestic prices is given by:

$$Q^D_i = (1 + ERP_i)\, Q^I_i$$

1/ We saw above that in a majority of cases the null hypothesis of a Cobb-Douglas technology could not be rejected from the data.

where ERP_i is the effective rate of protection for sector i, and superscripts D and I refer to valuation at domestic and world prices respectivey.

From the above equation we obtain:

$$\ln Q^D_i = \ln (1 + ERP_i) + \ln Q^I_i$$

Replacing in the correct model of equation (2) we obtain the following model:

$$\ln Q^D_i = \alpha^I + \beta^I \ln K_i + \gamma^I \ln L_i + \ln(1+ERP_i) + v^I_i - u^I_i$$

Thus in equation (1) we have left out the variable $(1+ERP_i)$. The implication of this specification error (Maddala pp. 459-460) is that:

$$E [\hat{\beta}^D] = \beta^I + P_{42}$$

$$E [\hat{\gamma}^D] = \gamma^I + P_{43}$$

Where P_{42} and P_{43} are the coefficients of $\ln K_i$ and $\ln L_i$ respectively in the auxiliary linear regression of $\ln(1+ERP_i)$ on a constant, $\ln K_i$ and $\ln L_i$.

If $\ln(1+ERP_i)$ and $\ln K_i(\ln L_i)$ are positively (negatively) associated as has usually been found, then $P_{42} > 0$, $P_{43} < 0$ and $\hat{\beta}^D$ is upward biased ($\hat{\gamma}^D$ is downward biased).

Furthermore $E[\hat{\sigma}^2_D] > \sigma^2$. Where $\sigma^2 = V(v)+V(u)$ and $\hat{\sigma}^2_D$ is an estimator of the variance obtained from model 1.

Therefore, the estimator of the variance obtained from the "wrong"model is upward biased.

REFERENCE

Aigner, D. J., and S. F. Chu (1968). "On Estimating the Industry Production Function," American Economic Review 58, 826-839.

Behrman, J. (1976). "Foreign Trade Regimes and Economic Development: Chile." Columbia University Press for National Bureau of Economic Research.

Bhagwati, J. and T. N. Srinivassan (1979). "Trade Policy and Development" in R. Dornbusch and J. A. Frenkel (eds.) International Economic Policy: Theory and Evidence. Baltimore: John Hopkins University Press.

Carlsson, B. (1972). "The Measurement of Efficiency in Production: An Application to Swedish Manufacturing Industries 1968." Swed. J. of Economics 468-485.

Corbo, V. and P. Meller (1977). "Alternative Trade Strategies and Employment Implications: Chile." Manuscript, N.B.E.R.

Corbo, V. and P. Meller (1979a). "The Translog Production Function: Some Evidence from Establishment Data." Journal of Econometrics 10 (June) pp. 193-199.

Corbo, V. and P. Meller (1979b). "La Sustitución de Trabajo, Capital Humano y Capital Físico en la Industria Manufacturera Chilena." Estudios de Economía N° 14, Segundo Semestre 1979.

Corbo, V. and P. Meller (1981). "Alternative Trade Strategies and Employment Implications: Chile" in A. G. Krueger, H. B. Lary, T. Monson, and N. Akrasanee (eds.) Trade and Employment in Developing Countries: Vol 1 Individual Studies. Chicago: University of Chicago Press.

Corbo, V. and P. Meller (1982). "The Substitution of Labor, Skill, and Capital: Its Implications for Trade and Employment" in Anne O. Krueger (ed.) Trade and Employment in Developing Countries: Vol 2 Factor Supply and Substitution, Chicago: University of Chicago Press.

Corbo, V. and J. de Melo (1983), "Technical Efficiency in a Highly Protected Economy: Preliminary Results for the Chilean Manufacturing Sector: 1967" (mimeo), January, The World Bank.

Farrell, M. J. (1957). "The Measurement of Productive Efficiency," Journal of the Royal Statistical Society. A 120, 253-281.

Forsund, F., C. A. Knox-Lovell, and P. Schmidt (1980). "A Survey of Frontier Production Functions and of Their Relationship to Efficiency Measurement." Journal of Econometrics 13 (May) 5-25.

Frank, C., K. Suk Kim, L. Westphal (1975). Foreign Trade Regimes and Economic Development: South Korea. Columbia University press for NBER.

Greene, W. H. (1980a). "Maximum Likelihood Estimation of Econometric Frontier Functions," Journal of Econometrics Vol 13 N°1 (May) pp. 27-56.

Griliches, Z. (1967). "Production Functions in Manufacturing: Some Preliminary Results" in M. Br-wn (ed.) The Theory and Empirical Analysis of Production, New York: Columbia University Press for National Bureau of Economic Research.

Jondrow, J. C. A. Knox Lovell, I. Materov, and P. Schmidt (1982). "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model." Journal of Econometrics 19 (August) 233-238.

Kopp, R. (1981). "The Measurement of Productive Efficiency: A Reconsideration." Quarterly Journal of Economics Vol XCVI, August, N°3.

Kopp, R. and V. Smith (1982). "Frontier Production Function Estimates for Steam Electric Generation: A Comparative Analysis." Southern Economic Journal, 1982, pp 1049-1059.

Krueger, A. (1972). "Evaluating Restrictionist Trade Regimes: Theory and Measurement." Journal of Political Economy, Vol. 80 N°1.

Krueger, A., et al. (1981). Trade and Employment in Developing Countries, Vol. 1. Individual Studies. Chicago: University of Chicago Press.

Lee, L. F. and W. G. Tyler (1978). "The Stochastic Frontyier Production Function and Averag Efficiency: An Emprical Analysis," Journal of Econometrics 7, 385-389.

Little, I., T. Scitovsky and M. Scott (1970). Industry and Trade in Some Developing Countries. New York: Oxford University Press.

Meller, P. (1976). "Efficiency Frontiers for Industrial Establishments of Different Sizes." Explorations in Economic Research, Occasional Papers of the National Bureau of Economic Research 3, 379-407.

Meller, P. and C. Swinburn (1975), "5000 Indicadores de Concentracion Industrial," Documento de Trabjao, Universidad Catolica de Chile, Santiago, Chile.

Meussen, W. and J. van den Broeck (1977). "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error, International Economic Review, 18, N° 2, June, 435-444.

Olson, J., P. Schmidt, and D. Waldman (1980). "A Montecarlo Study of Estimators of Stochastic Frontier Production Functions." Journal of Econometrics 13. 67-82.

Pitt, M. and L. F. Lee (1981). "The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry." Journal of Development Economics, 9, August, 43-64.

Richmond, J. (1974). "Estimating the Efficiency of Production." International Economic Review 15, 515-521.

Salter (1969). Productivity and Technical Change, Cambridge, 1969.

Schmidt, P. (1976), "On the Stastistical Estimation of Parametric Frontier Production Functions." Review of Economics and Statistics, 238-239.

Schmidt, P. and C. A. Knox Lovell (1979). "Estimating Technical and Allocative Inefficiency Relative to Stochastic Production and Cost Frontiers," Journal of Econometrics 9, N°3, Feb. 343-366.

Stern, R. (1975). "Capital-Skill Complementarity and U.S. Trade in Manufactures." In Quantitative Studies of International Economic Relations, ed., H. Glejser. Amsterdam: North Holland.

Stigler, G. J. (1976), "The Xistence of X-efficiency, American Economic Review, 66, 1, March 213-216.

Taylor, L. and E. Bacha (1973), "Growth and Trade Distortions in Chile and their Implications in Calculating the Shadow Price of Foreign Exchange" in R. S. Eckaus and P. N. Rosenstein-Rodan (eds.). Analysis of Development Problems.

Tyler, W. (1974), "Labour Absorption with Import Substituting Industrialization: An Examination of the Elasticities of Substitution in the Brasilian Manufacturing Sector," Oxford Economic Papers.

Tyler, W. and L. F. Lee (1979), "On Estimating Stochastic Frontier Production Functions and Average Efficiency: An Empirical Analysis with Colombian Micro Data." The Review of Economics and Statistics, Vol. LXI, No. 3, 436-438.

Waldman, D. M. (1977), "Estimation in Economic Frontier Functions." Unpublished manuscript.

Zellner, Kmenta and Dreeze (1966). "Specification and Estimation of Cobb-Douglas Production Function Models," Econometrica, 34, 784-795.