The Average Efficiency of Firms within an Industry: An Application of Data Envelopment Analysis

Bert M. Balk* Paul M. C. de Boer† Sonja Greve‡

Draft, July 20, 2000

Abstract

Firms acting according to the same production technology usually exhibit differing degrees of efficiency. For various purposes it is interesting to obtain insight into the cross-sectional dispersion of these efficiency scores and its evolution over time. It is thereby important to employ a method that rests on a minimal number of assumptions. In this paper we show how this task can be performed by Data Envelopment Analysis (DEA).

The three basic models will be presented, and two of them will be applied to the Dutch rubber-processing industry to obtain technical and scale efficiency scores of the individual firms.

Keywords: Production frontiers; technical efficiency; scale efficiency; Data Envelopment Analysis.

---

* Department of Statistical Methods, Statistics Netherlands, P. O. Box 4000, 2270 JM Voorburg, The Netherlands, e-mail: bblk@cbs.nl. The views expressed in this paper are those of the author and do not necessarily reflect the policies of Statistics Netherlands.

† Econometric Institute, Erasmus University Rotterdam, P. O. Box 1738, 3000 DR Rotterdam, The Netherlands, e-mail: pmdeboer@few.eur.nl, and GAME, Université Libre de Bruxelles, Belgium.

‡ Schere 114, 3085 DW Rotterdam, The Netherlands.

1 Introduction

Firms acting according to the same production technology usually exhibit differing degrees of efficiency. For various purposes it is interesting to obtain insight into the cross-sectional dispersion of these efficiency scores and its evolution over time. It is thereby important to employ a method that rests on a minimal number of assumptions. In this paper we show how this task can be performed by Data Envelopment Analysis (DEA). But let us first explain what is to be understood by production technology and efficiency.

The production technology that is relevant for a certain group of firms[1] consists of all the feasible combinations of input quantities and output quantities. Formally,

$$T = \{(x, y) : x \text{ can produce } y\}, \tag{1}$$

where $x \geq 0$ denotes a vector of input quantities and $y \geq 0$ denotes a vector of output quantities. This set is assumed to be closed. A production frontier, or isoquant, consists of all x that can produce a certain y in a (technically) efficient way, that is without any waste. In the case of a single output the set of frontiers can be represented by a production function $y = f(x)$. Figure 1 depicts such a function for a single-input single-output technology.

[Insert Figure 1]

Figure 1. Firm A and the production frontier

Consider a firm A that produces output quantities y_A and uses input quantities x_A. Firm A is not efficient from the input-oriented point of view, because it could produce the same output quantities y_A by contracting its input quantities by the factor OX_1/OX_A. It is not efficient from the output-oriented point of view, because it could expand its output quantities by the factor OY_1/OY_A while still using the same input quantities x_A. In this paper we will use the input-oriented view.[2]

Next, consider a firm F that produces one output with the aid of two inputs. Let $y = f(x_1, x_2)$ denote the production function and let w_1 and

[1] More generally, "Decision Making Units".

[2] For a treatment of the output-oriented view, as well as for more details on the input-oriented point of view, we refer to Coelli et al. (1998) and Balk (1998).

w_2 denote the (exogenously given) input prices. In Figure 2 II' denotes the frontier $y_F = f(x_1, x_2)$, and $C_F = w_1x_1 + w_2x_2$ the isocost line of firm F.

[Insert Figure 2]

Figure 2. Firm F and its production frontier

Following Farrell (1957), technical efficiency is measured along a ray[3] from the origin to the observed point F. The input quantities can be contracted by a factor $TE = OB/OF$ to produce the same output quantity without any waste, i.e. TE measures the technical efficiency of firm F. The line C_B is the isocost line of B. But operating at point B is not efficient from an economic point of view. A firm operating at point D is not only technically efficient, but produces the output quantity also at minimal cost. This firm is the cost efficient firm. Its isocost line is C_D. The so-called allocative efficiency of firm F is measured by $AE = OG/OB$, and the cost efficiency by $CE = OG/OF$. Obviously, in this framework $CE = AE \times TE$. In terms of associated costs, we have:

$$TE = \frac{OC_B}{OC_F} \qquad AE = \frac{OC_D}{OC_B} \qquad CE = \frac{OC_D}{OC_F}.$$

It is clear that in order to measure the technical, allocative, and cost efficiency of firms we need an estimate of the production frontiers (or, in the single-output case, of the production function). The various approaches can be classified into two groups:

1. The parametric approach, where it is assumed that the production frontier has a specific functional form which depends on a finite number of parameters, such as the Cobb-Douglas, CES, HCDES (de Boer and Harkema, 1993) or translog functional form (Lesuis and de Boer, 1994, and Lesuis et al., 1996);

2. The non-parametric approach, where such an assumption is not made. In this paper we shall consider one such approach, called Data Envelopment Analysis (DEA).

[3] The advantage of radial efficiency is that it is invariant with respect to the units of measurement.

Before introducing DEA, however, we will briefly discuss the parametric approach. It is assumed that there is a continuum of firms, all facing the same production frontiers, say, for ease of exposition, having the Cobb- Douglas functional form:

$$\ln(y) = \beta_1 + \beta_2 \ln(x_1) + \beta_3 \ln(x_2) + \varepsilon. \tag{2}$$

We dispose of data on n firms which constitute a sample from the population facing (2). In order to estimate the parameters β_1, β_2 and β_3, one might use the method of least squares. In this case, the disturbance term ε is commonly assumed to be normally distributed. For more complicated functional forms, one often has to resort to the method of maximum likelihood, necessitating the assumption of normality.

When using this method, one implicitly assumes that all n firms operate in a technically efficient way, since ε only accounts for "statistical" noise.

This method has been criticized by Aigner, Lovell and Schmidt (1977), and, independently, by Meeusen and Van den Broeck (1977). They suggested to decompose ε as

$$\varepsilon = v - u \tag{3}$$

where v is a symmetrical disturbance term, accounting for "statistical" noise, such as measurement error and other possible random influences, and u is an one-sided disturbance term that accounts for inefficiency ("economical" noise); u and v are assumed to be independently distributed. Usually, v is assumed to be distributed as $N(0, \sigma_v^2)$, and u is commonly assumed to be half-normally distributed.[4] This method is known as Stochastic Frontier Analysis (SFA).

In Figure 3 we give an idea of the difference between the two methods in the simple case of one output and one input. The acronym "LS" stands for "least squares" and "SF" for an SFA model.

[Insert Figure 3]

Figure 3. Least squares versus SF

[4] Other distributions that have been considered are the exponential, the truncated normal, and the two-parameter Gamma distribution. For more details we refer to Kumbhakar and Lovell (2000).

The econometric estimation of this type of models, however, is computationally burdensome and can lead to estimates that are not satisfactory. When using the half-normal distribution, for instance, one frequently encounters the problem that the estimated value of the variance, σ_u^2, is negative. See for instance Brouwer (1999).

A school of researchers moreover feel that the imposition of a specific functional form to the production frontiers as well as a specific probability distribution, such as the half-normal, is unduly restrictive. They propagate DEA as a non-parametric alternative.

By way of example, in this paper we apply DEA to show how to assess the performance of firms in a certain industry. To that end we present in section 2 the basic models; in section 3 we discuss the data used as well as some results, whereas section 4 concludes.

2 Data Envelopment Analysis: the Basic Models

Let be given a population of firms that produce M-dimensional output quantity vectors $y \geq 0$ with K-dimensional input quantity vectors $x \geq 0$, governed by the closed production possibility set T defined in (1). We dispose of data on a sample D of n firms.

Define the $K \times n$ matrix X, whose columns are the input quantity vectors x_i, and the $M \times n$ matrix Y, whose columns are the output quantity vectors y_i $(i \in D)$.

The principle of DEA is to compare the performance of firm i $(i \in D)$ to that of a linear combination of other firms. That linear combination will be called the "reference unit" in the sequel. This implies that the input quantities of firm i should not be smaller than those of the reference unit, and that the output quantities should not be larger than those of the reference unit. Mathematically:

$$x_i \geqq X\lambda \text{ and } y_i \leqq Y\lambda, \tag{4}$$

where $\lambda' = (\lambda_1 . \cdots, \lambda_n)$. With respect to T we assume:

Assumption 1: "Monotonicity" or "disposability".

If $(x, y) \in T$, and $x' \geqq x \vee y' \leqq y$, then $(x', y') \in T$.

Assumption 2: "Data envelopment".

$$(x_i, y_i) \in T \ (i \in D).$$

Imposition of assumptions 1 and 2 leads to the Full Disposable Hull (FDH) model due to Deprins, Simar and Tulkens (1984). In Figure 4 we depict the FDH model for the case of one output and one input.

[Insert Figure 4]

Figure 4. The production possibility set according to FDH

In the FDH model there is only one reference unit: either one other firm, or, in case the firm lies on the production frontier, the firm itself. In the latter case the firm is, according to the FDH model, efficient. Mathematically this means that $\lambda_i \in \{0, 1\}$ and $\iota'\lambda = 1$, where $\iota' = (1, \cdots, 1)$ is the summation vector.

Consequently, the estimator of the production possibility set by the FDH model is given by:

$$T_{FDH} = \{(x, y) \mid x \geqq X\lambda; y \leqq Y\lambda; \iota'\lambda = 1; \lambda_i \in \{0, 1\}; i \in D\}. \qquad (5)$$

Assumption 3: "Convexity".

If $(x_j, y_j) \in T$ where $j \in F \subseteq D$, and if $\sum_F \mu_j = 1$ $(\mu_j \geq 0)$ then $(\sum_F \mu_j x_j, \sum_F \mu_j y_j) \in T$.

Assumptions 1, 2, and 3 lead to the model proposed by Banker, Charnes and Cooper (1984). It is called the BCC model and is characterized by variable returns to scale (VRS). In Figure 5 we depict the model for the case of one output and one input.

[Insert Figure 5]

Figure 5. The production possibility set according to BCC

The estimator of the production possibility set by means of the BCC model is given by:

$$T_{BCC} = \{(x, y) \mid x \geqq X\lambda; y \leqq Y\lambda; \lambda_i \geq 0; \iota'\lambda = 1; i \in D\}. \qquad (6)$$

Assumption 4: "Ray unboundedness".

If $(x,y) \in T$, then $(kx,ky) \in T$ for $k \geq 0$.

The assumptions 1, 2 and 4 lead to the model due to Charnes, Cooper and Rhodes (1978). It is called the CCR model and is characterized by constant returns to scale (CRS). In Figure 6 we depict the model for the case of one output and one input.

[Insert figure 6]

Figure 6. The production possibility set according to CCR

The estimator of the production possibility set according to the CCR model is

$$T_{CCR} = \{(x,y) \mid x \geqq X\lambda; y \leqq Y\lambda; \lambda_i \geq 0, i \in D\}. \tag{7}$$

In order to estimate the frontiers (5)-(7) we need an estimate of the vector of parameters λ.

As mentioned in the introduction we use the Farrell (1957) measure of technical efficiency which, for the input-oriented view, reads:

$$\theta(x,y) = \min\{\theta : (\theta x, y) \in T\}. \tag{8}$$

In order to estimate this efficiency measure we have, for each firm i, to solve the following linear programming (LP) problem:

$$\min_{\lambda} \theta \tag{9}$$

subject to

$$y_i \leqq Y\lambda \tag{10}$$

$$\theta x_i \geqq X\lambda \tag{11}$$

$$\lambda_{i'} \geq 0 \; (i' \in D). \tag{12}$$

Solving these problems means that we are imposing the CCR model. When we additionally impose[5]

$$\iota'\lambda = 1, \tag{13}$$

we have the BCC model, while imposing additionally

$$\lambda_{i'} \in \{0,1\}\ (i' \in D) \tag{14}$$

we have the FDH model.

Let $\hat{\theta}_{i,j}$ denote the solution of the LP problem for firm $i \in D$ according to model j (j = CCR, BCC, or FDH), then $\hat{\theta}_{i,j}$ is the estimate of the technical efficiency of firm i according to model j:

$$TE_{i,j} = \hat{\theta}_{i,j}\ (i \in D). \tag{15}$$

Because BCC is a restricted version of CCR , and FDH in turn of BCC, it follows that

$$\hat{\theta}_{i,CCR} \leq \hat{\theta}_{i,BCC} \leq \hat{\theta}_{i,FDH}. \tag{16}$$

When CCR (constant returns to scale) and BCC (variable returns to scale) yield the same technical efficiency score, then the firm is said to be *scale efficient*. Therefore, a natural estimate of scale efficiency is

$$SE_i = \frac{\hat{\theta}_{i,CCR}}{\hat{\theta}_{i,BCC}} \leq 1. \tag{17}$$

In order to measure *cost efficiency*, the function to be minimized becomes instead of (9)

$$\min_{\lambda,x^*} w'x^* \tag{18}$$

whereas the constraint (11) is replaced by

$$x^* \geqq X\lambda. \tag{19}$$

Let $\hat{x}^*_{i,j}$ denote the solution of the LP problem for firm $i \in D$ according to model j (j = CCR, BBC, or FDH), and let x_i denote the actual input

[5]Imposition of $\iota'\lambda \leq 1$ leads to a model exhibiting non-increasing returns to scale.

quantity vector, then the cost efficiency of firm i according to model j is estimated by

$$CE_{i,j} = \frac{w'\hat{x}^*_{i,j}}{w'x_i} \; (i \in D). \tag{20}$$

The allocative efficiency of firm i according to model j is finally estimated by

$$AE_{i,j} = CE_{i,j}/TE_{i,j} \; (i \in D). \tag{21}$$

3 Application

We work with a balanced panel of 18 Dutch firms, classified as belonging to the rubber-processing industry, over the period 1978-1992. Averaged over the whole time period they account for over 80% of the industry's value added.

The various outputs are aggregated to a single one, namely the deflated money value of gross output. The inputs are aggregated to three types: materials (including energy), labour, and capital.

The money value of gross output, of materials and labour come from the yearly production surveys of Statistics Netherlands. In order to arrive at volumes (at 1980 prices), the value of gross output is deflated by firm-specific weighted averages of sectoral price index numbers for inland sales and sales abroad; the labour cost (wage bill including social security contributions) by sectoral index numbers of contractual wages; and materials cost by appropriate price index numbers.

User costs of capital and the price of capital are calculated by a method that makes use of firm and time specific depreciation costs, capital goods scrapping rates, corporate tax rates, interest rates, and price index numbers of investment goods. For details, we refer to Greve (1998).

We notice that the number of firms is rather small. Since the application of the FDH model requires quite some observations (otherwise almost any firm will be classified as being efficient, see Post 1999), we refrain from applying FDH.

For each year and each firm we have calculated the technical efficiency score according to the CCR-model (CRS), and according to the BCC-model (VRS). From these scores we derived the scale efficiencies as well. In Table 1 we summarize our findings.

Table 1: Technical and scale efficiency in the Dutch rubber-processing industry, 1978-1992

Year	Technical efficiency						Scale efficiency		
	CCR-model (CRS)			BCC-model (VRS)					
	Min	Mean	Neff	Min	Mean	Neff	Min	Mean	Neff
1978	0.711	0.885	5	0.828	0.960	12	0.737	0.923	5
1979	0.730	0.913	4	0.857	0.978	12	0.730	0.935	4
1980	0.670	0.883	5	0.778	0.941	9	0.692	0.941	7
1981	0.808	0.929	6	0.846	0.967	11	0.872	0.961	6
1982	0.752	0.921	6	0.848	0.970	11	0.795	0.948	7
1983	0.776	0.946	5	0.941	0.989	13	0.776	0.957	5
1984	0.806	0.952	10	0.897	0.990	15	0.806	0.962	11
1985	0.869	0.962	7	0.926	0.995	15	0.869	0.967	7
1986	0.710	0.947	6	0.711	0.977	14	0.838	0.970	6
1987	0.872	0.959	6	0.919	0.986	12	0.872	0.973	6
1988	0.796	0.926	7	0.879	0.959	10	0.819	0.967	7
1989	0.763	0.933	6	0.837	0.957	9	0.835	0.975	6
1990	0.718	0.913	6	0.738	0.952	8	0.859	0.960	6
1991	0.846	0.952	5	0.855	0.967	9	0.907	0.985	7
1992	0.871	0.949	7	0.876	0.970	10	0.930	0.979	7

Min = minimum efficiency score; Mean = arithmetic average of the efficiency scores; Neff = number of efficient firms (efficiency score 1)

Efficiencies according to the CCR-model

We observe that the minimum efficiency score varies considerably; from 0.670 in 1980 (which means that that particular firm could have realized a gain of 33% in terms of a reduction in inputs) to 0.872 in 1987. The mean scores vary from 0.883 in 1980 to 0.962 in 1985; the average over the whole period being 0.933. It follows that the firms, on average, show a rather high level of efficiency, which is to be expected since the firms that are considered, are survivors. The number of efficient firms is the lowest in 1979 (4) and the highest in 1984 (10), with an average of 6.1 over the whole period.

Efficiencies according to the BCC-model

As noted above there is an additional restriction imposed to the minimization problem so that the scores according to the BCC model are higher than those resulting from the CCR-model, see (16). The minimum efficiency score shows almost the same variation as before: from 0.711 in 1986 to 0.941 in 1983, but the mean score varies from 0.941 in 1980 to 0.995 in 1985, and exhibits a low variation about the average of 0.971 over the whole period. The number of efficient firms is the lowest in 1990 (8) and the highest in 1984 and 1985 (15). On the average 11.3 firms are efficient according to the BCC-model.

Scale efficiencies

We observe that the minimum scale efficiency and the mean increase over the time period considered. The number of scale efficient firms is the lowest in 1979 (4), and peaks in 1984 (11).

4 Conclusion

In this paper we have used a non-parametric technique, Data Envelopment Analysis, in order to obtain insight into the location and dispersion of technical and scale efficiencies of the firms belonging to a certain industry. The technique is inexpensive and rest on a minimal set of assumptions. The information it provides constitutes a welcome addition to the usual measures of sectoral performance.

References

Aigner, D.J., Lovell, C.A.K. & Schmidt, P. (1977) Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics*, 6, 21-37.

Balk, B.M. (1998) *Industrial price, quantity, and productivity indices, the micro-economic theory and an application*, Kluwer Academic Publishers, Boston / Dordrecht / London.

Banker, R.D., Charnes, A. & Cooper, W.W. (1984) Some models for estimating technical and scale efficiencies in data envelopment analysis, *Management Science*, 30, 1078-1092.

de Boer, P.M.C. & Harkema, R. (1993) Estimating input-output factor demand equations; some results of Hanoch's HCDES production function, *Economic Systems Research*, 5, 355-363.

Brouwer, P. (1999) Productivity and efficiency in the Dutch construction industry, Research paper no. 9916, Department of Statistical Methods, Statistics Netherlands.

Charnes, A., Cooper, W.W. & Rhodes, E. (1978) Measuring the efficiency of decision making units, *European Journal of Operational Research*, 2, 429-444.

Coelli, T., Rao, D.S.P. & Battese, G.E. (1998) *An introduction to efficiency and productivity analysis*, Kluwer Academic Publishers, Boston / Dordrecht / London.

Deprins, D., Simar, L. & Tulkens, H. (1984) Measuring labour efficiency in post-offices, in: M. Marchand, P. Pestieau and H. Tulkens (eds.), *The performance of public enterprises*, North Holland, Amsterdam, 243-267.

Farrell, M.J. (1957) The measurement of productive efficiency, *Journal of the Royal Statistical Society Series A*, 120, 253-290.

Greve, S. (1998) Efficiency and productivity measurement, an application, Research paper 9813, Department of Statistical Methods, Statistics Netherlands.

Kumbhakar, S. C. & Lovell, C. A. K. (2000) *Stochastic Frontier Analysis*, Cambridge University Press, Cambridge / New York / Melbourne / Madrid.

Lesuis, P. J. J. & de Boer, P. M. C. (1994)

Lesuis, P.J.J., de Boer, P.M.C., Harkema, R. & Hobijn, B. (1996) Structural models of factor demands and of technological change: an empirical assessment of dynamical adjustment specifications for sectors of the Dutch economy, *Economic Systems Research*, 8, 341-360.

Meeusen, W. & van den Broeck, J. (1977) Efficiency estimation from Cobb-Douglas production functions with composed error, *International Economic Review*, 18, 435-444.

Post, G.T. (1999) *Finding the frontier: methodological advances in Data Envelopment Analysis*, Tinbergen Institute Research Series, vol. 211, Rotterdam, The Netherlands.