A complete list of Working Papers on the last page.

No. 17, 1978

GENERALIZED FARRELL MEASURES OF EFFICIENCY: An Application to Milk Processing in Swedish Dairy Plants

by

Finn R. Førsund

and

Lennart Hjalmarsson

May, 1978

First draft

March 1978

Revised May 1978

GENERALIZED FARRELL MEASURES OF EFFICIENCY: An Application to Milk Processing in Swedish Dairy Plants*)

by

Finn R Førsund, University of Oslo, Oslo, Norway

and

Lennart Hjalmarsson, University of Gothenburg and the Industrial Institute for Economic and Social Research, Stockholm, Sweden.

*) We wish to thank H J Bakke for valuable programming and computational assistance. We also wish to thank A Bagstam, B Hovmark, O Norrgren and Y Rasmusson of the Swedish Dairy Federation, SMR, for valuable assistance with establishing the data base, background understanding of milk processing and comments on the empirical results.

Financial support from the Swedish Council for Social Science Research and the Industrial Institute for Economic and Social Research is gratefully acknowledged.

A related paper has been presented at the European meeting on Antitrust and Economic Efficiency at Bruxelles, September 2-4, 1976. (Førsund & Hjalmarsson [1976].)

ABSTRACT

This paper is concerned with the measurement of productive efficiency. Farrell's measures of efficiency are generalized to nonhomogeneous production functions. Several new measures of efficiency have been introduced and applied to the Swedish milk processing industry. The empirical analysis is based on a complete set of cross section - time series data for a period of 10 years of 28 individual plants producing a homogeneous product, pasteurized milk. Industrial structure and structural change are examined by both studying the shape of the efficiency distributions for the individual units and their changes through time. The aggregate performance of the sector is studied by the development of the different measures of structural efficiency.

1. Introduction

The seminal paper by Farrell [1957] on the measurement of productive efficiency has inspired several studies during the last years on best-practice technology and efficiency measures. See e.g. Aigner & Chu [1968], Seitz [1970] and [1971], Timmer [1971], Todd [1971], Carlsson [1972], Meller [1976], Førsund & Jansen [1977], Meeusen & van den Broeck [1977] and Schmidt & Lovell [1977].

However, all these studies are cross section studies and based on estimation of efficiency relative to a frontier production function of the Cobb-Douglas type (with the exception of Førsund & Jansen [1977]), excluding the possibilities of studying the impact of scale economies on technical efficiency.

In this study Farrell's measures of efficiency are generalized (see Førsund & Hjalmarsson [1974a] to nonhomogeneous production functions, and applied to the Swedish milk processing industry during the period 1964-73. The analysis is based on a complete set of cross section time series data for the 10 years for 28 individual plants producing a homogeneous product, pasteurized milk.

The purpose of efficiency estimates at the industry level is to measure the relative performance of the plants or firms within an industry, and thereby to give a picture of the structure of the industry. For background considerations regarding efficiency measures, their interpretations and implications, see Førsund & Hjalmarsson [1974a]. In this study industrial structure and structural change is examined by studying both the shape of the efficiency distributions for the individual units and their changes through time. The aggregate performance of the sector is studied by the development of the different measures of structural efficiency.

There are two different methods to form a basis for measurement of efficiency, either to estimate an efficiency frontier or to estimate an explicit frontier production function.

Farrell's method is based on estimating a convex hull of the observed input coefficients in the input coefficient space when assuming production functions homogeneous of degree 1, and expanding the space to include output when assuming increasing returns to scale, i.e. an efficient surface is obtained for each value of output (Farrell & Fieldhouse [1962]. One disadvantage of this method is that a convex hull is an unduly pessimistic estimation of smoothly curved efficiency frontier isoquants. Another more serious disadvantage is that direct estimation of the efficiency frontier does not in general give enough information to get the whole representation of the production function which is necessary for establishing the efficiency measures employed in this study, when working with inhomogeneous functions. Even if the method of Farrell & Fieldhouse in principle, is applicable for *one* of the technical efficiency measures, for discrete levels of output, it is very cumbersome empirically and to our knowledge only one study has followed that approach. See Seitz [1970].

The method adopted here for estimation of efficiency measures is based on the estimates of the best-practice frontier or frontier production function which is a natural reference or basis for efficiency measures within an industry. The estimation method and empirical results are set out in Førsund & Hjalmarsson [1977]. The estimated function is given in the Appendix.

Three types of efficiency measures are usually distinguished: technical efficiency, scale efficiency, and price - or allocative efficiency. Below we are only concerned with technical efficiency and scale efficiency. These measures concern a certain plant or firm within a sector. For the industry as a whole, different measures of structural efficiency are also constructed and computed.

Originally Farrell's measures of efficiency were generalized to nonhomogeneous production functions in Førsund & Hjalmarsson [1974a]. However, in this paper these measures are further elaborated, some new measures of pure scale efficiency are defined and the relationship between the different measures and the scale properties of the production function are also shown.

All our measures hold generally for nonhomogeneous production functions. The measures are ray measures i.e. the distance between an observed unit and the reference path is measured along a factor ray. In general this can be justified by the splitting of total efficiency (originally due to Farrell) into two components, one showing potential cost reduction due to a proportional movement along a factor ray (technical efficiency and scale efficiency) and another showing the potential cost reduction due to movement along an isoquant (price efficiency). In this study we are not concerned with price efficiency.

2. The efficiency frontier

Efficiency measures and especially scale efficiency are often based on unit requirements of inputs, i.e. the production function, f, for a micro unit is transformed from the factor space into a space of input coefficients $\xi = (\xi_1,\ldots,\xi_n)$:

$$x = f(\frac{v}{x}\, x) = f(\xi x) \qquad (1)$$

where x is output and v is a vector $(v=v_1,\ldots,v_n)$ of inputs.

This transformation forms a set of feasible input coefficients bounded towards the origin and the coordinate axes of the factor space under certain restrictions on the forms of the micro unit production functions. A sufficient restriction is that the functions conform to the "regular ultra passum law" defined by Frisch [1965]; see also the analysis in Førsund [1971]. The set of input coefficients is not bounded for functions homogeneous of a degree $\neq 1$, but collapsing to a single curve for homogeneity of degree 1.

Assuming functional forms resulting in input coefficient sets bounded towards the origin and the coordinate axes of the factor space, the following definition is made:

Definition: The efficiency frontier for an industry consisting of m production units is made up of all points where the input coefficients $(\xi_1,\ldots,\xi_n)$ obtain their minimum values along rays through the origin. Under our regularity assumptions all such efficiency frontier points are boundary points of the feasible production set.

The efficiency frontier is the locus of all points where the elasticity of scale, $\varepsilon=1$, i.e. it is a technical relationship between inputs per unit of output for production units of optimal scale. Thus the efficiency frontier represents the optimal scale of the frontier production function.

The frontier production function and the efficiency frontier are illustrated in Figures 1 and 2. These figures are also utilized for illustration purposes of the different measures of efficiency below. In Figure 1 the production function, f, $(x=f(v))$, is cut with a vertical plane through the origin, i.e. v indicates a factor ray. The point P is an observed unit with inputs and output denoted by (v^o, x^o). The technically optimal scale is denoted by $\hat{x}$. In Fig. 2 optimal scale of the production function is transformed to the input coefficient space for the two-input case. Point A, of course, lies on the efficiency frontier. B and

C are the transformed points of the production surface in Fig.1 corresponding to output levels x^o and x^* respectively and D is an observed point $(v_1^o/x^o, v_2^o/x^o)$ corresponding to P in Fig.1.

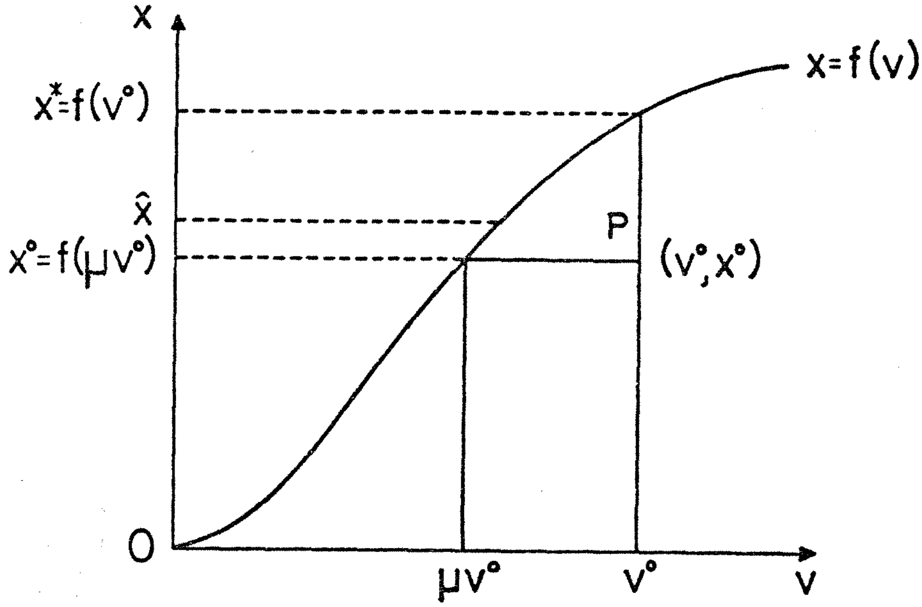


Figure 1. The frontier production function cut with a vertical plane along a ray through the origin. $\mu \in (0,1)$.

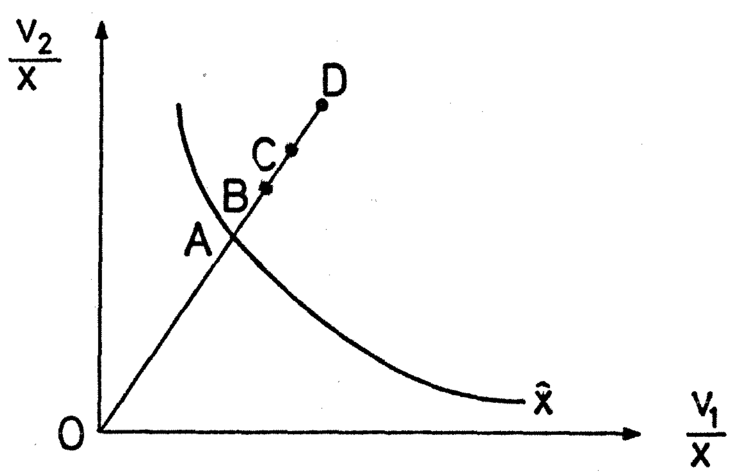


Figure 2. The efficiency frontier.

3. Technical efficiency

As shown in Førsund and Hjalmarsson [1974a] two different measures of technical efficiency denoted by E_1 and E_2 can be defined when allowing for production functions homogeneous of a degree different from 1 or inhomogeneous production functions. An illustration of the measures is provided in Fig.1 and 2.

One measure, E_1, is obtained by comparing an observed point of input requirements and output (v^o, x^o), with the input requirements on the frontier production function corresponding to the observed output. In Fig.1

$$E_1 = \mu \tag{2}$$

where μ is found by solving for μ from $x^o = f(\mu v^o)$

The measure shows the ratio between the amount of inputs required to produce the observed output with frontier function technology and the observed amount of inputs. (Input saving measure.)

In the input coefficient space this means comparing an observed input coefficient point with the point on the transformed isoquant of the frontier function corresponding to the observed output with the observed factor proportions. By definition this transformed isoquant must lie closer to the origin. The measure will then show the relative reduction in the amount of inputs needed to produce the observed output with frontier function technology with the observed factor proportions. In Fig. 2

$$E_1' = OB/OD \tag{3}$$

Another measure, E_2, is obtained by comparing an observed point of input requirements and output (v^o, x^o) with the output obtained on the frontier production function for the same amounts of input. This output is only possible to obtain when having an explicit production function. In Fig.1

$$E_2 = \frac{x^o}{f(v^o)} = \frac{f(\mu v^o)}{f(v^o)} \tag{4}$$

The measure shows the ratio between the observed output and the potential output obtained by employing the observed amount of inputs in the frontier function. (Output increasing measure.)

In the input requirement space this means comparing an observed point with the point on the transformed isoquant of the frontier production function corresponding to the output obtained by employing the observed amount of inputs in the frontier function. In Fig.2

$$E_2 = OC/OD \tag{5}$$

These two measures will in general not coincide except in the case of linear homogeneity. There exists an interesting relationship between E_1 and E_2 and the elasticity of scale (or the passus coefficient in the terminology of Frisch [1965]). In Frisch [1965], p 73, there is a formula (an identity) called the *second form of beam variation equation* which states that under a proportional variation of inputs

$$f(\mu v_1^o, \mu v_2^o, \ldots \mu v_n^o) \equiv f(\mu v^o) \equiv f(v^o) \cdot \mu^{\bar{\varepsilon}} \tag{6}$$

where $\bar{\varepsilon} = \int_\mu^1 \frac{\varepsilon(\tau)}{\tau} d\tau \;/\; \int_\mu^1 \frac{1}{\tau} d\tau$ which is a weighted average of the elasticity of scale in the interval between x^o and x^* in Fig.1.

Rearranging (6) yields

$$\frac{f(\mu v^o)}{f(v^o)} = \mu^{\bar{\varepsilon}} \quad \text{or} \tag{7}$$

$$\bar{\varepsilon} = \frac{\ln \left(\frac{f(\mu v^o)}{f(v^o)}\right)}{\ln \mu} \tag{8}$$

Substituting for E_1 and E_2 we obtain

$$\bar{\varepsilon} = \frac{\ln E_2}{\ln E_1} \tag{9}$$

Thus $E_1 \gtreqless E_2$ for $\bar{\varepsilon} \gtreqless 1$.

As stated above the two measures coincide when f is homogeneous of degree one.

The ranking of units according to the two measures of technical efficiency coincides if the elasticity of scale is constant or does not pass through the value of 1 in the sample. As we have chosen E_1 and E_2 to be figures with values between 0 and 1, E_1 is greater (smaller) than E_2 when the average of the elasticity of scale is greater (smaller) than one. Thus in Fig.2 we have arbitrarily chosen $E_1 < E_2$.

In empirical studies the choice between the measures has to be determined by the objective. If the amount of resources is assumed to be fairly constant, e.g. a fixed total employment, then E_2 is the relevant measure. If the framing of the problem is such that output is assumed to be constant, then E_1 is the relevant measure.

The efficiency measures derived from the production function specification employed in this paper are shown in the Appendix.

4. Scale efficiency

A measure of scale efficiency shows how close, in some sense, an observed plant is to the optimal scale. Three different measures of scale efficiency are defined here though, due to limited space, only one of the measures is shown in the empirical part of the study. These measures are of special interest in a long run analysis of potential possibilities of increased productivity.

The first measure of scale efficiency, E_3, shows the distance, in terms of input coefficient reductions, from an observed plant to the optimal scale on the frontier function and in Fig.2

$$E_3 = OA/OD \qquad (10)$$

The interpretation of the measure is the relative reduction in input coefficients made possible by producing at optimal scale on the frontier production function with the observed factor proportions. This measure is shown in the empirical part below.

E_3 is not a measure of pure scale efficiency. To obtain such a measure one has to eliminate the technical inefficiency of the observations by moving each observed unit to the surface of the frontier function. This can be done in two different ways corresponding to the two definitions of technical efficiency, i.e. by moving the units to the frontier either in the vertical or in the horizontal direction in Fig.1.

When moving a unit in the horizontal direction the second measure of scale efficiency, E_4, shows the distance from the transformed isoquant corresponding to x^o to the optimal scale and in Fig.2

$$E_4 = OA/OB \qquad (11)$$

When moving a unit in the vertical direction the third measure of scale efficiency, E_5, shows the distance from the optimal scale to the transformed isoquant corresponding to x^* and in Fig.2

$$E_5 = OA/OC \tag{12}$$

The interpretation of E_4 and E_5 is the relative reduction in input coefficients by producing in optimal scale on the frontier function for th observed factor proportions of a plant whose technical inefficiency has been eliminated in two different ways corresponding to the definition of E_1 and E_2 respectively.

From the definitions of the efficiency measures (2),(3),(10), (11) and (12) it follows easily from Fig.2 that

$$E_4 = E_3/E_1 \tag{13}$$

and

$$E_5 = E_3/E_2 \tag{14}$$

As the efficiency frontier constitutes the limit towards the origin of the feasible input coefficients, E_3 always shows a lower value than E_1 and E_2 except for units producing exactly in optimal scale on the frontier production function.

From (9), (13) and (14) we also get that

$$\bar{\varepsilon} = \frac{\ln E_3 - \ln E_5}{\ln E_3 - \ln E_4} \; . \tag{15}$$

This formula shows the relationship between the scale elasticity and the three different measures of scale efficiency. Thus, all measures of scale efficiency can be expressed as a function of the average elasticity of scale.

One must remember here that the average elasticity of scale $\bar{\varepsilon}$ depends on the observation chosen, i.e. a specific $\bar{\varepsilon}$ is obtained for each observation.

5. Structural efficiency

In his original article Farrell also suggested a measure of technical efficiency of the whole industry, i.e. a measure of structural efficiency, by simply taking a weighted average (by output) of the technical efficiencies of its constituent production units. In this paper we have extended the analysis of Farrell on this point and elaborate several other measures of structural efficiency.

According to Farrell [1957], p 262 the purpose of a structural efficiency measure is to measure "the extent to which an industry keeps up with the performance of its own best firms." In our context we want the structural measures to reflect the same for the industry as the individual efficiency measures show for a micro unit, i.e. potential input saving (E_1), potential increase of output (E_2) and potential reduction in input coefficents (E_3-E_5).

The approach suggested by Farrell is to weight the individual measures by observed output levels. Thus, the **first** measure of structural efficiency, here denoted by S_0, is obtained by taking the average of the E_1 technical efficiency measures with outputs as weights.

However the main problem with this approach is that the result of this weighting scheme does not have a straight-forward interpretation in terms of the objectives of the structural measures.

Another approach (indicated by Farrell's qualifications of the weighted measure) is to construct an average plant for the industry and regard this average plants as an arbitrary observation on the same line as the other observations and then compute E_1, E_2 and E_3 for this average unit. (In this paper the average plant is constructed by taking the *arithmetic* average of each amount of inputs and outputs). These measures of structural efficiency are denoted by S_1, S_2 and S_3 respectively where S_1 and S_2 are measures of structural technical efficiency and S_3 is a measure of structural scale efficiency.

These three latter measures seem to be more satisfactory as measures of structural efficiency as specified above than the S_0 measure. However, the reason for calculating S_0 is that is seems to be the only measure of structural efficiency that has been utilized in earlier studies. See e.g. Carlsson [1972].

By eliminating structural technical inefficiency by adjusting the average plant to the frontier in the two different ways corresponding to the E_1 and E_2 measures we obtain two other measures of pure structural scale efficiency corresponding to E_4 and E_5 denoted by S_4 and S_5. It is obvious that

$$S_4 = S_3/S_1 \tag{16}$$

and

$$S_5 = S_3/S_2 \tag{17}$$

Even in this case there exists a clear relationship between the scale properties of the production function and the efficiency measures. Because the average unit can be regarded as an arbitrary observation the relationship between the different measures of structural efficiency and the average of the elasticity of scale is the same as the relationship between the corresponding E_i measures. Thus

$$\bar{\varepsilon} = \frac{\ln S_2}{\ln S_1} \tag{18}$$

and

$$\bar{\varepsilon} = \frac{\ln S_3 - \ln S_5}{\ln S_3 - \ln S_4} \, . \tag{19}$$

By analogy with the E_i measures S_3 always shows a lower value than S_1 or S_2 except in the case in which the industry consists of a number of plants of optimal size employing the same best-practice technique, a situation that characterizes a long run equilibrium of an industry. (See Hjalmarsson [1973] for a discussion of optimal structure and structural change of an industry and long run equilibrium.)

While the relationship between S_1 and S_2 is given by eq. (18) it seems very difficult to establish analytically how S_0 is related to the other measures. Averaging units with $E_1 = 1$ yields an average with $E_1 < 1$ if they have different factor ratios i.e. the frontier units tend to contribute more to the S_0 measure than the S_1 measure. The relative impact on S_0 and S_1 of units below the frontier is difficult to assess. The results in Försund and Hjalmarsson [1976] indicated that when the dominant large unit was on or near the frontier S_0 was larger than S_1. When the largest unit really was highly inefficient S_0 showed a smaller value than S_1. In our empirical results below S_0 is always greater than S_1 even in the year when the largest unit has the lowest E_1 measure. This illustrates the impact of the whole structure on the difference between the measures.

6. The Data[1)]

In the empirical part of this study we have utilized primary data for general milk processing from 28 individual dairy plants during the period 1964-1973. We have received all data from SMR (Svenska mejeriernas riksförening), a central service organization for the dairies in Sweden.

1) The same data set is utilized in Førsund & Hjalmarsson [1977]

The processing of milk in a dairy can be divided into different stages of which each one can be referred to as a production process. The data used in this study refer to one such production process, namely general milk processing. This process includes reception of milk from cans or tanks, storage, pasteuration and separation. All milk passes this process before it goes further to different processes for consumption milk, butter, cheese or milk powder etc. Thus this stage defines the capacity of the plant. Moreover, general milk processing is often treated as a separate munit in cost accountings.

A strong reason for our choice of this part of a dairy is that it makes it possible to measure output in physical or technical units (tonnes) avoiding value added or gross output. This means that our estimated production function is more of a technical production function in the original sense.

Thus milk is regarded as a homogeneous product which is a very realistic assumption. Output is measured in tonnes of milk delivered to the plant each year. The amount of milk received is equal to the amount produced. There is no measurable waste of milk at this stage. According to SMR any difference is due to measurement errors.(Differences were of the magnitude of kilos.)

The labour input variable is defined as the hours worked by production workers including technical staff usually consisting of one engineer.

Capital data of buildings and machines are of user-cost type, including depreciation based on current replacement cost, cost of maintenance and rate of interest. The different items of capital are divided into five different subgroups depending on the durability of capital which varies between 6 and 25 years, so the capital measure is an aggregated sum of capital costs from these subgroups.

Capital costs, divided into building capital and machine capital, are calculated on the basis of these subgroups as a sum of the capital costs of the subgroups. The capital measure has been centrally accounted for by SMR according to the same principles for all plants and after regular capital inventory and revaluations by engineers from SMR. Afterwards we have aggregated building capital and machine capital into one measure. Thus, we have assumed that the conditions of the composite commodity theorem are fulfilled. In fact the relative prices of buildings and machine capital have developed almost proportionally during the 10-year period. The price index has moved from 100 in 1964 to 158 in 1973 for buildings and to 161 for machine capital. An alternative would be to retain the disaggregation of building and machine capital but in the case of a C-D kernel function, implying a unitary elasticity of substitution, this seems to be a less realistic assumption. Note that this capital measure is proportional to the replacement value of capital, which can serve as a measure of the volume of capital. See Johansen & Sørsveen [1967].

As the data are not adjusted for capacity utilization we have investigated a measure based on monthly maximum amount of milk received compared with the yearly average. This ratio is fairly stable over time, and the differences between plants are not very great. In consequence we have not corrected for capacity utilization. The increasing output over time for most of the plants support the assumption.

7. Empirical results

Structural efficiency

Let us first look at the aggregated picture of the industry. The estimates of structural efficiency are presented in Table 1 below.

Table 1. Estimates of structural efficiency.

Year	S_0 Weighted sum of efficiency measures	S_1 The distance of the average plant to the frontier function for given output. (Corresponds to E_1)	S_2 The distance of the average plant to the frontier function for given amount of inputs. (Corresponds to E_2)	S_3 The distance of the average plant to the efficiency frontier. (Corresponds to E_3)	S_4 S_3/S_1 Pure scale efficiency. (Corresponds to E_4)	S_5 S_3/S_2 Pure scale efficiency. (Corresponds to E_5)
1964	.7826	.7006	.6488	.6469	.9234	.9971
1965	.7465	.6941	.6337	.6305	.9084	.9950
1966	.7190	.6327	.5756	.5755	.9096	.9998
1967	.7018	.6264	.5622	.5619	.8970	.9995
1968	.6662	.6016	.5397	.5397	.8971	1.0000
1969	.6386	.5907	.5186	.5186	.8779	1.0000
1970	.6183	.5660	.4827	.4826	.8527	.9998
1971	.6561	.6004	.5020	.4994	.8318	.9948
1972	.6687	.6259	.5113	.5030	.8036	.9838
1973	.6475	.5928	.4715	.4658	.7858	.9879

The interpretation of the S_1 measure is the relative reduction in the amount of inputs needed to produce the observed industry output with frontier function technology with the observed factor proportions. Thus the table shows that the same output in the different years could have been produced by 70-59 % of the observed amounts used.

The S_2 measure shows the ratio between the observed output and the output obtained for the observed amount of inputs by using frontier function technology with the observed factor proportions. The table reveals that observed output is between 65 % and 47 % of potential output if the inputs were employed in units with frontier production technology.

The S_3-S_5 measures show the relative reduction in input coefficients by producing at optimal scale on the frontier function with the observed factor proportions. Thus e.g. for S_3 the table shows that at optimal scale on the frontier production function the potential input coefficients are 65-47 % of the observed input coefficients.

The most remarkable result is the high level of structural inefficiency measured by all the four measures S_0-S_3. Moreover, it seems to be a clear decreasing trend in the values of structural efficiency and not the contrary as most commentators on productivity differences seem to assume. Thus the distance between average performance and best practice has increased during the period. This result is confirmed in a related paper, Försund and Hjalmarsson [1978], which studies the development of the distance between the frontier production function and the average production *function*.

Even if the development of the efficiency measures S_0-S_3 is the same, the levels for each year differ rather much. For all years $S_0 > S_1 > S_2 > S_3$, However, the difference between S_2 and S_3 is rather small. This means, which S_5 shows, that if the average plant is moved to the efficiency frontier in the vertical direction rather little is to be gained by moving it to the optimal scale. This stems from the fact that the average amounts of inputs are about the same as required at optimal scale for the first year and have developed in the same way, as the amounts of inputs required at optimal scale.

On the other hand, if the average plant is moved to the frontier in the horizontal direction there still remains some pure scale inefficiency which increases rather much, from .92 to .79 during the period. Thus most units become too small when they are moved horizontally to the frontier, a tendency which is strengthened during the period. While optimal scale has increased from about 49 000 tonnes in 1964 to 99 000 tonnes in 1973 the average output has only increased from 29 000 tonnes to 39 000 tonnes.

The low level of structural efficiency has been confirmed for one year in an earlier study by Carlsson [1972] who estimated S_0 for 26 Swedish industries in 1968 relative to a Cobb-Douglas frontier production function. His estimate of S_0 for the whole dairy industry in this year was 0.6184, not too far from our own estimate that year. Moreover, it turned out that the dairy industry showed the second lowest degree of structural efficiency of the 26 industries. What is then the reason for this high degree of structural inefficiency?

Carlsson [1972] tries to explain the differences in the efficiency between industries by differences in competitive pressure and finds that protection seems to breed inefficiency. Of course, this can be one part of the explanation of efficiency differences. However, if a putty-clay production structure and embodied technical progress are empirically relevant, which seems to be the case in most manufacturing industries (Salter [1960]) there will normally be differences between production units within an industry. As pointed out in a comment on Carlsson's result (Førsund & Hjalmarsson [1974b]) the more rapid the technical progress the less efficient the industry may appear in an analysis based on cross section data as in Carlsson [1972] depending on what happens to investment and the rate of scrapping. Thus, if a faster rate of technical progress increases the differences in efficiency between the best practice plants and the industry average for a given rate of industry output expansion one can as well state that technical progress breeds inefficiency.

The differences in efficiency can be perfectly efficient from an economic point of view, as shown in Johansen [1972] and Førsund & Hjalmarsson [1974a]. Important explanatory factors of industry structure at a point in time are then the forms of the establishment *ex ante* production functions within the industry, the rate of embodied technical progress, and the expansion rate of the industry output.

A main characteristic of the technological structure of dairy plants is that there are different substitution possibilities before and after investments in new production techniques, i.e. one must distinguish between *ex ante* and *ex post* production possibilities (Johansen [1972]). A putty-clay structure, embodied technical progress and economies of scale in plant construction give rise to different vintages of capital.

However, it is not possible at our level of aggregation to identify unique vintages. Technical change is characterized by successive improvements of different parts of the dairies as e.g. changes of milk reception from cans to tanks and introduction of self-cleaning separators.

In Førsund & Hjalmarsson [1977] it is shown that technical progress has been rapid during the period. See also the Appendix below. In fact, average cost at optimal scale decreases progressively from about 9% per year in the beginning of the period to about 13% at the end of the period.

Thus one reason, and probably the most important one, for the large and increasing differences between best-practice technology and average performance must be the underlying technological structure in combination with a rapid technical progress. Further aspects of the efficiency differences will be discussed below.

All plants included in this study have survived the whole period. During the same time a lot of dairies have been closed down in Sweden. Thus, the development of structural efficiency for all plants may have been another than for the set utilized here.

Technical efficiency and scale efficiency

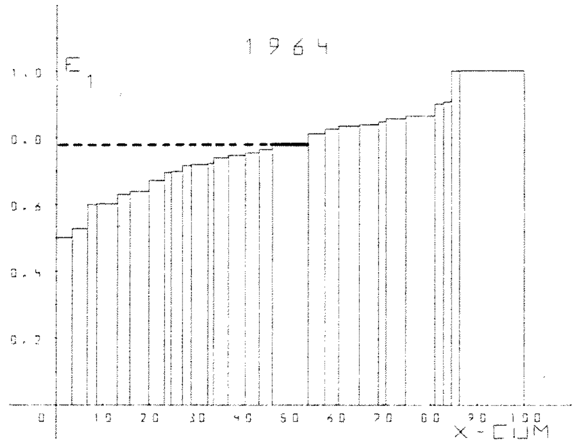
The estimates of the individual measures of technical efficiency and scale efficiency are presented in Fig 3-6 below for three different years, 1964, 1968 and 1973. In the figures (which are computer plotted) the units are arranged in increasing order of their efficiency values. Each rectangle or step in the diagrams represents an individual unit. Efficiency is measured along the ordinate axis and the percentage share of output (accumulated) along the abscissa axis.

In these figures both the range and shape of the efficiency distributions are illustrated. At the same time we can observe the positions of the small and large units.

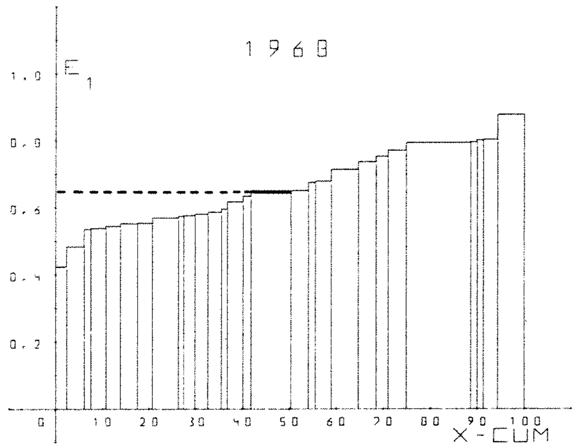
Let us first look at Fig.3-5 where the measures are shown separately. The interpretation of the measures are shown in a few examples.

In 1964 the least efficient unit according to E_1 produces about 3% of total industrial output and has an efficiency value E_1 about .50. This means that the same output could have been produced by 50% of the observed amount of input when utilizing best-practice technology.

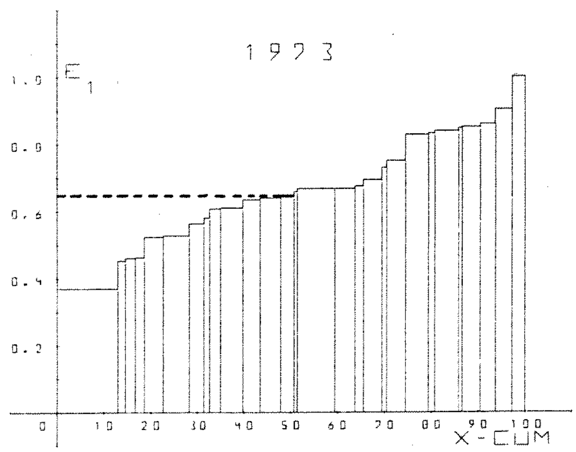
The least efficient unit according to E_2 also produces about 3% of total output and has an efficiency value of E_2 about .46, which means that the observed production is only 46% of the output obtained by employing the same amount of inputs in the frontier function.



(a)

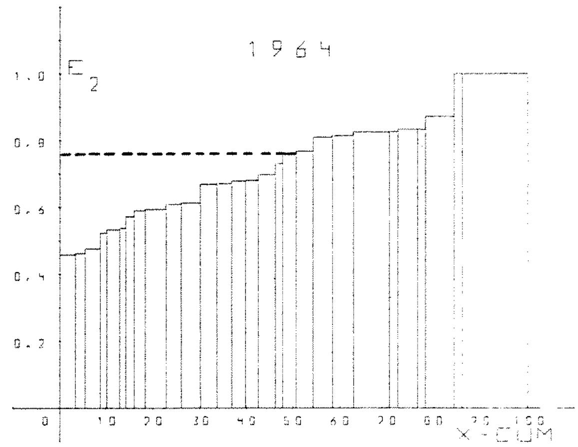


(b)

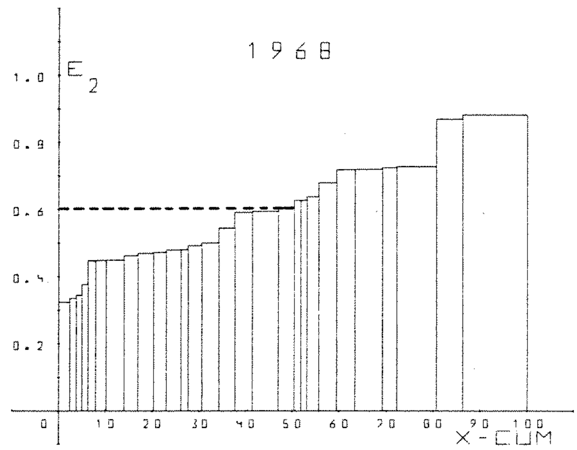


(c)

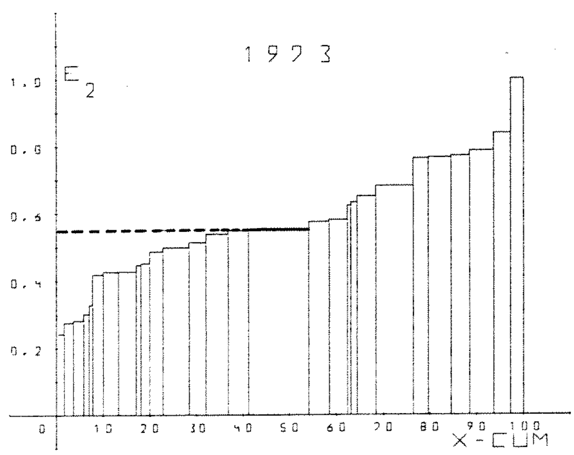
Figure 3. E_1 measures of technical efficiency for selected years.



(a)



(b)



(c)

Figure 4. E_2 measures of technical efficiency for selected years.

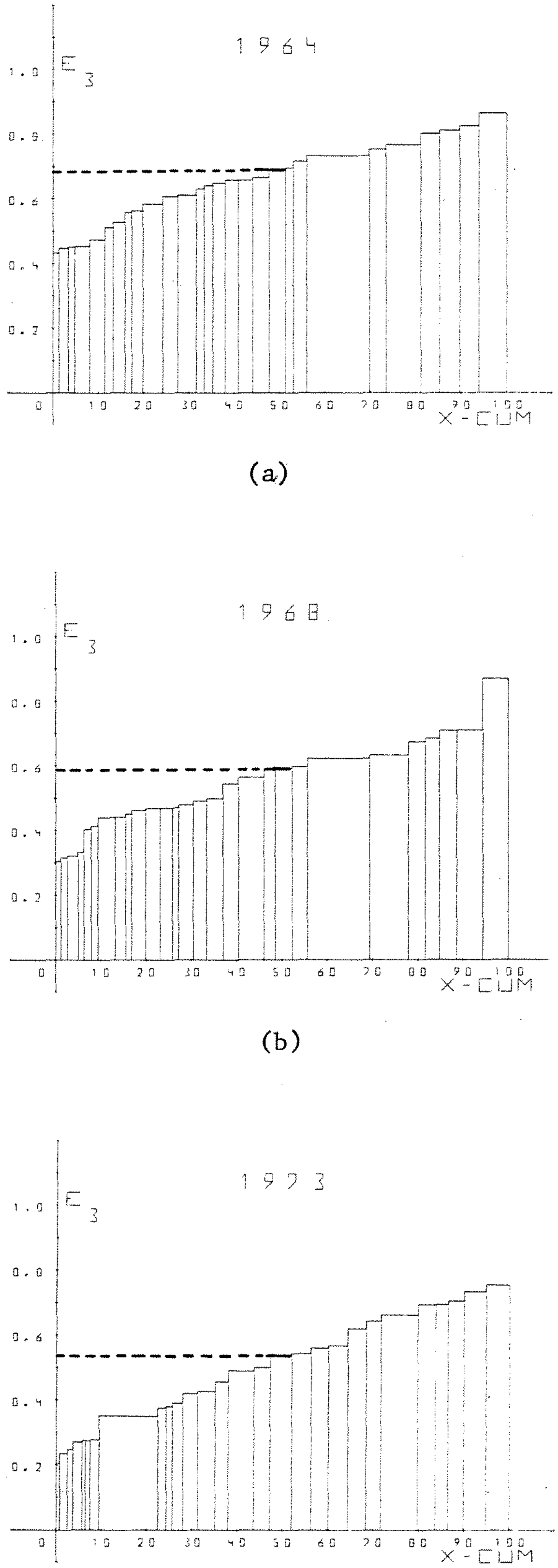


Figure 5. E_3 measures of scale efficiency for selected years.

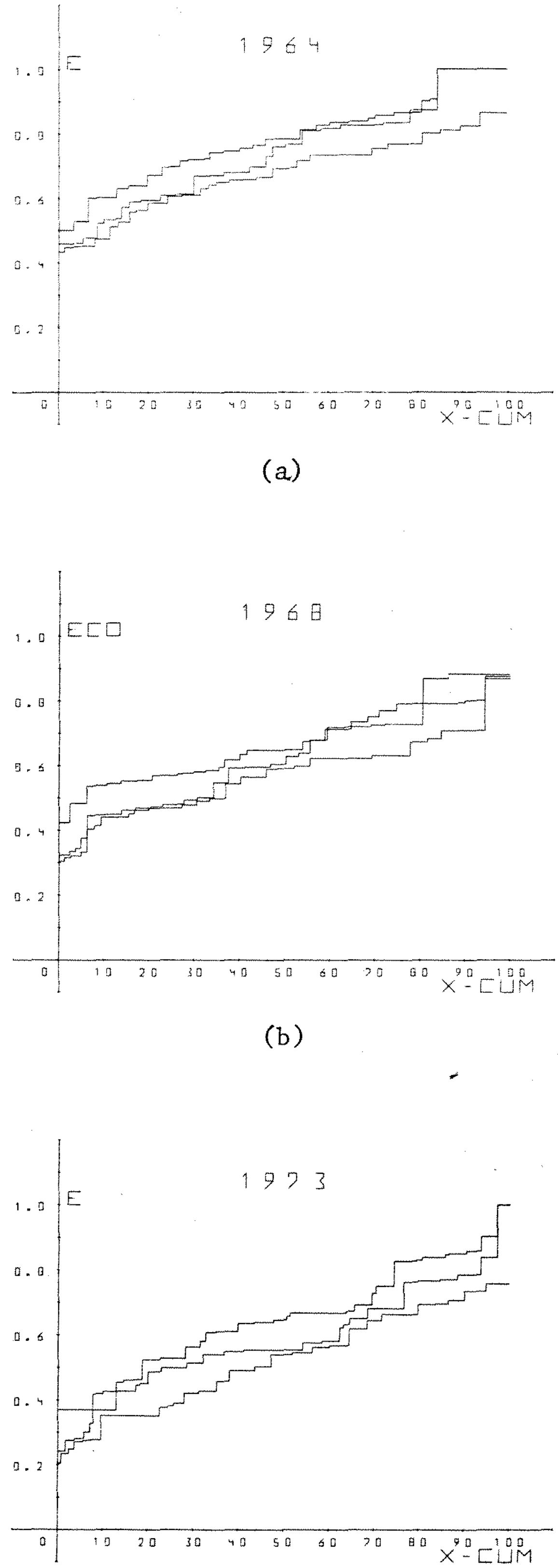


Figure 6. E_1,E_2 and E_3 measures for selected years.

Let us also look at scale efficiency, E_3, in 1973. The least efficient unit with E_3 about .20 produced only about 1% of total output. If this unit had employed frontier function technology at optimal scale, the level of the potential input coefficients would have been only 20% of the actual observed. The most efficient unit that year with E_3 about .76 produced about 5% of total output. The level of its potential input coefficients was then 76% of the actual observed if the observed amount of inputs had been employed at optimal scale in the frontier production function.

As the figures reveal there is a large variation in efficiency between the units for all years. The most striking example here is the E_2-values for 1973 when the most efficient unit was on the frontier ($E_2=1$) and the most inefficient one had a value of $E_2=.24$. Moreover the range increased during the period in consistence with the development of the measures of structural efficiency.

The shape of the distributions also changed during the period. Seen from the left to the right in the figures, efficiency decreases rather continuously in 1964 but in 1973 the efficiency distribution becomes more irregular except for scale efficiency which has a very regular shape during the whole period.

As regards the position of the small and large units in the efficiency distributions there is no clear relationship between size and technical efficiency. In 1963 the largest plant and a very small one were on the production frontier and in 1973 a plant of medium size.

The development of the largest plant is interesting. In 1964 this plant was on the frontier i.e. $E_1=E_2=1$. Also in 1968 this plant was rather efficient but in the last year its efficiency was reduced dramatically especially measured by E_1, but not so much by E_2. A closer look at the data shows that the input coefficients of labour and capital were fairly constant for this unit during the period while the input coefficients for labour decreased for most other units being approximately constant for capital. Thus the productivity of this unit has been fairly constant at the same time as the frontier has moved upwards.

Because the frontier is estimated by LP-techniques the number of units on the frontier are at most equal to the number of estimated parameters, five here. The frontier is also usually built up of plants of different size, one large, one small and a few medium sized. A very small plant with high input coefficients of both labour and capital can be on the frontier because that plant is the most efficient of that size.

The differences in ranking the units according to E_1 (constant output) and E_2 (constant input) are also clearly demonstrated in Figs. 3 and 4. Especially for the largest unit in 1973 the difference is striking. According to E_1 this unit is the most inefficient one, according to E_2 it has about medium efficiency. Thus it is not a matter of indifference which measure is utilized when talking about efficiency for individual plants.

As regards scale efficiency there is a clear tendency for the large units to show high values. An exception is the largest unit in 1973 which has a rather low value of scale efficiency.

A further comprehensive view of the development of the efficiency distributions is obtained in Fig. 6 where all the three measures of efficiency E_1, E_2 and E_3 are plotted at the same time as step functions i.e. the top levels from the histograms are plotted in the same figure, as an alternative to the histograms. The step diagrams give a good picture of the dispersion in the different measures and the ranking of the units according to the different measures.

The total dispersion for all measures is somewhat reduced by the changes in the ranking between the different measures, on the same time as the range increases through time.

An alternative to the measures of structural efficiency above is to look at the the efficiency value of that unit which covers the 50 % accumulated capacity point on the abscissa axis. These values are indicated by dotted lines in Fig. 3-5. This median capacity value of E_1, which is very similar to the value of S_0, has decreased from .79 in 1964 to .65 in 1973. This is about the same percentage decrease as in the S_0 and S_1 measures (about 20 %). The median capacity value of E_2 has decreased from .77 in 1964 to .55 in 1973 which is about the same percentage decrease as in the S_2 measure (about 40 %) but on higher level. The median capacity value of E_3 has decreased from .69 to .54 (28 %) which is a smaller reduction than for S_3 (38 %).

Let us also look more thoroughly at the rankings between different years,of the individual units in the efficiency distributions. We are interested in investigating whether there have been any dramatic changes in the rankings of the units during the period. Thus we have calculated Spearman's rank correlation coefficient between the different years consecutively and between 1964 and 1973 together with Kendall's coefficient of concordance, denoted by W, for the whole period. The results are shown in Table 2 below.

Table 2. Spearman's rank correlation coefficient between different years and Kendall's coefficient of concordance, W.

Years	E_1	E_2	E_3
1964/65	.8544	.7969	.8402
1965/66	.7614	.8199	.9201
1966/67	.8856	.9595	.9625
1967/68	.8681	.8380	.8730
1968/69	.8027	.8210	.9373
1969/70	.7756	.7367	.7983
1970/71	.9146	.8544	.9086
1971/72	.8643	.9135	.9351
1972/73	.8593	.9245	.9688
1964/73	-.0282	.1073	.4072
W	.5429	.6011	.7003

The table reveals a high correlation of efficiency rankings between successive years, and highest for scale efficiency. Usually the correlation coefficient is in the interval between .80 and .95. There has not been any dramatic changes in the efficiency rankings between a pair of years but scale efficiency has been most stable. The value of the coefficient of concordance is rather high, but somewhat lower than the correlation coefficients for successive years, also indicating a high stability in the rankings.

On the other hand, there has been a gradual change in the rankings during the period, relatively small for scale efficiency but large for technical efficiency. The correlation coefficient between the start and end years 1964 och 1973 even shows a negative sign for E_1. An example here is the largest unit which was on the frontier in 1964 but had the lowest E_1 value in 1973. The lower values of the coefficient of concordance, in comparison with the correlation coefficients for successive years, also indicate this gradual change of the rankings.

We have also confronted the dairy experts of the Swedish Dairy Federation with our empirical results and discussed the reasons for differences in efficiency between the units.

We got a confirmation that our results regarding the most and least efficient plants were reasonable. Some differences in efficiency were explained by the modernity of equipment while others were explained by more or less skilful managements (degree of X-efficiency). With some simplification the small best-practice plants seemed to have good managements while large efficient plants also had modern equipment.

8. Concluding remarks.

In this paper Farrell's measures of productive efficency have been elaborat and generalized to inhomogeneous production functions. Several new measures of efficiency have been introduced and applied to the Swedish milk processing industry. The development of the industrial structure is studied by the change in the efficiency distributions for the individu plants through time and the aggregate performance of the sector is studied by examining the development of the different measures of structural efficiency.

The most remarkable result is the rather high distance between best-practice and average performance measured by different measures of structural efficiency. Moreover, this distance shows an increasing trend during the period. These results are explained by rapid technical progress in combination with an underlying putty clay technological structure and a slow growth of investment.

The distribution of the individual measures of technical efficiency ans scale efficiency reveals a large variation in efficiency between the units for all years. Some of these differences in efficiency can be explained by the modernity of equipment and others by differences in management capability.

Appendix

As regards the form of the production function the following specification is employed (cf Zellner & Revankar [1966]):

$$x^{\alpha(t)} e^{\beta(t)x} = A(t) \prod_{j=1}^{2} v_j^{a_j(t)} \qquad \text{(A1)}$$

Technical change is accounted for by specifying the possibility of changes in the constant term, A, and the kernel elasticities, a_j, for labour, L, and capital, K, and the scale function parameters α, β.

The corresponding elasticity of scale function is:

$$\varepsilon(x,t) = \frac{1}{\alpha(t)+\beta(t)x} \qquad \text{(A2)}$$

Frontier estimates

The estimated frontier production function is

$$x^{0.32-0.0056\cdot t} e^{(1.47-0.073\cdot t)\cdot 10^{-5}\cdot x} = 0.0024\ L^{0.81+0.0019\cdot t}\cdot\ K^{0.19-0.0019\cdot t}$$

t=1 in 1964,

t=10 in 1973

The elasticity of scale function is

$$\varepsilon(x,t) = \frac{1}{0.32-0.0056\cdot t + (1.47-0.0073\cdot t)\cdot 10^{-5}\cdot x}.$$

This means that optimal scale, x for $\varepsilon=1$, increases from 48 644 in 1964 to 99 325 in 1973.

The specification of the derived efficiency measures

To simplify the notation, the production function specified in (A1) can be written as

$$x^{\alpha}e^{\beta x} = A \prod_i v_i^{a_j} \tag{A3}$$

where α, β, A, and a_j are all functions of time as indicated in (A1).

If optimal scale $\hat{x} = (1-\alpha)/\beta$ ((A2) is solved for x with $\varepsilon=1$) is inserted in (A3) considering that

$\xi_i = \frac{v_i}{x}$ the efficiency frontier is obtained as

$$(\xi_1^{a_1},\ldots,\xi_n^{a_n}) \cdot A \cdot \left(\frac{e\beta}{1-\alpha}\right)^{\alpha-1} = 1 \tag{A4}$$

The following efficiency measures are then derived (v_i^o, x^o, and ξ_i^o denote actual observations):

$$E_1 = \frac{x^{o\alpha} e^{\beta x^o}}{A \cdot \pi_i (v_i^o)^{a_i}} \quad \text{and} \tag{A5}$$

$$E_2 = \frac{x^o}{x^*} \quad \text{where } x^* \text{ is the solution of } x^{\alpha}e^{\beta x} = A \cdot \pi_i (v_i^o)^{a_i} \tag{A6}$$

$$E_3 = \frac{\left(\frac{e\beta}{1-\alpha}\right)^{1-\alpha}}{A \cdot \pi_i (\xi_i^o)^{a_i}} \tag{A7}$$

$$S_0 = \sum_j E_{1j} \frac{x_j^o}{\sum_j x_j^o} \tag{A8}$$

$$S_1 = \frac{\left(\frac{1}{n}\sum_j x_j^o\right)^{\alpha} e^{\beta\left(\frac{1}{n}\sum x_j^o\right)}}{A \cdot \pi_i \left(\frac{1}{n}\sum v_{ij}^o\right)^{a_i}} \tag{A9}$$

$S_2 = \frac{\bar{x}^o}{\bar{x}^*}$ where $\bar{x}^o$ is observed average production and $\bar{x}^*$ is obtained as the solution of $x^{\alpha}e^{\beta x} = A \cdot \pi_i \left(\frac{1}{n}\sum v_{ij}^o\right)^{a_i}$ (A10)

$$S_3 = \frac{\left(\frac{e\beta}{1-\alpha}\right)^{1-\alpha}}{A\pi_i \left(\frac{1}{n}\sum_j v_{ij}^o \Big/ \frac{1}{n}\sum_j x_j^o\right)^{a_i}} \tag{A11}$$

REFERENCES

Aigner, D.J. and Chu, S.F. (1968). "On Estimating the Industry Production Function." *American Economic Review*, vol. 58, pp. 226-39.

Carlsson, B. (1972). "The Measurement of Efficiency in Production: An Application to Swedish Manufacturing Industries,1968." *Swedish Journal of Economics*, vol. 74, pp. 468-85.

Todd, D. (1971). "The Relative Efficiency of Small and Large Firms." *Research Report No 18, Committee of Inquiry on Small Firms*. Her Majesty's Stationary Office. London.

Farrell, M.J. (1957). "The Measurement of Productive Efficiency." *Journal of the Royal Statistical Society*, Series A, 120, pp. 253-90.

Farrel, M.J. and Fieldhouse, M. (1962). "Estimating Efficient Production Functions under Increasing Returns to Scale." *Journal of the Royal Statistical Society*, Series A, 125, pp. 252-67.

Frisch, R. (1965). *Theory of Production*. Dordrecht: D. Reidel Publishing Co.

Førsund, F.R. (1971). "A Note on the Technically Optimal Scale in Inhomogeneous Production Functions." *Swedish Journal of Economics*, vol. 73, pp. 225-40.

Førsund, F.R. and Hjalmarsson, L. (1974a). "On the Measurement of Productive Efficiency." *Swedish Journal of Economics*, vol. 76, pp. 141-54.

— (1974b). "Comment on Bo Carlsson's 'The Measurement of Efficiency in Production: An Application to Swedish Manufacturing Industries, 1968." *Swedish Journal of Economics*, vol. 76, pp. 251-54.

— (1976). "Structural Change and Economic Efficiency of Swedish Dairy Plants." Mimeographed . (Sept.)

— (1977). "Frontier Production Functions and Technical Progress: A Study of General Milk Processing in Swedish Dairy Plants." The Industrial Institute for Economic and Social Research, Stockholm, forthcoming in *Econometrica*.

— (1978). "Technical Progress and Structural Efficiency of Swedish Dairy Plants." In *Capital in the Production Function*. Institut de Recherches en Économic de la Production, Paris X - Nanterre. *Forthcoming*.

Førsund, F.R. and Jansen, E.S. (1977). "On Estimating Average and Best Practice Homothetic Production Functions via Cost Functions." *International Economic Review*, vol. 18, pp. 463-76.

Hjalmarsson, L. (1973). "Optimal Structural Change and Related Concepts." *Swedish Journal of Economics*, vol. 75, pp. 176-92.

Johansen, L. (1972). *Production Functions*. Amsterdam: North-Holland.

Johansen, L. and Sørsveen, Å. (1967). "Notes on the Measurement of Real Capital in Relation to Economic Planning Models." *Review of Income and Wealth*, vol. 13, pp. 175-97.

Komiya, R. (1962). "Technological Progress and the Production Function in the United States Steam Power Industry." *Review of Economics and Statistics*, vol. 44, pp. 156-66.

Meeusen, W. and Van den Broeck, J. (1977). "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error." *International Economic Review*, vol. 18, pp. 435-44.

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

Salter, W.E.G. (1960). *Productivity and Technical Change*. Cambridge: University Press.

Sato, R. (1970). "The Estimation of Biased Technical Progress and the Production Function." *International Economic Review*, vol. 11, pp. 179-208.

Schmidt, P. and Lovell, C.A.K. (1977). "Estimating Technical and Allocative Inefficiency Relative to Stochastic Production and Cost Frontiers." Mimeographed (May).

Seitz, W.D. (1970). "The Measurement of Efficiency Relative to a Frontier Production Function." *American Journal of Agricultural Economics*, vol. 52, pp. 505-11.

—— (1971). "Productive Efficiency in the Steam-Electric Generating Industry." *Journal of Political Economy*, vol. 79, pp. 878-86.

Timmer, C.P. (1971). "Using a Probabilistic Frontier Production Function to Measure Technical Efficiency." *Journal of Political Economy*, vol. 79, pp. 776-94.

Zellner, A. and Revankar, N.S. (1969). "Generalized Production Functions." *Review of Economic Studies*, vol. 36, pp. 241-50.