Aus dem Institut für Betriebswirtschaft, Agrarstruktur und ländliche Räume

Alexander Gocht
Kelvin Balcombe

On the quality of DEA estimates

Manuskript, zu finden in www.fal.de

Braunschweig
Bundesforschungsanstalt für Landwirtschaft (FAL)
2003

On the quality of DEA estimates

Alexander Gocht [a)], Kelvin Balcombe [b)]

[a)] Federal Agricultural Research Centre (FAL), Braunschweig, Germany

[b)] Imperial College London, Wye, England

Abstract

This paper explores the reliability of Data Envelope Analysis (DEA) within Agriculture. Particular interest is directed towards the impacts of using constant (CRS) or variable returns to scale (VRS) along with choices regarding data aggregation using a data set of Slovenian farms. Statistical inference is implemented by using the smoothed homogenous bootstrap procedure introduced by SIMAR and WILSON (2000). The coefficient of separation (CoS), a statistic that indicates the degree of statistical differentiation within the sample, is used to demonstrate the findings. The CoS suggests a substantive dependency of the results on the methodology and assumptions that are employed. Accordingly, some observations are made about how to conduct DEA in order to get more reliable efficiency rankings, depending on the purpose for which they are to be used. In addition, attention is drawn to the ability of the SLICE MODEL, implemented in GAMS, to enable researchers to overcome the computational burdens of conducting DEA (with bootstrapping) in large samples.

Keywords: Data Envelopment Analysis, DEA, bootstrapping, Agriculture, Technical Efficiency, Confidence Intervals, Slice DEA model, GAMS

1. Introduction

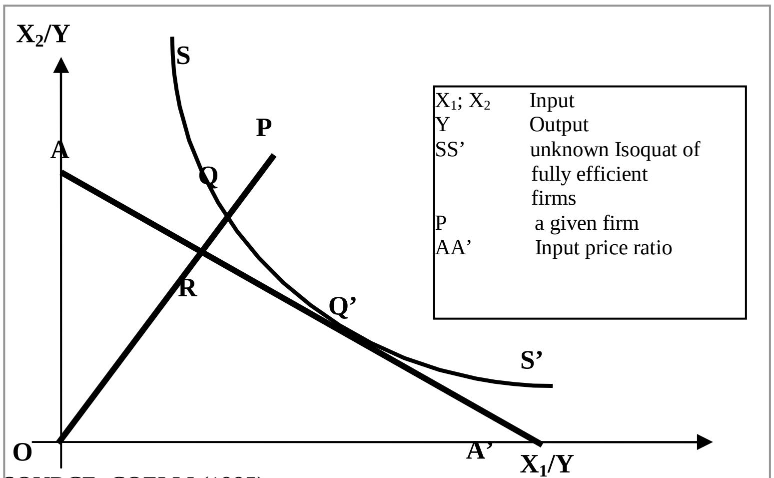
Data Envelopment analysis is a potentially useful technique for measuring efficiency. However, some fundamental concerns need to be addressed before DEA can be accepted as a routine tool. In this paper we investigate the robustness of DEA efficiency subject to the assumption of returns to scale and data aggregation. Technical efficiency scores are calculated applying DEA to 69 decision making units (DMUs). Different assumptions regarding returns to scale and input aggregation are made. The purpose here is to show how heavily these assumptions influence the ranking of the DMUs. Because the underlying frontier model is nonparametric, bootstrapping is required to evaluate the statistical properties of the estimates. Thus, the SIMAR and WILSON smoothed homogeneous bootstrap procedure is used to calculate bias, variance and confidence intervals for the attained efficiency scores. Based on the confidence intervals for the efficiency scores, it is demonstrated how the choice of input aggregation and returns to scale lead to quite different DMU rankings. A Slovenian data set will serve as background upon which these issues will be discussed.

The paper is structured as follows: DEA is compared with the stochastic frontier approach (SFA) to get deeper understanding why DEA is a real competitor. The literature on Data Envelopment Analysis applications in agriculture is then reviewed in order to highlight the heterogeneous specifications employed in previous studies. It is observed that little comment has been made concerning the sensitivity of the results to alternative specifications. In the next section, different model specifications for Slovenian farm data are calculated and the bootstrap procedure from SILMAR and WILSON (2000a) is applied in order to obtain confidence intervals. The findings are discussed and along with the implications for the practical implementation of DEA. Conclusions are drawn and areas requiring further investigation are identified.

2 Defining Efficiency

The concept of economic efficiency is generally assumed to consist of two components: technical efficiency and allocative efficiency. Broadly, the former is defined as the capacity and willingness of an economic unit to produce the maximum possible output from a given bundle of inputs and technology. The latter is defined as the ability and willingness of an economic unit to equate its specific marginal value product with its marginal cost. FARRELL developed an isoquant method in 1957 to measure efficiency in frontier models (FARRELL 1957). He suggested either the use of a non-parametric piecewise–linear convex isoquant or the use of a parametric function fitted to the data in a way that no point should lie left or below the frontier. He illustrated his ideas by using a simple example involving firms which use two inputs (x_1,x_2) to produce a single output[1].

Figure 1 Technical and Allocative Efficiency



SOURCE: COELLI (1995)

The production function of the fully efficient firm is not known in practice. SS' in Fig. 1 represents the unknown isoquant. If a firm uses quantities of inputs, defined by point P, to produce a unit of output the technical efficiency of that firm is defined to be the ratio of OQ/OP. The point Q is technically efficient because it lies on the efficient isoquant[2]. By considering the input price ratio AA', allocative efficiency may be calculated by OR/OQ. The total economic efficiency is then OR/OP. The perfect market, represented by the unknown isoquant, is the norm against which firm P is compared. A decision maker would be considered efficient only if he had perfect knowledge of the best technology, of the future action and reaction of other people and of future natural events (PASOUR 1981). The

---

[1] Farrell assumed constant return to scale.

imperfect market itself is the foundation for inefficient behaviour of market participants. RÖDERS (1995) discussed the problem that farmers often behave sub-optimally and the one dimensional objective behaviour of profit maximisation does not reflect the true utility function of the farmer and stated that multidimensional objective functions are difficult to model and to estimate. COELLI (1995) discussed alternative dual forms of the production function technology, such as cost and profit function to reflect alternative behavioural objectives (such as cost minimization). Inefficiency implies entrepreneurial error in the sense that the entrepreneur fails to notice profit opportunities. PASOUR argued that performance standards derived on the assumption of profit maximisation should not be used to measure the performance of a farm, because other objective functions may exist. FARRELL (1957) stated that price efficiency[3] is, in any case, a measure with rather limited usefulness, although the measurement of allocative and economic efficiency is not as controversial as the measurement of technical efficiency. The measurement of technical efficiency has proved difficult and complex and the literature provides a range of methods both at the firm level and the industry level (KALIRAJAN 1999).

FARRELL (1957) introduced technical efficiency as a relative notion, relative to best-observed praxis in the group. To get the "relative" technical efficiency of the *i*th firm we have to calculate the actual output divided by the maximum feasible observable output (though more generally with DEA the problem can be treated as input or output orientated). Because the actual output is observable, the maximum output must be estimated. To get the maximal output there are different methods. All these methods try to find a production frontier, which should represent the maximum possible observed output. There are two main approaches to estimate the frontier, the stochastic frontier approach and data envelopment analysis. The former uses statistical methods and the latter mathematical programming. To estimate the frontier and to obtain efficiency scores the following concepts are to be distinguished: The Deterministic[4] parametric frontier approaches, the Deterministic statistic frontier approaches, the Stochastic frontier approaches, the Deterministic non parametric frontier approaches (DEA).

---

[2] Farrell also discussed the extension to the case of more than two inputs and more than two outputs and non constant return to scale.
[3] FARRELL (1957) used the term price efficiency instead of allocative efficiency.
[4] The term deterministic is generally used to describe a group of methods that assume to have a strict one sided error term and the error term represents the inefficiency of the DMU (Decision Making Unit)

4 Data Envelopment Analyses: a real competitor to SFA

The majority of early economists followed a parameteric approach. However, economists at Berkely advanced a programming approach for piecewise linear frontier production functions that went largely unnoticed by the research community (FØRSUND, SARAFOGLOU 2002).

CHARNES et al. (1978) (CCR) showed that the FARRELL unit isoquant model was a special case of the ordinary linear programming problem. At first in the Operational research and management science, but later also within economics, CCR started a new active research field, popularly called DEA (Data Envelopment Analysis). For the applied economists the great advantage, compared with the aforementioned frontier approaches, was the possibility to use multiple outputs. DEA encompasses a variety of alternative related models for evaluating performance of Decision Making Unit (DMU). The basic DEA models are shown in the Table 1.

Table 1 Basic DEA models

Model	Return to scale	Envelope	Efficiency	Type of efficiency	Data	
					Input	Output
CCR [1]-Output	CRS	Piecewise linear	0 - 1	TE	Semi-p	Free
CCR-Input	CRS	Piecewise linear	0 - 1	TE	Semi-p	Free
BCC [2] -Input	VRS	Piecewise linear	0 - 1	PTE,SE	Free	semi-p
BCC-Input	VRS	Piecewise linear	0 - 1	PTE,SE	Semi-p	Free
ADD [3]	CRS, VRS	Piecewise linear	Free	TE, AE	Free	Free
VARMULT	CRS (log linear)	Piecewise log linear	0 - 1	TE	Free	Free
INVARMULT	VRS (log linear)	Piecewise Cobb Douglas	0 - 8	TE	Free	Free

1 CCR means CHARNES, COOPER and RHODES (1978)
2 BCC means BANKER R.D., A. CHARNES, W.W. COOPER (1984)
3 Non oriented model

Variable return to scale (VRS); Constant Return to Scale (CRS)
Technical efficiency (TE)
Allocative efficiency (AE)
Economic efficiency (EE)
Pure Technical Efficiency (PTE)
Scale efficiency (SE)
Variante Multiplicative (VARMULT)
Invariante Multiplicative (INVARMULT)

Source: Cooper et al. (1999)

DEA and SFA use the concept of the frontier production function to define individual firm specific technical efficiency for a group of firms. The technical efficiency is calculated by comparing the firm's efficiency with its potential. An advantage of the DEA approach is that it places no restrictions on the functional form of the frontier and it does not impose any (explicit) distributional assumption on the firm specific efficiency. DEA can accommodate multiple outputs and inputs but is extremely sensitive to variable selection and errors. The SFA has the advantage that it is a statistical approach and formal hypothesis tests can therefore be performed. However, SFA requires a functional form and needs assumptions about the distribution of the 'random' errors and the efficiencies.

The comparison of Efficiency scores becomes problematic as soon as the model, the data or variables are changed. DEA focuses on deriving results for each DMU while within the SFA approach behavioural hypotheses can be tested. COELLI (1998) argues that SFA is likely to be more appropriate than DEA for applications in agriculture, particularly in developing countries where the data are heavily influenced by measurement errors and effects like weather, shocks and diseases.

Several researchers have tried to compare results of applications of different estimation methods having the same set of data. BANKER et al. (1985), SHARAM et al. (1999) and PLESSMANN (2000) compared DEA with other estimation methods, whereby the structure of production was unknown. The common finding is that efficiency measurement depends on the choice of functional form of the SFA. GONG and SICKLES (1992) utilized Monte Carlo techniques to control the underlying technology and compared SFA and DEA. If the functional form is closed to the underlying technology, SFA outperforms DEA. If the functional form is closed to the underlying technology, SFA outperforms DEA and DEA seems to be more appropriate when the knowledge about the underlying technology is weak (KALIRAJAN, SHAND 1999).

While the two approaches yield different results, DEA remains a competitor to stochastic frontier methods. Also, the use of more than one technique enables more robust inferences to be drawn. However, for DEA to be viewed as a true competitor, point estimates of efficiencies are not enough. In order to make inference on the DEA estimates, SIMAR and WILSON (1998) proposed a general methodology for bootstrapping in frontier models.

The bootstrap procedure for nonparametric frontier models was outlined by SIMAR and Wilson (1998), (2000a), (2000b)(SM). Our paper is not indented to explain the bootstrap procedure in detail. However, because it is important for the understanding of the remainder the general idea behind the bootstrap approach and the extension made by SIMAR and WILSON will be illustrated. In the following bootstrapping is discussed in terms of input efficiency. Given a column vector of p inputs (denoted by $x \in \Re_+^p$) and q outputs ($y \in \Re_+^q$), the activity of the production set can be described by means of the

$$\Psi = \{(x, y) \in \Re_+^{p+q}\} \qquad (1)$$

production set (SIMAR and WILSON 2000b). The input requirement set defined $\forall y \in \Psi$ is

$$X(y) = \{x \in \Re_+^p \mid (x, y) \in \Psi\} \qquad (2)$$

The Farrell efficiency boundaries are subsets of $X(y)$ denoted by

$$\partial X(y) = \{x \mid x \in X(y), \theta x \notin X(y) \forall 0 < \theta < 1\} \qquad (3)$$

The efficiency for a given point (x_k, y_k):

$$\theta_k = \min\{\theta \mid \theta x_k \in X(y_k)\} \qquad (4)$$

If $\theta_k = 1$ the unit k is input efficient. An $\theta_k \leq 1$ represents the feasible proportionate reduction of inputs the DMU could realize if y_k were produced efficiently. Hence the efficient level of input of corresponding to the output level is:

$$x^\theta(x_k, y_k) = \theta_k x_k .$$

θ_k is unknown because $\Psi, X(y)$ and $\theta_k x_k$ are unknown.

Let P denote the DGP (data generating process) defined by specific assumption (see SIMAR and WILSON 2000b, page 7-8), from which the random sample $\chi = \{(x_i, y_i) \mid i = 1,\ldots,n\}$ is obtained. By using a nonparametric method M to obtain $\hat{\Psi}, \hat{X}(y), \partial\hat{X}(y)$ it is possible to estimate its efficiency $\hat{\theta}_k = \min\{\theta \mid \theta x_k \in \hat{X}(y_k)\}$. (5)

The bootstrap procedure is based on the simple idea that there is a Data Generating Process which can be determined by Monte Carlo approximation. Therefore, it may be a reasonable estimator of the true unknown DGP generated from the data χ. By using the nonparametric method M

$$\hat{\theta}(x_0, y_0) = \min\left\{\theta > 0 \mid y_0 \leq \sum_{i=1}^{n} \gamma_i y_i, \theta x_0 \geq \sum_{i=1}^{n} \gamma_i, x_i, \sum_{i=1}^{n} \lambda_i = 1, \gamma_i \geq 0, i = 1,\ldots,n\right\}. \qquad (6)$$

We can obtain an estimate $\hat{\theta}(x_0, y_0)$ of the true $\theta(x_0, y_0)$ efficiency for the farm f_0.

Consider now a new data set $\chi^* = \{(x^*{}_i, y^*{}_i) \mid i = 1,...,n\}$ drawn from $\hat{P}$. The convex hull of χ^* gives an estimator $\hat{\Psi}^*$ of $\hat{\Psi}$, where $\hat{\Psi} \subseteq \Psi$.

Corresponding to $\hat{\Psi}$ we can derive

$$\hat{\theta}^*(x_0, y_0) = \min\left\{\theta > 0 \mid y_0 \le \sum_{i=1}^{n} \gamma_i y^*{}_i, \theta x_0 \ge \sum_{i=1}^{n} \gamma_i, x^*{}_i, \sum_{i=1}^{n} \lambda_i = 1, \gamma_i \ge 0, i = 1,...,n\right\} \quad (7)$$

The sampling distribution of $\hat{\theta}^*(x_0, y_0)$ is known since $\hat{P}$ is known. This distribution can easily approximated by Monte Carlo methods. By using $\hat{P}$ to generate B samples $\chi^*{}_b, b = 1,..., B$, yields a set of pseudo estimates $\hat{\theta}_b{}^*$. The bootstrap method introduced by EFRON (1979) is based on the idea that if $\hat{P}$ is a consistent estimator of P, the known bootstrap distribution will mimic the original unknown sampling distribution of the estimators of interest with: $\left(\hat{\theta}^*(x, y) - \hat{\theta}(x, y)\right) \mid \hat{P} \overset{approx}{\sim} \left(\hat{\theta}(x, y) - \theta(x, y)\right) \mid P$. (8)

Unfortunately this "naive" bootstrap would yield inconsistent estimates (SIMAR and WILSON 2000a). Therefore SM introduced a homogeneous smoothed bootstrap procedure. An easily implemented algorithm for consistently generating the bootstrap values $\hat{\theta}_b{}^*$ from a kernel density estimate is given in SIMAR and WILSON 1998 and SIMAR and WILSON 2000b.

The complete smoothed bootstrap algorithm is summarized by the following steps:

(a) First for each $(x_k, y_k) k = 1,..., n$ compute $\hat{\theta}_k$ by the linear program (6)

(b) Using the smooth bootstrap generate a random sample of size n from $\hat{\theta}_i, i = 1,..., n$ providing $\hat{\theta}^*{}_{1b},..., n\hat{\theta}^*_{nb}$.

(c) Then compute the bootstrap estimates $\hat{\theta}^*{}_{k,b}$ for $\hat{\theta}_k$ for $k = 1,..., n$ by solving

$$\hat{\theta}^*(x_0, y_0) = \min\left\{\theta > 0 \mid y_k \le \sum_{i=1}^{n} \gamma_i y_i, \theta x_k \ge \sum_{i=1}^{n} \gamma_i, x^*{}_{k,b}; \sum_{i=1}^{n} \lambda_i = 1, \gamma_i \ge 0, i = 1,...,n\right\} \quad (9)$$

where $x^*_{k,b} = (\hat{\theta} / \theta^*_i) x_i, i = 1,..., n$ and θ^*_i is a factor that is necessary to correct the generated bootstrap sequence. (see SIMAR and WILSON 2000a, page 56) with

$$\theta^* = \overline{\beta}^* + 1/(\sqrt{1 + h^2 / \hat{\sigma}_{\hat{\theta}}{}^2})(\tilde{\theta}^*{}_i - \overline{\beta}^*) \quad (10)$$

with $\overline{\beta}^* = 1/n \sum_{i=1}^{n} \beta^*_i$ (11)

and let $\beta_i^*,...,\beta_n^*$ a simple bootstrap sample from $\hat{\theta}_1,...,\hat{\theta}_n$ obtaining be drawing with replacement from $\hat{\theta}_1,...,\hat{\theta}_n$ and the random generator

$$\tilde{\theta}^*{}_i = \begin{cases} \beta_i^* + h\varepsilon_i^* & if -> \beta_i^* + \varepsilon_i^* \leq 1, \\ 2 - \beta_i^* - h\varepsilon_i^* & otherwise \end{cases} \tag{12}$$

where h is called the bandwidth factor and ε_i^* is a random deviate drawn from the standard normal. SIMAR and WILSON (2000b) discussed in detail how to calculate the bandwidth factor. We used the normal reference rule which assigns

$$\hat{h} = \left(\frac{4}{p+q+4}\right)^{\frac{1}{p+q+4}} n^{\frac{1}{p+q+4}} \quad . \tag{13}$$

The rule is only valid when the data are normally distributed with unit variance and zero covariance. Because that will probably not be the case, several other bandwidth factors were chosen in order to assess the impact of that parameter for applied DEA analysis in our analysis.

(d) Finally repeat step b-c B times to provide for k=1,..n a set of estimates

$$\left\{\hat{\theta}_{b,b}^* b = 1,..., B\right\} \tag{14}$$

The procedure described by SIMAR and WILSON (1998) for constructing confidence intervals introduces additional noise into the procedure (SIMAR and WILSON 2000a). Therefore, they introduced an improved procedure (SIMAR and WILSON 1999) to derive confidence intervals which automatically correct for bias without explicit use of a noisy biased estimator. Using the pseudo estimates $\left\{\hat{\theta}_{b,b}^* b = 1,..., B\right\}$ it is possible to find $\hat{b}_\alpha, \hat{a}_\alpha$.

$$\Pr(-\hat{b}_\alpha \leq \hat{\theta}_{DEA}^*(x_0, y_0) - \hat{\theta}(x_0, y_0) \leq -\hat{a}_\alpha \mid \hat{P}(\chi_n)) = 1 - \alpha \tag{15}$$

Finding $\hat{b}_\alpha, \hat{a}_\alpha$ means sorting the values $\hat{\theta}_{DEA}^*(x_0, y_0) - \hat{\theta}(x_0, y_0)$, b 01,…,B in increasing order and then deleting $\left(\frac{\alpha}{2} * 100\right)$ percent of the rows at either end of the list and set $-\hat{b}_\alpha, -\hat{a}_\alpha$ to the endpoints of the array with $, \hat{a}_\alpha \leq \hat{b}_\alpha$ and the $1-\alpha$ -percent confidence interval is then $\hat{\theta}_{DEA}(x_0, y_0) + \hat{a}_\alpha \leq \theta(x_0, y_0) \leq \hat{\theta}_{DEA}(x_0, y_0) + \hat{b}_\alpha$. This procedure is then repeated n times to obtain n confidence intervals, one for each form.

5 DEA Applications in Agricultural Science

PETRAJA-CHAPARRO et al. (1999) state that: [… inspite of the maturity of the DEA literature, there remain same lacunae] and amongst the most important is the absence of a convincing model selection methodology. Therefore, the purpose of this section is to argue that insufficient attention has been payed to decisions regarding data aggregation, sample size and returns to scale within the applied agricultural literature. Moreover, sensitivity analysis has been the exception rather than a rule. The findings are displayed in Table 2, whereby interest is directed toward the different models, data aggregations and sample sizes.

Table 2 Overview over DEA efficiency measurement applications in agriculture

Author	**Title**	**DEA method**	**Place of investigation**	**Explaining efficiency scores**	**Sensitivity analysis**	**Input/Out in DEA** **Sample size**	**Efficiency**
LLEWELYN, WILLIANS (1996)	Nonparametric analysis of technical, pure technical and scale efficiency for food crop production in East Java	CRS, VRS	Indonesia Java	Socio economic regression	-	6/1 (77)	TE, PTE, SE
TOWNSEND, KIRTEN, VINK (1998)	Farms size productivity and return to scale in agriculture revisited: a case study of wine producers in South Africa	CRS, VRS	South Africa	Regression on productivity and farm size	-	7/1 (400)	TE, SE
JAFORULLAH, WHITEMAN (1999)	Scale efficiency in New Zealand dairy industry: a non-parametric approach	CCR, BCC	New Zealand	-	-	3/7 (264)	TE,PTE, SE
SHARMA, LEUNG, ZALESKI (1999)	Technical, allocative and economic efficiencies in swine production in Hawaii: a comparison of parametric and nonparametric approaches	DEA (VRS and CRS), SFA	Hawaii	Regression	-	4/1 (53)	TE, AE, EE
FRASER,CORDINA (1999)	An Application of DEA to irrigated diary farms	CRS, VRS	North Victoria Australia	-	-	6/1 (50)	TE
FRASER, HONE (1999)	Farm level efficiency and productivity measurement using panel data: wool production in south west Victoria	CRS	South west Victorian wool farms	-	-	4/1 (26)	TE, TFP
ZAIBET,DHARMAPALA (1999)	Efficiency of government-supported horticulture: the case of Oman	SFA, DEA (CCR, BCC)	Oman	Regression	-	3/1 (35)	TE

SHAFIQ, REHMAN(2000)	The extent of resource use inefficiencies in cotton production in Pakistan's Punjab: an Application of DEA	CRS, VRS	Pakistan	Regressi on	-	6/1 (117)	TE,AE
BRÜMMER (2001)	Estimating confidence intervals for technical efficiency: the case of private farms in Slovenia	DEA, SFA	Slovenia	No	BS[5]	4/1 (185)	TE

Table 2 shows that both CRS and VRS model specifications and different input data aggregations were applied. Sample sizes range from 26 up to 4027. Most analysis applied both, CRS and VRS. While this should be viewed positively, the majority of the literature has not discussed, with the exception of scale efficiency, which assumptions were likely to be more appropriate. In applied research, the estimation of confidence intervals for the nonparametric frontier approach has been mainly absent, with the exception of BRÜMMER (2001). Although many different DEA models were developed in the past and used for applied analysis, there seems to be little theoretical or empirical guidance with regard to DEA specification, particularly when using bootstrapping.

6 On the quality of efficiency scores

In view of the observations above, this paper uses Slovenian farm data to investigate how efficiency ranking depends on the model specifications and how confidence intervals can be used to give further insights into the validity of the efficiency scores. Efficiencies and confidence intervals for constant and variable returns to scale, for different bandwidth factors h (see equation 12) and for different input aggregations were calculated for a corrected data set of 69 Slovenian farms.

The sample size satisfies the rule proposed by BANKER et al. (1989). He proposed that the sample size (n) should be greater tan 3(m+s), where m is the number of inputs and s the number of outputs. Furthermore, the sample is representative of the reviewed applications in Table 2, where data collection was at the farm level. In addition SIMAR and WILSON 2000b illustrated the bootstrap method by examining data with 70 observations originally from CHARNES et al. (1981). Therefore, the chosen sample size is representative for this kind of investigation. The model specifications include a calculation for CRS and VRS for a 2input/1output and 4input/1output case. The confidence intervals were estimated by using the

[5] Bootstrap/Confidence interval

homogeneous smoothed bootstrap procedure introduced by SILMAR and WILSON with 2000 bootstraps and three different bandwidth factors (h), whereby h(0.58) was estimated by using the normal reference rule (see equation 13). To compare the different confidence intervals the coefficient of separation was estimated (LATRUFFE et al. 2002). The coefficient of separation is a summary statistic, which is calculated by taking each firm and identifying the farms in the sample that are significantly more efficient (at a given significance level). The statistic tells us, what percentage of the sample are significantly less efficient than a given percentage of the sample, after the sample has been ranked. The coefficient of separation is calculated by the ratio of the area under the curve and the area of full separation (see figure 4) and serves to demonstrate the fact that wider intervals mean higher probability of overlapping intervals. In essence, the smaller the coefficient of separation (at a given level of significance), the less we can differentiate between farm efficiencies.

7 Data

The data used in this study are based on the RIAFE (Research Institute for Agricultural and Food economics) farm cost database in Slovenia. It is not the main purpose of this work to investigate the structure of efficiencies in Slovenia, but to show how reliable the DEA results are by applying different model specifications. After the data set was corrected for outliers the mean normalized procedure was (SARKIS 2002). The inputs for the four input case are (1) purchased seed, home grown seed (implicit price), (2) purchased fertilizer, manure (implicit price), (3) chemicals, other direct costs, wages and (4) services and other cost (all inputs in Slovenian Krowns) Output defined as production of wheat in metric tons. For the two input case the inputs 1/2 and 3/4 were aggregated.

8 Computation

To compute the confidence intervals it is necessary to solve n x b linear programs. The GAMS/DEA tool was added to the GAMS system, which solves linear and mixed integer Data Envelopment Analysis (DEA) programs very efficiently (FERRIS et al 2000). By using the SLICE module in GAMS and CPLEX it was possible to reach a very high performance regarding the calculation time[6]. Several runs were made to test the power of the GAMS/DEA SLICE module and our recommendation are that there is no computational burden up to 2500

[6] Hardware Intel® Pentium® 3 processor 800 Mhz. The Gams Programs for the bootstrapping procedure are provided in the Appendix. This shall document the work in the most precise and reproducible way possible.

DMUs and 8 input/1output. Therefore sensitivity analysis on DEA estimates using bootstrapping may be implemented as a standard routine, at least from the computational point of view.

Table 3 Solution time for (CPLEX) Slice Interface DEA (BBC)

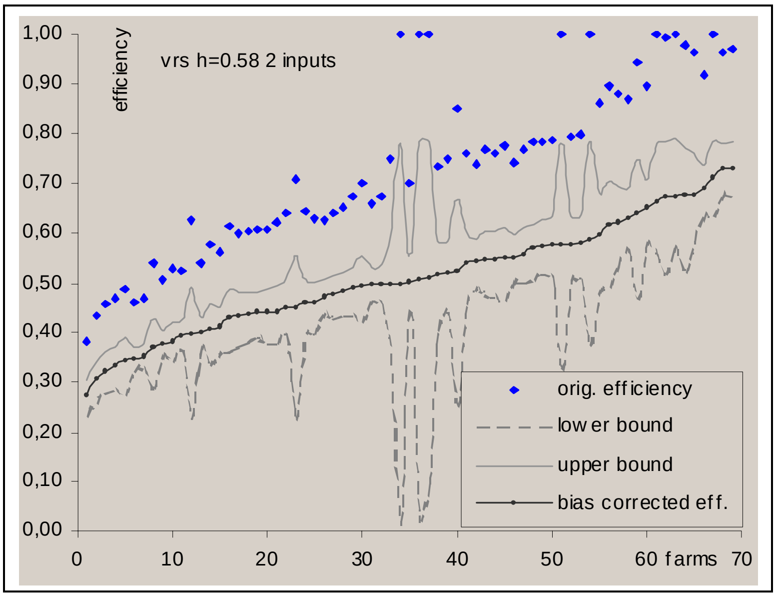
Number of bootstraps	Number of DMUs	Number of outputs	Number of inputs	Solving time
2000	80	4	1	47 min
2000	1000	4	1	7 hours, 24 min

9 Results

To investigate the robustness of the results, the smoothed homogeneous input orientated bootstrap method for each model specifications were applied with the calculated bandwidthfactor[7] of 0.58. If one applies this rule to estimate the bandwidth it is assumed that the data are normally distributed. Nevertheless SIMAR and WILSON found for their sample that the normal reference rule produced a reasonably good estimator for the bandwidth. Additional bandwidths of (0.005/0.05) were used to find the impact of the bandwidths on the confidence intervals for the efficiency scores of the DMUs. The results for the estimated confidence interval for the two input case, CRS/VRS and h(0.58) are shown in Figure 2.

Figure 2: Confidence intervals and point estimates for CRS (a) and VRS (b) 2 inputs h(0.58)

(a)



---

[7] Calculated according to normal reference rule. See equation (13)

(b)

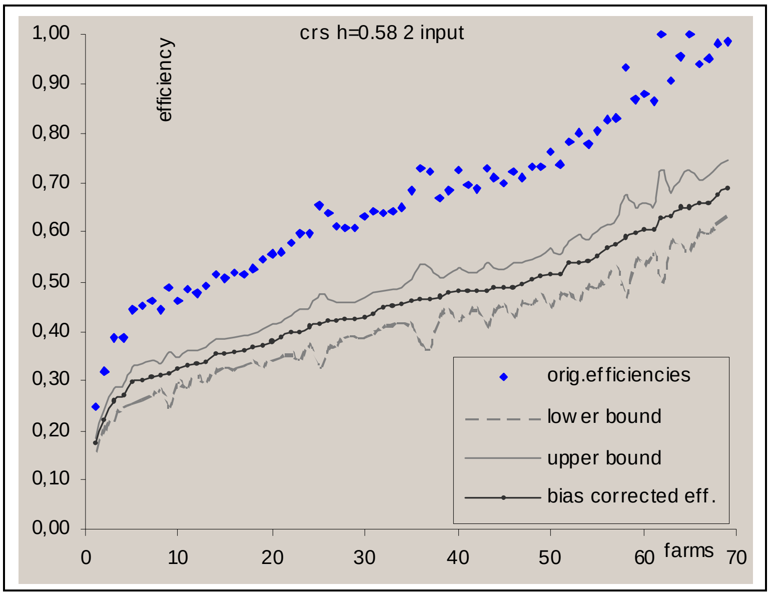
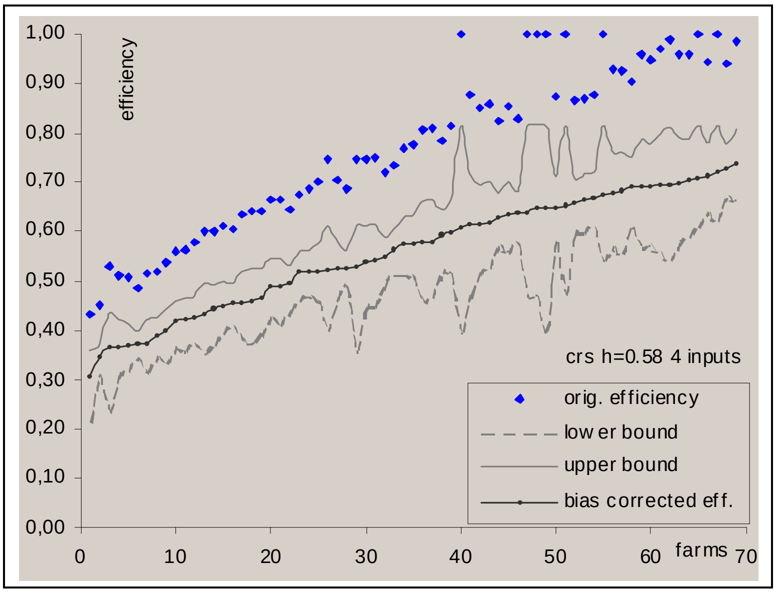


Figure 2 depicts the sample observations ordered by the bias-corrected efficiency score. The 95 percent confidence intervals for each farm are represented by the lower dashed line and the upper line and original efficiencies are indicated by the respective symbols. It is evident that the original efficiencies are not included in the confidence interval. The estimated confidence intervals for the CRS case are narrower than the confidence intervals of the VRS. Figure 2 reveals that the estimated bias is negative and in many cases quite large. Amongst the observations which were originally efficient, the lower bound for the estimated 95 percent confidence intervals range, for the CRS case, from 0.5 to 0.57 and, for the VRS case, from 0.02 to 0.63 for the two input model.

For DMU 57 an original efficiency score of 1.00 was estimated. The bias corrected efficiency was 0.50 and the lower and upper bound of the confidence interval are 0.02 and 0.78. These wide confidence intervals for particular DMUs have also been found by SIMAR and WILSON (2000b). Nevertheless, there are observations where the confidence interval is quite small, in particular for the two input CRS case. The widths of the confidence intervals vary considerably over the sample size especially for the VRS case and for more than two inputs. BRÜMMER (2001) stated that it is easier to identify the observations with low efficiency scores than to identify high performers in his sample. The same observation can be made for the Slovenian farm sample, in particular for the VRS model.

Figure 3 Confidence intervals and point estimates for CRS (a) and VRS (b) 4 inputs h(0.58)

(a)



(b)

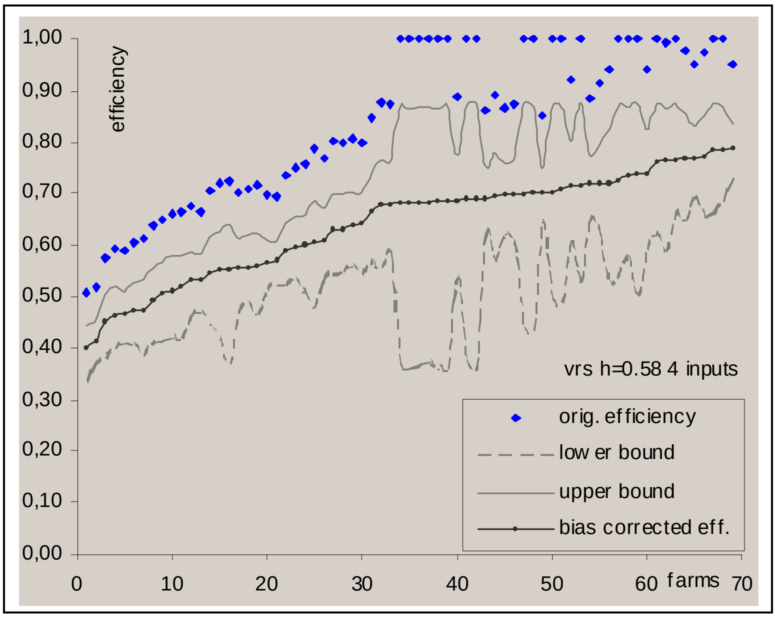


Figure 3 depicts the 4 input case for the estimated bandwidth factor h(0.58). The Weight of the confidence intervals for the VRS as well as for the CRS increases, and hence the coefficients of separation decline.

Table 4 Coefficient of Separation for the different model specifications

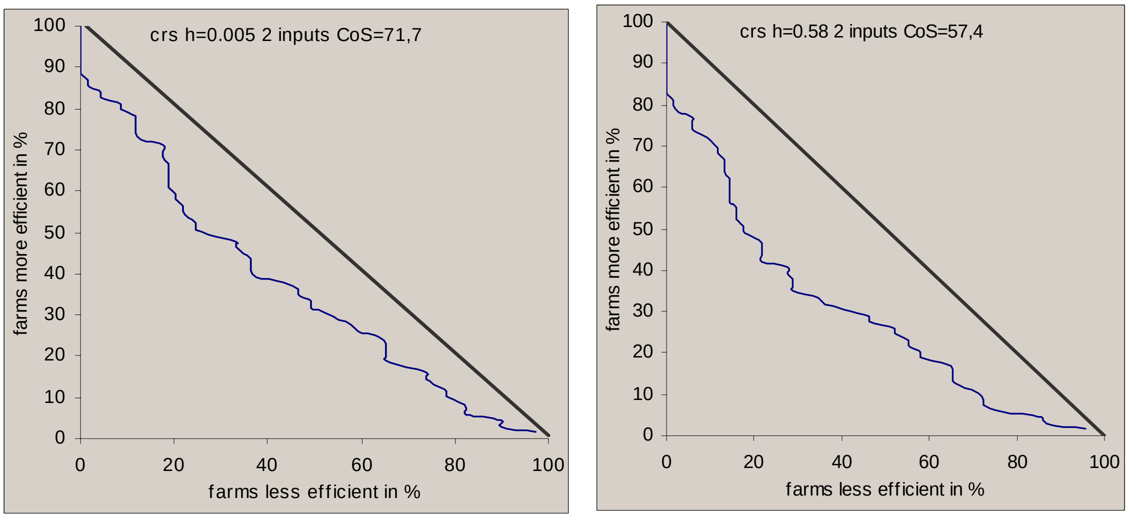
		Coefficient of Separation in %		
Number of input	Return to scale	h=0,005	h=0,05	h=0,58
2	CRS	71.7	70.8	57.4
2	VRS	50.1	48.2	37
4	CRS	52.2	49.4	36.4
4	VRS	29.7	26.9	20.3

The coefficient of separation ranges between 71,7 to 20,3. The highest coefficient of separation is reached in the case of two input, CRS and h(0,005) and the lowest by the four Inputs-VRS h(0,5). For h(0.58) the separation coefficient is still 57 percent. Unfortunately, this is only valid for the model specification case with two inputs and CRS. If the discriminatory power was improved by increasing the number of inputs, the coefficient of separation declined by 20 percent, independent of the choice of return to scale or the bandwidth factor.

The situation deteriorates when applying VRS. The coefficient of separation declines then by 21 percent. The coefficient of separation decreased if a larger bandwidth factor was chosen. The implications of the analysis above are that with moderate sample sizes, expecting to get accurate point estimates using VRS is optimistic. However, this does not undermine the use of DEA in these circumstances. Many studies use these estimates in subsequent analysis, taking DEA scores and regressing them against potential explanatory variables such as education and so on. The implication of the analysis above is that the dependent variable is measure with considerable noise. The results above do suggest that these studies are invalid. However, they do highlight that there are important decisions to be made with regard to using CRS or VRS. The former may be more biased, but if the consequences of using VRS are that the intervals are very wide, then CRS might actually outperform it according to a mean square error criteria. Thus, there is a bias vs efficiency trade off here that is much the same as the tradeoff between using a flexible or parsimonious functional form using a SFA.

On the basis of the results above, we would suggest always doing both CRS and VRS, whereby if the bootstrap standard errors for VRS are too large (CoS < 50 percent), use the CRS for subsequent analysis. Based on a CoS greater than 50 percent DMUs can be grouped and these groups can be used to rank the performance, to set targets or to examine the efficiency of a particular group. Furthermore, one must be aware that higher aggregation of inputs or output means a loss of discriminating power but a gain regarding to the CoS or narrower intervals. The central argument of this paper is therefore: DEA in conjunction with bootstrapping give the applied researcher the possibility to justify the model specification subject to the purpose for which the results are used. Using the Slice Model, there is no overwhelming computational burden of doing this, even in very large sample sizes.

Figure 4: Statistics for the different model specifications to get the Coefficient of Separation (2 inputs)
(a) h(0.005) (b) h(0.58)



10 Summary and Conclusions

Despite the growing literature on the statistical properties of DEA estimators, most agricultural scientists have ignored the sampling noise and often had little theoretical or empirical guidance concerning how to correctly conduct DEA. As different model specifications on Slovenian farm sample demonstrate, ignoring the statistical properties of the estimators or ignoring different assumptions about returns to scale and data aggregation can lead to erroneous conclusions.

For the different model specifications, CVS, VRS 2/4 input and h(0.005;0.05;0.58) it was shown that using VRS instead of CRS decreases the coefficient of separation, and efficient units also have larger confidence intervals than in the CRS case. If the bandwidth factor was increased the slope of the ordered bias-corrected efficiencies decreased. High bandwidth factors increase the bias and the original efficiencies are therefore not included in the confidence intervals. In the Slovenian data set, the coefficient of separation depends on the model specification and ranges from 20% to 71% non-overlapping confidence intervals. Therefore, DEA results must be interpreted cautiously. Here, we suggest always doing both CRS and VRS subject to different input and output aggregations, whereby if the bootstrap standard errors for VRS are too large (e.g. CoS under 50 percent), use the CRS for subsequent analysis and try to increase the aggregation subject to the purpose for which the results are to be used. Apart from the different model specifications, it is important to set up a computational framework which ensured a convenient calculation of confidence intervals for DEA. By using the slice model in GAMS, the statistical properties of the estimator may easily be investigated for any applied study and could shed light on the usefulness of the DEA efficiencies.

The results of applied data envelopment analysis in agriculture can heavily depend on the methodology. Recently developed techniques for interval estimation of technical efficiency can be used to test DEA results. But again, the confidence intervals depend on the model (CCR, BBC) and on the aggregation assumptions. Hence it could be shown, that it is possible to rank 57 percent of the farms significantly under the assumption of CRS and two inputs. Whereas for VRS and four inputs only 20 percent of the DMUs in the sample could be ranked without overlapping intervals although the same data set was used. Therefore, it is possible that using different procedures, the coefficient of separation may heavily depend on the treatment of returns to scale and input aggregation.

Finally, further research might exploit the high performance of the programmed GAMS/SLICE bootstrap procedure. This might extend the work conducted by SIMAR and WILSON by conducting Monte Carlo experiments on more than 2 dimensions of inputs and outputs while also increasing the number of DMUs. Question marks also remain over the relationship between the sample size and the CoS.

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

BANKER, R. D., CONRAD, R. F., STRAUSS, R. P. 1985. A comparison application of DEA and translog methods: An illustrative study of hospital production, Management Science 32, 30-44

BANKER R. D. et al. 1989. An introduction to data envelopment analysis with some of its models and their uses, Research in Government and nonprofit Accounting 5,125-164

BRÜMMER, B. 2001. Estimating confidence intervals for technical efficiency: the case of private farms in Slovenia, European Review of Agricultural Economic 28 (3), 285-306

COELLI, T. J. 1995. Recent Developments in Frontier Modelling and Efficiency Measurement, Australian Journal of Agricultural Economic 39(3), 219-45

COELLI, T. J., PRASADA, R. D. S., BATTESE, G. E. 1998. An Introduction to Efficiency and Productivity Analysis, Kluwer Academic Publishers, Boston/ Dordrecht/ London, 270pp.

CHARNES, A., COOPER, W. W., RHODES, E. 1978. Measuring the Efficiency Of Decision Making Units, European Journal of Operational Research 46, 185-200

COOPER, W., SEIFORD, L., TONE, K. (1999): “Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver”, Kluwer Academic Publisher, Bosten/ Dordrecht/ London

LLEWELYN, R. V., WILLIANS, J. R. 1996. Nonparametric analysis of technical. Pure technical and scale efficiency for food crop production in East Java, Agricultural economics 15, 113-126

FARREL, M. J. 1957. The Measurement of productive efficiency, The journal of the royal statistical society ACXX(3), 253-290

FÄRE, R., GROSSKOPF, S., LOVELL, C. A. K. 1994, Production Frontier, Cambridge University Press, 227pp

FERRIS, M. C., VOELKER, M. M. 2000. Slice Models in General Purpose Modeling Systems, www.wisc.edu/ferris/pub

FØRSUND, F. R., SARAFOGLOU, N. 2002. On The Origins of the Data Envelopment Analysis, Journal of Productivity Analysis 17, 23-40

FRASER, I., HONE, P. 1999. Farm level efficiency and productivity measurement using panel data: wool production in south west Victoria, The Australian Journal of Agricultural and Resource economics 45 (2), 215-32

FRASER, I., CORDINA, D. 1999. An Application of DEA to irrigated diary farms, Agricultural Systems 59, 267-82

GONG, B.-H., SICKLES, R. C., 1992. Finite sample evidence on the performance of stochastic frontiers and data envelopment analysis using panel data, Journal of Econometrics 51 1-2, 259-284
JAFORULLAH, M., WHITEMAN, J. 1999. Scale efficiency in New Zealand dairy industry: a non-parametric approach, The Australian Journal of Agricultural and Resource Economics 43 (3), 523-41

JONDROW, J., LOVELL, C. A. K., MATEROV, I. S., SCHMIDT, P. 1982. On the Estimation of Technical Inefficiency in Stochastic Frontier Production Function, Journal of Econometrics 19, 233-38

KALIRAJAN, K. P., FLINN, J. C. 1982. The Measurement of Farm Specific Technical Efficiency, Pakistan Journal of applied economics 2, 427-42

KALIRAJAN, K. P. and SHAND, R. T. 1999. Frontier Production Functions and Technical Efficiency Measure, Journal of Economic Surveys 13 (2), 148-72

LANSINK, A. O., PIETOLA, K., BÄCKMAN, S. (2002). Efficiency and productivity of conventional and organic farms in Finland 1994-1997, European Review of Agricultural Economics 29(1), 151-65

LATRUFFE, L., BALCOMBE, K., DAVIDOVA, S., ZAWALINSKA, K. 2002. Technical and Scale Efficiency of crop and livestock farms in Poland: Does the choice of specialisation matters? , Imperial College at Wye, mimeo, unpublished

MEEUSEN, W. and VAN DEN BROECK, J. 1977. Efficiency Estimation from Cobb Douglas Production Function with Composed Error, International Economic Review 18, 435-44

PASOUR, E. C. 1981. A Further Note On The Measurement of Efficiency And Economics of farm Size, Journal of Agricultural Economics XXXII(2), 135-46

PEDRAJA-CHAPARRO, F.,SALINAS-JIMENEZ, J., SMITH, P. 1999. On the quality of data envelopment analysis model, Journal of Operational Research Society (1999) 50, 636-644

PLESSMANN, F. 2000, Vergleichende Produktionskostenanalyse des Marktfruchtbaues in Mecklenburg-Vorpommern und Schleswig-Holstein (compared production cost analysis of crop farms in Mecklenburg-Vorpommern and Schleswig-Holstein). AGRIMEDIA Bergen/Dumme, 235pp

RÖDERS, I., 1995 . Determinanten von Erfolgs- und Effizienzunterschieden markfruchtproduzierender Unternehmen. Dissertation, Kiel

SARKIS, J. 2002. Preparing Your Data For DEA, http://www.clarku.edu/jsarkis/deadata.htm last download 10th of August

SHARAM, K. R., LUENG, P., ZALESKI, H. M. 1999. Technical, Allocative and Economic Efficiencies In Swine Production In Hawaii: A Comparison Of Parametric And Nonparametric Approach, Agricultural economics 20, 23-35

SIMAR. L. and WILSON, P. W. 1998. Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models, Management Science 44(1), 49-61

SIMAR, L. and WILSON, P. W. 2000a. Statistical inference in nonparametric frontier models: The state of the art, Journal of Productivity Analysis 13, 49-78

SIMAR, L. and WILSON, P. W. 2000b. A general methodology for bootstrapping in Non-parametric frontier models, Journal of Applied Statistics 27 (6), 779-802

SHAFIQ, M., REHMAN, T. 2000. The Extent Of Resource Use Inefficiencies In Cotton Production In Pakistan's Punjab: An Application Of DEA, Agricultural economics 22, 321-330

TOWNSEND, R. F., KIRTEN, J., VINK, N. 1998. Farms Size Productivity And Return To Scale In Agriculture Revisited: A Case Study Of Wine Producers In South Africa, Agricultural economics 19, 175-180

ZAIBET, L., DHARMAPALA, P. S. 1999, Efficiency Of Government-Supported Horticulture: The Case Of Oman, Agricultural Systems 62, 159-68