Decomposing Production Efficiency into Technical, Allocative and Structural Components

G. Anandalingam, Nalin Kulatilaka

Journal of the Royal Statistical Society. Series A (General), Volume 150, Issue 2 (1987), 143-151.

J. R. Statist. Soc. A (1987)
150, *Part* 2, *pp.* 143–151

Decomposing Production Efficiency into Technical, Allocative and Structural Components

By. G. ANANDALINGAM and NALIN KULATILAKA†

University of Virginia, USA *Boston University and Energy Laboratory, M.I.T., USA*

SUMMARY

This paper derives a technique by which production efficiency can be decomposed into (a) allocative and (b) technical components that are within the control of the firm, and (c) a structural component that is determined by the economic environment. The methodology is implemented by estimating the restricted cost frontier, solving a cost minimization problem to infer the efficient cost frontier, and measuring all efficiencies along the same ray through the origin of the input quantity space.

Keywords: DECOMPOSITION OF PRODUCTION EFFICIENCY; TECHNICAL EFFICIENCY; ALLOCATIVE EFFICIENCY; STRUCTURAL EFFICIENCY; PRODUCTION FRONTIER; COST FRONTIER

INTRODUCTION

Ever since the initial work of Farrell (1952), there have been numerous studies that estimated production efficiency (see the recent reviews by Forsund *et al.*, 1980; Schmidt, 1985). In a pioneering study Farrell (1957) identified two components of production efficiency: *Technical efficiency* which measures a firm's success in producing maximum output (or set of outputs) from a given set of inputs; *Allocative efficiency*, which Farrell calls "price" efficiency and measures the firm's success in choosing an optimal set of inputs. The underlying premise of Farrell and the ensuing literature is that the removal of technical and allocative inefficiencies will yield efficient production. When cast in the dual cost minimization framework the corresponding assumption is that deviations of actual production costs from the minimum cost is due to technical and allocative inefficiencies. [Kopp and Diewert (1982) show the correspondence between the primal production frontier and the dual cost frontier methods of efficiency analysis.]

The reduction of inefficiency then is completely within the control of the firm. Firms need to choose a technology that can produce at minimum cost in order to eliminate technical inefficiency and then adjust the mix of factor inputs (such as labour, capital, materials etc.) to suit the prevailing market prices in order to eliminate allocative inefficiency. In many production situations the presence of government intervention, in the form of regulation and rationing, and factor adjustment costs could prevent firms from adjusting factors instantaneously. These structural rigidities are external to the firm and are prevalent in many countries. For example, see the analysis by Spierer and Wood (1984) of adjustment problems in Swiss Manufacturing under natural gas regulations, and the study by Kulatilaka and Anandalingam (1986) of the impact of capital rationing on factor use in India. Hence, structural rigidities will contribute to measures of allocative inefficiency and should be estimated, separately.

Farrell (1957) noticed that "there remains the question of whether a high price (i.e. allocative) efficiency is necessarily desirable." He argued that in a dynamic context, firms may over invest in factors in the short-run in order to achieve long-run goals and thus appear to be allocatively inefficient in a static sense. After remarking that allocative "inefficiency" is a measure that is both unstable and dubious of interpretation, Farrell did not pursue it further. Forsund and Hjalmarsson (1974) also noted that an analysis of efficiency which does not take into account

† *Address for correspondence*: School of Management, Boston University, 704 Commonwealth Avenue, Boston, Ma 02215, USA

 0035–9238/87/150143 $2.00

the adjustment behaviour will tend to overestimate allocative inefficiency. Forsund *et al.* (1980) argue that one must be cautious when interpreting measures of allocative inefficiency. In this paper we provide, for the first time, an explicit method for measuring this third component of efficiency which we call *structural efficiency.*

The word "efficiency" may be a misnomer in this context. Firms do not adjust factors corresponding to the prevailing prices because it may be optimal for them not to do so. Thus structural efficiency merely estimates the mismeasure of allocative inefficiency. For example, if the structural efficiency is 80 per cent, then the static measure of pure allocative efficiency would be underestimated by 20 per cent. However it should be noted that while technical and allocative inefficiencies are production decisions within the control of the firm, the structural components are primarily due to the external environment which can be changed by appropriate government policies. Thus we are justified in using the term structural *efficiency* because this component measures how the external environment reduces the efficiency of factor mobility. For instance, the identification of structural inefficiencies in an economy with regulations and rationing will point towards government policies that can reduce these distortions.

In this paper, we provide a method for estimating structural inefficiency in a manner that is consistent with Farrell measures of efficiency. We explicitly allow for partial adjustment in the observed data and analyze production efficiency using a partial static equilibrium framework where some of the imputs are quasi-fixed. Such a framework is described by a *restricted* cost frontier. (See McFadden, 1978). Short-run allocative and technical efficiencies are estimated with respect to the restricted cost frontier at which variable costs are minimized subject to constraints on the levels of quasi-fixed factors. From such a restricted cost frontier we can infer the unrestricted full equilibrium frontier by solving the total cost minimization problem. The full equilibrium frontier will provide long-run allocative and technical efficiencies. From the short-run and long-run efficiencies, structural efficiency can be estimated.

We illustrate the decomposition of efficiency into technical, allocative *and* structural components using the well known data set first introduced by Berndt and Wood (1975). These data have been used in previous frontier estimation efforts by Kopp and Diewert (1982) and Burley (1980) and, thus, provide a basis for comparison. Results of this pedagogical example indicate that structural inefficiency could account for much of what was previously thought to be allocative inefficiency.

The rest of the paper is organized as follows: In the next section we briefly discuss the Farrell measures of efficiency and present their interpretations in the restricted cost framework. We then define structural efficiency in a manner that is consistent with the Farrell framework. The Zeischang (1983) method of decomposing efficiency within a full static cost frontier is adopted to the present framework in obtaining the technically efficient point. Section 3 illustrates the method using a translog functional form to specify the restricted cost frontier. In section 4 we summarize and make some concluding remarks.

2. MEASURES OF EFFICIENCY

Consider a production process which employs two short-run variable factors (levels of which are entirely within the control of the firm) and one or more quasi-fixed factors (levels of which are constrained due to government regulation or adjustment costs and delays) to produce a single output of level Y. For example, in a four factor production setting, materials and energy are usually variable while capital and contractually bound labour would be quasi-fixed. If this constrained production technology can be characterized by a linear homogeneous *restricted* production frontier, the *restricted* production efficiency can be measured on a ray drawn from the origin of the variable input space to the observed inefficient production point.

Fig. 1 depicts an inefficient production point A and a unit isoquant (SS') which is conditioned on levels of the quasi-fixed factors, Z. When facing a particular variable input price ratio (represented by the isocost line PP') it will be optimal (for a cost minimizing firm) to produce at the tangency point E. Following Farrell (1957) we can define the restricted production efficiency and its components as follows:

$$\text{restricted production efficiency } (RPE) = \frac{OC}{OA} = \| X^C \| / \| X^A \|, \tag{1}$$

$$\text{restricted allocative efficiency } (RAE) = \frac{OC}{OB} = \| X^C \| / \| X^B \|, \tag{2}$$

$$\text{restricted allocative efficiency } (RTE) = \frac{OB}{OA} = \| X^B \| / \| X^A \|, \tag{3}$$

where X^I denotes the vector coordinates of point I and $\| X^I \|$ is its vector norm.

In a full equilibrium setting where all factors are variable Kopp and Diewert (1982) derive computable expressions for similar measures in terms of the efficient full equilibrium cost frontier. Their result when applied to the restricted cost frontier yields

$$RPE = \frac{W.X^C + P.Z}{W.X^A + P.Z} = \frac{CR(W, Y, Z) + P.Z}{W.X^A + P.Z}, \tag{4}$$

$$RAE = \frac{CR(W, Y, Z) + P.Z}{W.X^B + P.Z}, \tag{5}$$

$$RTE = \frac{W.X^B + P.Z}{W.X^A + P.Z}, \tag{6}$$

where W and P are vectors of variable and quasi-fixed factor prices respectively, and $CR(.)$ is the restricted cost *frontier*. The point X^A represents the observed inefficient data and X^B is computed by noting that B lies on both the ray OA and on the restricted frontier SS'.

We stress that the quasi-fixed factors remain at their observed levels. As noted by Forsund *et al.* (1980) the existence of adjustment costs could make it efficient for a cost minimizing firm to keep some factors quasi-fixed in the short-run and to operate at input levels (X^E, Z). Although, they may be able to lower input costs in the long run by adjusing levels of quasi-fixed factors, doing so will incur higher adjustment costs than the input cost savings. Hence, E can be interpreted as the *short-run* efficiency point. In the long run firms may adjust (incurring lower adjustment costs) to a fully efficient point and, thus eliminate the *structural* inefficiency.

In other situations quasi-fixed factors may result from explicit government regulations. Such a regulated environment would also make it optimal for firms to operate at the restricted cost minimizing point E. The remaining *structural* inefficiencies (due to sub-optimal use of the quasi-fixed factors) are exogenous to the firm and should be the focus of government policy decisions.

The curve $S^*S^{*\prime}$ in Fig. 1 represents the unrestricted unit isoquant where quasi-fixed factors are at the optimal level Z^*. In the dual this corresponds to the unrestricted cost frontier which can be inferred from the restricted cost frontier by minimizing total costs with respect to Z; i.e. Z^* is the solution to the problem

$$\min_{Z} CR(W, Y, Z) + P.Z. \tag{7}$$

In any practical application, where $CR(.)$ is differentiable, Z^* will solve the first order condition

$$\frac{\partial CR}{\partial Z} + P = 0. \tag{8}$$

It should be noted that the solution to equation (8) will yield not only unrestricted values of $Z(=Z^*)$ but also values of the variable factors X at the new frontier (point E^*). The unrestricted efficiency measures can be defined in a way analogous to the restricted cases above.

$$\text{unrestricted production efficiency } (URPE) = \frac{OC^*}{OA} = \| X^C \| / \| X^A \| \tag{9}$$

$$\text{unrestricted allocative efficiency } (URAE) = \frac{OC^*}{OB^*} = \| X^{C*} \| / \| X^{B*} \| \tag{10}$$

$$\text{unrestricted technical efficiency } (URTE) = \frac{OB^*}{OA} = \| X^{B*} \| / \| X^A \| \tag{11}$$

Following an argument similar to that made by Kopp and Diewert (1982) we can express these efficiency measures in the dual:

$$URPE = \frac{CR(W, Y, Z^*) + P.Z^*}{W.X^A + P.Z}, \tag{12}$$

$$URAE = \frac{CR(W, Y, Z^*) + P.Z^*}{W.X^{B*} + P.Z^*}, \tag{13}$$

$$URTE = \frac{W.X^{B*} + P.Z^*}{W.X^A + P.Z}, \tag{14}$$

where X^{B*} is computed using one of the above mentioned decomposition methods.

Finally, we can measure the impact (inefficiency) due to the "restrictions" (e.g. adjustment costs, rationing, etc.) by the ratio OC^*/OC. This structural efficiency can be expressed in the dual framework as

$$SE = \| X^{C*} \| / \| X^C \| = \frac{CR(W, Y, Z^*) + P.Z^*}{CR(W, Y, Z) + P.Z} \tag{15}$$

The relationships between the various measures are summarized below.

(i) $RPE = RAE \times RTE$

(ii) $URPE = URTE \times URAE$

(iii) $URPE = RPE \times SE.$

The problem remaining is to find computationally feasible points for B and B* and estimate the technically efficient point. First, let us consider the restricted cost function and the computation of point B. We follow Zeischang's (1983) observation in noting that B (in Fig. 1) lies on the efficient surface and also on the ray OA. Since X^B is on the restricted cost frontier an application of the envelope theorem with respect to variable factor prices yield

$$X_i^B = \partial CR(W^B, Y, Z)/\partial W_i \qquad i = 1, \ldots, N. \tag{16a}$$

Notice that the cost function is evaluated at the prices W^B corresponding to the price plane which is tangent to the restricted cost function at B.

Since B lies on OA, the factor ratios at B and A must be identical:

$$X_i^A/X_N^A = X_i^B/X_N^B, \tag{16b}$$

where N is the variable factor against which all other factor prices are normalized. Combining (16a) and (16b) gives

$$\frac{X_i}{X_N} = \frac{\partial CR(W^B, Y, Z)/\partial W_i^B}{\partial CR(W^B, Y, Z)/\partial W_N^B}. \tag{17}$$

When the variable input prices are normalized with respect to the Nth variable factor price we can solve these N–1 equations to give the relative prices $w^B(= W^B/W_N^B)$ at each observation. A similar analysis at the point B^* will yield the prices for the unrestricted case.

3. AN ILLUSTRATION

In order to illustrate the decomposition of production efficiency into its three components, we have chosen a four factor *KLEM* model where capital (K) is assumed to be quasi-fixed. We use a translog functional form to represent the restricted cost frontier in order to highlight the generality of efficiency decomposition in the dual framework. Thus the restricted cost frontier is given by

$$\begin{aligned}\ln CR = a_0 + a_L \ln w_L + a_E \ln w_E + a_Y \ln (Y/K) + \tfrac{1}{2}a_{LL}[\ln w_L]^2 \\ + \tfrac{1}{2}a_{EE}[\ln w_E]^2 + a_{LE} \ln w_L \ln w_E + a_{YL} \ln w_L \ln (Y/K) \\ + a_{YE} \ln w_E \ln (Y/K) + \tfrac{1}{2}a_{YY}[\ln(Y/K)]^2 + \ln K + \ln W_M,\end{aligned} \tag{18}$$

where $w_i = W_i/W_M$ and W_M is the price of materials.

In order to obtain the relative factor prices at point B, (i.e. to satisfy equation 17), the following system of simultaneous equations has to be solved:

$$X_L/X_M - [a_L + a_{YL} \ln(Y/K) + a_{LL} \ln w_L + a_{LE} \ln w_E]/w_L \,.\, \Delta C_M = 0 \tag{19}$$

and

$$X_E/X_M - [a_E + a_{YE} \ln w_E + a_{LE} \ln w_L]/w_E \,.\, \Delta C_M = 0, \tag{20}$$

where

$$\Delta C_M = 1 - a_L - a_E - (a_{YL} + a_{YE}) \ln (Y/K) - a_{LL} \ln w_L - a_{EE} \ln w_E - a_{LE}(\ln w_E + \ln w_L). \tag{21}$$

The restricted and unrestricted cases are obtained by substituting in turn K and K^* is obtained by solving:

$$P_K - CR[a_Y + a_{YL} \ln w_L + a_{YE} \ln w_E - a_{YY} \ln(Y/K^*) - 1]/K^* = 0, \tag{22}$$

where CR is given by (18).

For purposes of illustrating the efficiency computations we used the time series data on U.S. manufacturing constructed by Berndt and Wood (1975). This data set covers the period 1947–1971 and consists of prices and expenditures of the four aggregate production factors, capital (K), labour (L), energy (E), and non-energy materials (M). A detailed description of the data sources and construction techniques together with a listing of the data can be found in the original Berndt-Wood paper.

We estimated the restricted cost *function* using a full information maximum likelihood technique and tested the resulting function for monotonicity and curvature conditions (see Kulatilaka, 1986). It is important to do so because all the derived results and Zeischang (1983) method can only work for strictly *regular* functions. Although the model estimated initially satisfied the monotonicity conditions, in several cases it failed the convexity/concavity conditions. We imposed *local* concavity and convexity (at the point of normalization) using a method similar to that suggested by Lau (1978) for static cost functions.

A number of methods exist for estimating production (and cost) frontiers (see the survey by Forsund *et al.* (1980)). For the purposes of this paper, in the spirit of Kopp and Diewert (1982), we have scaled several parameters of the above "practice" cost function with imposed convexity to make it appear as if it were a frontier. In particular, we altered the intercept term in order to form the convex hull (envelope) of all data points. This operation displaces the cost function in a neutral way and does not change the curvature of the cost function. As a consequence, at certain points restricted allocative efficiency would be equal to one.

The estimated efficiences are reported in Tables 1 and 2. Using the same restricted cost frontier we find that there are efficiency differences between the restricted and the unrestricted cases. More importantly, in the unrestricted case, structural inefficiency accounts for much of the total allocative inefficiency. Thus, analysts who do not decompose total allocative

TABLE 1

Farrell restricted efficiency indexes

Year	Restricted Technical Efficiency RTE	Restricted Allocative Efficiency RAE	Restricted Production Efficiency $RPE = RTE \times RAE$
1947	0.953	0.988	0.941
1948	0.877	1.000	0.877
1949	0.916	0.987	0.906
1950	0.952	0.963	0.916
1951	0.939	0.981	0.921
1952	0.922	0.996	0.918
1953	0.929	0.995	0.924
1954	0.904	1.000	0.904
1955	0.930	0.985	0.916
1956	0.923	0.991	0.915
1957	0.911	0.998	0.909
1958	0.899	0.997	0.897
1959	0.912	0.994	0.907
1960	0.905	1.000	0.905
1961	0.903	1.000	0.903
1962	0.903	1.000	0.903
1963	0.916	0.998	0.915
1964	0.915	1.000	0.915
1965	0.929	0.990	0.920
1966	0.924	0.996	0.920
1967	0.919	0.994	0.913
1968	0.921	0.992	0.914
1969	0.917	0.992	0.910
1970	0.898	1.000	0.898
1971	0.927	0.971	0.900

inefficiency into the structural and non-structural components would tend to overestimate production inefficiencies due to incorrect factor distribution.

The interpretation of allocative efficiency must be done with extreme caution. While previous measures of allocative efficiency entirely ignored the possibility of external short-run constraints to factor adjustment, the present method assumes that all such constraints can be modeled on rigidities in adjusting the capital input.

Within this framework, changes in structural inefficiencies will stem from changes in tax provisions, technology, and government regulations which affect capital formation. Our results indicate year-to-year fluctuations in structural efficiency within a range of about 5%. Acknowledging this caveat, roughly 10 to 20 per cent of allocative inefficiency may be explained by quasi-fixity of capital. This framework, however, can be extended to allow for other quasi-fixed factors (e.g. contractually bound labour) and thus, allow for a wider array of sources of structural inefficiency.

5. CONCLUDING REMARKS

This paper presented a measure of efficiency that is additional to the usual Farrell indexes of production efficiency. Structural inefficiency is due primarily to the environment external to the firm's production activities Identifying and estimating structural inefficiency would be important for analysing industrial performance. Because structural inefficiency is primarily due to the external environment, in cases where structural efficiency is large, it would be necessary to set up government policy to overcome these rigidities. In addition, structural efficiency measures provide an additional feature for international comparisons of industrial efficiency and productivity. For instance it has been believed in the recent past that Japanese

TABLE 2
Farrell unrestricted efficiency index

Year	*Unrestricted Technical Efficiency URTE*	*Unrestricted Allocative Efficiency URAE*	*Unrestricted Production Efficiency URPE**	*Structural Efficiency SE***
1947	0.889	0.893	0.794	0.839
1948	0.835	0.963	0.804	0.874
1949	0.864	0.892	0.771	0.844
1950	0.898	0.871	0.782	0.848
1951	0.884	0.888	0.785	0.847
1952	0.870	0.902	0.784	0.849
1953	0.875	0.896	0.784	0.843
1954	0.859	0.925	0.794	0.868
1955	0.881	0.897	0.791	0.858
1956	0.871	0.891	0.776	0.843
1957	0.863	0.907	0.783	0.854
1958	0.861	0.922	0.795	0.879
1959	0.873	0.920	0.803	0.880
1960	0.864	0.922	0.796	0.872
1961	0.863	0.924	0.797	0.875
1962	0.861	0.923	0.795	0.867
1963	0.873	0.913	0.797	0.865
1964	0.871	0.914	0.796	0.863
1965	0.885	0.902	0.798	0.862
1966	0.880	0.907	0.798	0.862
1967	0.877	0.906	0.795	0.864
1968	0.883	0.909	0.802	0.872
1969	0.877	0.904	0.793	0.865
1970	0.859	0.913	0.784	0.865
1971	0.889	0.873	0.776	0.855

* URPE = URTE × URAE
= RPE × SE

** These numbers measure, to the extent that they depart from unity, the restraint imposed on efficiency by structural factors in the short-term.

industry is more efficient that U.S. industry primarily because the former can adjust faster to changes (Hayes, 1981). It would be imperative to identify structural inefficiency in order to resolve this issue. We are in the process of doing such research.

It should be noted however that *actual* cost frontiers should be estimated in order to obtain correct inefficiency results. The reader should be cautioned that the illustrative example presented here merely estimates a convex hull of the data on production.

The above measures of production efficiency are placed in perspective and summarized below using the production possibility diagram in Fig. 1.

Point on Figure 1	*Technically*	*Allocatively*	*Structurally*
A	Inefficient	Inefficient	Inefficient
B	Efficient	Inefficient	Inefficient
C	Efficient	Efficient	Inefficient
B*	Efficient	Inefficient	Efficient
C*	Efficient	Efficient	Efficient

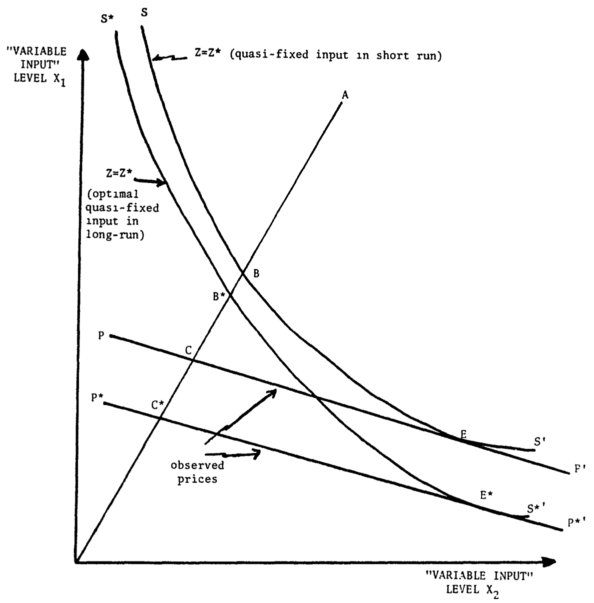

Fig. 1

Finally we examine the relationship between the above measures, and the technical efficiency (TE) and allocative efficiency (AE) measures derived from full static equilibrium (FSE) models. In FSE models the pairs of points B and B^*, and C and C^* (of Fig. 1) will coincide. That is, the model assumes structural efficiency ($SE = 1$) in estimating the frontier. Hence, AE and TE will be well defined only, if in fact, the observed data is structurally efficient. Recent empirical evidence has found significant departures from full static equilibrium invalidating the maintained hypothesis which forms the basis of previous frontier function models (see Kulatilaka, 1985). Hence, such techniques are likely to mismeasure components of production efficiency. [For example, Burley [1980] and Kopp and Diewart [1982] use the same Berndt-Wood data and obtain strikingly different results.]

ACKNOWLEDGEMENTS

We gratefully acknowledge financial support from *Resouces for the Future* and useful suggestions from a referee. We are solely responsible for any remaining errors.

REFERENCES

Berndt, E. R. and Wood, D. O. (1975) Technology, prices, and the derived demand for energy. *Rev. Econ. and Statist.*, **57**, 376–384.

Burley (1980) Productive efficiency in U.S. manufacturing: a linear programming approach. *Rev. Econ. and Statist.*, **62**, 619–622.

Farrell, M. J. and Fieldhouse, M. (1962) Estimating efficient production functions under increasing returns to scale. *J. R. Statist. Soc.* A, **125**, 252–267.

Farrell, M. J. (1957) The measurement of productive efficiency. *J. R. Statistic. Soc.* A, **120**, 253–281.

Forsund, F. R. and Hjalmarsson, L. (1974) On the measurement of productive efficiency. *Swedish J. Econ.*, **76**, 141–154.

Forsund, F. R. Lovell, C. A. K. and Schmidt, P. (1980) A survey of frontier production functions and their relationship to efficiency measurement. *J. Econometrics*, **13**, 5–25.

Hayes, R. H. (1981) Why Japanese factories work. *Harvard Bus. Rev.*, July–August, 56–66.

Kopp, R. J. and Diewert, E. (1982) The decomposition of frontier cost function deviations into measures of technical and allocative efficiency. *J. Econometrics*, **19**, 319–331.

Kulatilaka, N. H. (1985) Tests on the validity of static equilibrium models. *J. Econometrics*, **28**, 253–268.
———(1986) The specification of partial static equilibrium models, *Rev. Econ. and Statist.*, in press.
——— and Anandlingam (1986) Capital rationing and industrial factor demand. Working paper, School of Management, Boston University, January 1986.
Lau, L. J. (1978) Testing and imposing monotonicity, Convexity and quasi-convexity constraints. *In Production Economics: A Dual Approach to Theory and Applications*, Vol. 1. (M. Fuss and D. McFadden, eds.). Amsterdam: North-Holland.
McFadden, D. L. (1978) Cost, revenue, and profit functions. In *Production Functions: A Dual Approach to Theory and Applications*, Vol. 1, (M. Fuss and D. L. McFadden, eds.), Amsterdam: North-Holland.
Schmidt, P. (1985) Frontier production functions. *Econometric Rev.*, **4**, No. 2.
Spierer and Wood, D. (1984) Modeling swiss industrial interfuel substitution in the presence of natural gas supply constraints. MIT Energy Laboratory, WP 84-011, revised March 1986.
Zeischang, K. D. (1983) A note on the decomposition of cost efficiency into technical and allocative components. *J. Econometrics*, **23**, 401–405.