Darmstadt Discussion Papers in ECONOMICS

TECHNISCHE UNIVERSITÄT DARMSTADT

A Monte Carlo Study of Old and New Frontier Methods for Efficiency Measurement

Jens J. Krüger

Nr. 200

Arbeitspapiere der
Volkswirtschaftlichen Fachgebiete der TU Darmstadt

ISSN: 1438-2733

A Monte Carlo Study of Old and New Frontier Methods for Efficiency Measurement

February 2011

by

Jens J. Krüger

Darmstadt University of Technology
Department of Law and Economics
Residenzschloss, Marktplatz 15, D-64283 Darmstadt, Germany
tel.: +49 6151 163 693, fax: +49 6151 163 897
e-mail: jjk@vwl.tu-darmstadt.de

Abstract

This study presents the results of an extensive Monte Carlo experiment to compare different methods of efficiency analysis. In addition to traditional parametric-stochastic and nonparametric-deterministic methods recently developed robust nonparametric-stochastic methods are considered. The experimental design comprises a wide variety of situations with different returns-to-scale regimes, substitution elasticities and outlying observations. As the results show, the new robust nonparametric-stochastic methods should not be used without cross-checking by other methods like stochastic frontier analysis or data envelopment analysis. These latter methods appear quite robust in the experiments.

JEL classification: C14, D24, C15

Keywords: Monte Carlo experiment, efficiency measurement, nonparametric stochastic methods

1 Introduction

Until recently estimation methods for frontier production and cost functions widely in use could be divided in two broad groups: parametric-stochastic and nonparametric-deterministic. Both groups have their own advantages and disadvantages. Parametric-stochastic methods such as stochastic frontier analysis (SFA) require the functional specification of the production or cost function as well as the distributions of the stochastic parts but are considered robust against measurement errors. Nonparametric-deterministic methods such as data envelopment analysis (DEA) and free disposal hull (FDH) do not require functional form assumptions or distributional assumptions but are more sensitive to measurement errors.

There is a long history of attempts to combine nonparametric and stochastic methods for efficiency analysis. Grosskopf (1996) gives an early overview of those attempts. Most of these methods remained in an experimental state. Recently two approaches, the so-called order-m and order-α approaches which are explained below, appeared which have the potential of combining the advantages of the parametric-stochastic and the nonparametric-deterministic approaches. These nonparametric-stochastic approaches not require functional form or distributional assumptions *and* are robust with respect to measurement errors.

This paper reports the results of an extensive Monte Carlo study with scenarios containing different production functions, returns-to-scale regimes and the introduction of outliers in addition to the general presence of measurement errors. Considered are a wide range of efficiency measurement methods, including SFA, DEA and FDH as well as the recently developed order-m and order-α approaches which are not considered in such comparisons so far. It is not clear a priori that the conceptual advantages of the nonparametric-stochastic methods also materialize in small-sample situations with different extents of measurement errors and outliers. Including these methods in a unified Monte Carlo experiment is thus important for evaluating their properties in a direct comparison to the more traditional methods.

Indeed, the simulation results reveal that the traditional methods such as SFA and DEA are not as bad as frequently posited on theoretical grounds, at least in the setting chosen in this paper. Moreover, it is shown the robust nonparametric-stochastic order-m and order-α approaches can be much worse than the more traditional approaches when measurement errors are not excessively large and outliers are absent. Thus, these methods should not be employed as a standard procedure without cross-checking by other methods.

The plan of the analysis is as follows. The following section 2 presents the methods of efficiency measurement which are considered in the Monte Carlo study of this paper. Section 3 gives a brief survey of previous Monte Carlo studies and points out their deficiencies. Section 4 describes the design of the Monte Carlo experiment conducted in this paper and section 5 discusses the results of various scenarios. Finally, section 6 draws conclusions for the choice of appropriate methods of efficiency analysis.

2 Methods for Frontier Function Estimation

In this section we give an overview over the methods for frontier function estimation that are subsequently compared in the Monte Carlo experiment. We consider the three classes of parametric-stochastic, nonparametric-deterministic and nonparametric-stochastic methods in turn. We keep this overview brief and confined to the versions used subsequently in the Monte Carlo experiment. The volume edited by Fried, Lovell and Schmidt (2008) provides a comprehensive and up-to-date overview of the methods.

Parametric-Stochastic

As the first class of models we consider the parametric methods of corrected ordinary least squares (COLS) and stochastic frontier analysis (SFA). COLS simply estimates a linear regression through the cloud of input-output points of the data sample and then shifts this regression line upwards until none of the residuals is strictly positive. More specifically, in log-form the production function is

$$\ln y_i = \ln f(\boldsymbol{x}_i, \boldsymbol{\beta}) + u_i \tag{1}$$

for a sample of n ($i \in \{1, ..., n\}$) observations for decision making units (DMUs) which is specified here for the case of two inputs, e.g. capital and labor, $\boldsymbol{x}_i = (K_i, L_i)'$, and a translog functional form

$$\ln f(\boldsymbol{x}_i, \boldsymbol{\beta}) = \beta_0 + \beta_1 \ln K_i + \beta_2 \ln L_i + \beta_3 (\ln K_i)^2 + \beta_4 (\ln L_i)^2 + \beta_5 \ln K_i \ln L_i. \tag{2}$$

It is immediate that the Cobb-Douglas production function is a special case of the translog with the parameter restriction $\beta_3 = \beta_4 = \beta_5 = 0$.

This linear regression in logs can be easily estimated by OLS resulting in the parameter estimate $\hat{\boldsymbol{\beta}}$ and the residuals $\hat{u}_1, ..., \hat{u}_n$. The residuals are then corrected by taking $\hat{u}_i^* = \max_i\{\hat{u}_i\} - \hat{u}_i$ for all $i = 1, ..., n$ making them all non-negative. This is equivalent to correcting the intercept β_0 of the regression such that all input-output combinations are below the estimated function. In the final step the measure of technical efficiency is computed as $TE_i = \exp(-\hat{u}_i^*)$. This efficiency measure is bounded in the range $(0, 1]$ and represents the relative deviation of the actual output of DMU i from the output it should realize when producing technically efficient on the frontier function. A measure of unity indicates technically efficient DMUs. In COLS no measurement error is allowed for. Therefore COLS is a parametric but not really a stochastic method.

SFA departs from this approach in that the residual is now represented by a two-part error term. In logs the production function is here

$$\ln y_i = \ln f(\boldsymbol{x}_i, \boldsymbol{\beta}) + v_i - u_i \tag{3}$$

where $\ln f(\cdot)$ is the translog production function in equation (2) as before. The two-part error term in this stochastic frontier model consists of a normally distributed component, denoted $v_i \sim N(0, \sigma_v^2)$, intended to capture the usual measurement errors, and a half-normally distributed component, denoted $u_i \sim |N(0, \sigma_u^2)|$, which can assume only positive values and reflects systematic downward departures from the frontier function which are associated with inefficiency. Other distributional assumptions could be used but are less common in applications.

Aigner, Lovell and Schmidt (1977) show how to combine these distributional assumptions to form the likelihood function. After maximizing the log of this likelihood function the measure of technical efficiency can be computed with the aid of the Jondrow et al. (1982) formula. This is a formula for the expected value $\mathrm{E}(u_i|\varepsilon_i)$ with $\varepsilon_i = v_i - u_i$ which can be easily applied by plugging in the parameter estimates resulting in $\hat{u}_i$ for $i = 1, ..., n$. The measure of technical efficiency is readily computed as above by $TE_i = \exp(-\hat{u}_i)$ with an analogous interpretation. Again a measure of unity indicates technically efficient DMUs. See Kumbhakar and Lovell (2000) and the chapter of Greene in Fried, Lovell and Schmidt (2008) for more details of the parametric-stochastic approach.

Nonparametric-Deterministic

The nonparametric methods of the second class require no assumption about the form of the production function and also no assumption about the distributions of the error terms. They are based on a set of basic axioms discussed e.g. in Färe and Primont (1995) putting restrictions on a technology set which can be generally stated as

$$T = \{(y_i, \boldsymbol{x}_i) : \boldsymbol{x}_i > \mathbf{0} \,\text{can produce}\, y_i > 0\}\,. \tag{4}$$

This technology set contains all, efficient as well as inefficient, DMUs and summarizes the possibilities of transforming inputs $\boldsymbol{x}_i$ to an output y_i.[1] Based on this technology set an output-oriented distance function can be defined as

$$\theta(y_i, \boldsymbol{x}_i) = \sup\{\theta : (\theta y_i, \boldsymbol{x}_i) \in T\} \tag{5}$$

[1] Actually, in most of the nonparametric efficiency literature the case of multiple outputs where y_i is a vector of outputs is treated from the beginning. We restrict the notation here to the subsequent implementation with a single output in the Monte Carlo experiment.

where θ is a factor that increases the output holding inputs constant. The result is the largest possible increase of the output such that the resulting input-output combination remains feasible, i.e. an element of the technology set T. For efficient DMUs on the frontier of the technology set we have $\theta = 1$ indicating that the output can not be increased further without leaving the technology set. For inefficient DMUs we have $\theta > 1$ indicating the possible increase of the output while staying in the technology set.

Computing these distance functions under the assumptions of a convex technology set and constant returns to scale is the purpose of data envelopment analysis (DEA) under constant returns to scale, denoted DEAc. This is the DEA model originally proposed by Charnes, Cooper and Rhodes (1978). It amounts to solving the following linear programming problem for each observation $i \in \{1, ..., n\}$

$$\max_{\theta,\boldsymbol{\lambda}} \left\{ \theta : \boldsymbol{x}_i \geq \sum_{j=1}^{n} \lambda_j \boldsymbol{x}_j, \theta y_i \leq \sum_{j=1}^{n} \lambda_j y_j, \boldsymbol{\lambda} = (\lambda_1, ..., \lambda_n)' = \mathbf{0} \right\} \tag{6}$$

and denoting the solution value by θ_i which is equal or larger than unity. This solution can be easily computed by the simplex algorithm or other linear program solvers. The measure of technical efficiency is analogous to those defined for the parametric approaches and is here $TE_i = \theta_i^{-1} \in (0, 1]$. As in case of COLS, all deviations from the frontier function are considered as inefficiency. This efficiency measure is therefore deterministic in that it does not allow for measurement errors.

The other nonparametric and deterministic methods considered are DEA under variable returns to scale (denoted DEAv) and free disposal hull (FDH). In the former the restriction $\Sigma_{j=1}^{n} \lambda_j = 1$ is added to the linear program (6) (Banker, Charnes and Cooper 1984). In the latter we abandon the convexity assumption we are computing the FDH solution which retains the previous restriction and adds the integer constraint $\lambda_j \in \{0, 1\}$ to (6) (Deprins, Simar and Tulkens 1984). This makes the program an integer programming problem for which, however, a very fast numerical algorithm has been proposed by Tulkens (1993).

Nonparametric-Stochastic

The third class comprises methods which are both nonparametric and stochastic. In a sense they combine the advantages of the parametric-stochastic approaches like SFA (less sensitivity to measurement error) and the nonparametric-deterministic approaches like DEA or FDH (no requirement of functional and distributional assumptions). Being stochastic makes the approaches robust since the frontier function is not forced to envelop all observations, including maybe outlying ones. The book of Daraio and Simar (2007a) as well as the chapter of Simar and Wilson in Fried, Lovell and Schmidt (2008) give thorough overviews of the nonparametric-stochastic approach. These surveys provide much more detail compared to the following sketch of the nonparametric-stochastic approach.

The order-m and order-α approaches are developed in a series of papers by Cazals, Florens and Simar (2002), Aragon, Daouia and Thomas-Agnan (2005), Daouia and Simar (2007) and Daraio and Simar (2005, 2007b). Both approaches are based on a probabilistic definition of the technology set. For the present case of two inputs in the vector $\boldsymbol{x}_i$ and a single output y_i for a DMU i the probability of being dominated is $H(y_i, \boldsymbol{x}_i) = \Pr(Y \geq y_i, \boldsymbol{X} \leq \boldsymbol{x}_i)$ with upper-case letters denoting random variables, i.e. the probability of producing more output with less input than DMU i. Using this concept an output-oriented distance function can be defined as

$$\theta_1(y_i, \boldsymbol{x}_i) = \sup \{\theta : H(\theta y_i, \boldsymbol{x}_i) > 0\}. \tag{7}$$

This distance function gives the largest possible increase of the output such that the resulting input-output combination has a probability marginally larger than zero. For the actual computation of the order-m efficiency measure the probability is decomposed as

$$H(y_i, \boldsymbol{x}_i) = \Pr(Y \geq y_i | \boldsymbol{X} \leq \boldsymbol{x}_i) \cdot \Pr(\boldsymbol{X} \leq \boldsymbol{x}_i) = S(y_i | \boldsymbol{x}_i) \cdot F(\boldsymbol{x}_i), \tag{8}$$

where $S(\cdot)$ is a conditional survivor function and $F(\cdot)$ is a distribution function. Assuming $F(\boldsymbol{x}_i) > 0$ for the relevant range of inputs, the statement $H(\theta y_i, \boldsymbol{x}_i) > 0$ in the definition of the distance function is equivalent to the statement $S(\theta y_i|\boldsymbol{x}_i) > 0$. The algorithm to compute the order-m efficiency measure relies on drawing at random a prespecified number $m < n$ of DMUs with no more input than $\boldsymbol{x}_i$ from the sample. These DMUs form a *partial* frontier function against which the efficiency of DMU i under consideration is evaluated by DEA or FDH. This procedure is repeated B times resulting in efficiency measures $\hat{\theta}^1_{mi}, ..., \hat{\theta}^B_{mi}$ from which the final order-m efficiency measure is computed as the simple mean $\hat{\theta}_{mi} = B^{-1} \cdot \Sigma_{b=1}^{B} \hat{\theta}^b_{mi}$. Typical default values for m and B used in practice are $m = 25$ and $B = 200$ (Daraio and Simar 2007a), but results are generally quite robust to variations of m and B. To reduce the computational burden we use $m = 25$ and $B = 100$ for the Monte Carlo experiments in this paper where the efficiency measure recorded is $TE_i = \hat{\theta}^{-1}_{mi}$.

By that procedure the efficiency of each DMU is repeatedly evaluated against a partial frontier spanned by just $m < n$ of the sample DMUs. This prevents the entire sample being enveloped by the frontier function and thus giving the procedure its robustness, while preserving the nonparametric nature of the efficiency measurement. As Cazals, Florens and Simar (2002) have shown, the order-m efficiency measure is a consistent estimator and converges at the usual parametric rate of $n^{1/2}$ irrespective of the number of inputs and outputs. This is rather exceptional for a nonparametric estimator which is usually subject to the so-called "curse of dimensionality" meaning that the rate of convergence declines with the dimension of the problem (here the number of inputs and outputs).

The related order-α approach is based on the definition of the production frontier as a quantile instead of the repeated comparison to a partial frontier function. The output-oriented efficiency measure of order $\alpha \in (0, 1]$ is defined as

$$\theta_\alpha(y_i, \boldsymbol{x}_i) = \sup \{\theta : H(\theta y_i, \boldsymbol{x}_i) > 1 - \alpha\}, \tag{9}$$

where α controls the probability of DMU i after having reduced inefficiency in output direction to be dominated by other DMUs which are using not more of the inputs. The choice of α is usually within the interval $[0.90, 0.99]$. In the case $\alpha = 1$ we are back at the definition of the frontier function as enveloping all observations. A choice of $\alpha = 0.95$ compares DMU i with the five percent of DMUs which are producing more output with no more input. This means falsely classifying a DMU as efficient in five percent of the cases and is therefore analogous to committing a type-I error in statistical hypothesis testing. This value of α is chosen in the Monte Carlo experiments below.

Daouia and Simar (2007) derive the statistical properties of the order-α efficiency measure for the general case of multiple inputs and multiple outputs. They show that the order-α efficiency measure is strongly consistent and converges at the usual parametric rate $n^{1/2}$ to a normally distributed random variable. Moreover, they show that order-α efficiency measures have a bounded influence function, whereas order-m efficiency measures have an unbounded influence function and therefore are less robust.[2] To compute order-α efficiency measures they devise an exact algorithm which is implemented in the software used for this Monte Carlo study. This result is transformed into a measure of technical efficiency $TE_i = \hat{\theta}^{-1}_{\alpha i}$.

All computations are performed with R using the package `frontier` for the parametric (which is actually based on the Fortran code of the program Frontier 4.1 by Coelli (1996)) and the package `FEAR` for the nonparametric methods. The functions of the latter package are described in Wilson (2008).

3 Monte Carlo Literature

The first study with a Monte Carlo investigation of efficiency measurement methods was conducted by Banker et al. (1987). This study initiated a continuous stream of further studies which are reviewed in this section. We provide a brief survey focusing on a selection of the more important studies in this area which are discussed in chronological order.

Banker, Gadh and Gorr (1993) undertake a comparison of SFA[3] and DEA in a setting with one output and two inputs which are combined with a piece-wise Cobb-Douglas production function. As performance

[2] The influence function of an estimator, introduced by Hampel (1974), shows the effect of a small (tending to zero) fraction of outlying observations on the estimator. If the influence function is unbounded a small fraction of outlying observations is sufficient to cause divergence of the estimator.

[3] Banker, Gadh and Gorr (1993) call this corrected ordinary least squares but actually mean the stochastic frontier function estimator of Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977).

measures they use the mean absolute deviations and the Wilcoxon signed rank test of the true and estimated efficiencies. The SFA is based on half-normal and exponential inefficiency distributions. The results show that SFA is better for sample sizes of 50, 100 and 200 and large measurement errors, whereas DEA performs better for the small sample size of 25 and small measurement errors. Both methods perform bad for large measurement errors. Very important is the finding that SFA fails to decompose inefficiency and measurement error (see also Ruggiero (1999) for more on that issue). The number of replications is 5 and is thus extremely low for a reliable Monte Carlo study.

Banker, Chang and Cooper (1996) put special emphasis on the issue of returns to scale in a Monte Carlo setting very similar to that of Banker, Gadh and Gorr. Using mean absolute deviations as error measure they conclude that DEA under variable returns to scale performs best, followed by DEA under constant returns to scale and corrected ordinary least squares (COLS) even in the presence of collinearity and misspecification in the form of omitted and irrelevant variables. A weakness of this study is the again very low number of 25 Monte-Carlo replications.

Bojanic, Caudill and Ford (1998) investigate the small-sample properties of parametric and nonparametric methods for frontier function estimation in the presence of heteroskedastic measurement errors. The main finding is that heteroskedasticity generally leads to biased, i.e. overstated, measurement of inefficiency.[4] Comparing across methods, the parametric methods appear to be superior to DEA. This statement is, however, only a relative one since all estimators do not perform very well. Further caveats are that the simulation setting favors the parametric methods against DEA and the here also not overly large number of 100 replications.

Ruggiero (1999) devotes special attention to the inability of SFA to separate measurement error and inefficiency. In a Cobb-Douglas framework with two inputs and one output, different inefficiency distributions and different variances of the measurement error he finds that SFA does not outperform COLS in most cases. Only in the most favorable situation and large sample sizes SFA performs clearly better. The findings are based on a small number of 100 replications and do not consider nonparametric methods.

Resti (2000) employs a cost function setting based on a piece-wise multi-product Cobb-Douglas technology with three outputs and two inputs for a comparison of SFA, DEA and three stochastic variants of DEA. Main finding is that the performance of the stochastic techniques relative to classical DEA variants under constant and variable returns to scale depends on the parameter used to control the compromise between error and inefficiency minimization. He points to a reasonable agreement of the classical techniques and concludes that "stochastic techniques usually do not outperform the classic ones" (Resti 2000, p. 570).

Ruggiero (2007) uses the same setting as in his previous study to show that while SFA is not able do separate inefficiency and measurement error for cross-sectional data, the additional information contained in panel data allows to do this task more effectively. Comparing SFA and DEA he concludes that the "... results suggest that the stochastic frontier model holds no real advantage over DEA. In particular, the purported advantage of the stochastic frontier, i.e. the ability to allow measurement error, can be overcome by averaging the data to smooth production" (Ruggiero 2007, p. 266).

Van Biesebroek (2007) compares a number of different methods for productivity measurement in a dynamically optimizing framework by means of a Monte Carlo analysis. Especially for comparing DEA and SFA he obtains some interesting general results. As expected, he finds that DEA tends to outperform SFA in small samples and in situations with small measurement errors (conversely, SFA is superior in larger samples and for larger errors) and that specification error is more detrimental to SFA. He finds that the loss of precision in the presence of large measurement errors is more a problem of DEA than of SFA, but also observes the frequent failure of SFA to separate inefficiency from noise, especially when measurement error is large. He points to the possible problem of multicollinearity when using the translog function[5] and to the problem of the very low number of replications in many other Monte Carlo comparisons of productivity and efficiency measurement methods.

Summarizing it can be said that the ranking of parametric-stochastic and nonparametric-deterministic approaches depends on the specific setting chosen and that there is no clear superiority of one of the approaches. Furthermore, it is striking that the number of Monte Carlo replications is rather low in these

[4] See Caudill and Ford (1993) for further results on the biases caused by heteroskedasticity.

[5] Gong and Sickles (1992) focus on functional form choice for econometric frontier estimation. They find in an experiment with simulated panel data that constant elasticity of substitution (CES) and translog functional forms outperform other functional forms and DEA in most cases. A caveat here is that only quite large sample sizes in the range of 500 to 2500 observations are considered.

studies. The present study remedies for this efficiency by basing all results on $B = 1000$ replications of the Monte Carlo experiments. This makes the results of this study much more reliable than those reviewed in this section. The next section explains the general design of the Monte Carlo experiment and the different scenarios which are explored in this experiment.

4 Monte Carlo Design

The design of the Monte Carlo experiment covers a baseline scenario in which the data are generated by a Cobb-Douglas production function under constant returns to scale, where both inputs are treated symmetrically and no outliers are induced. Subsequently the results from the baseline scenario are compared with various extensions which treat the inputs asymmetrically, consider decreasing and increasing returns to scale, use a more general constant elasticity of substitution (CES) production function and induce outliers with different frequencies. The parameter choice is oriented at the values used by Banker, Gadh and Gorr (1993) and Ruggiero (1999).

Each scenario is simulated for different sample sizes $n \in \{50, 100, 200\}$, extents of inefficiency defined by the standard deviation of inefficiency $\sigma_u \in \{0.2, 0.3\}$ and extents of measurement error defined by the standard deviations of measurement error $\sigma_v \in \{0.05, 0.1, 0.15\}$ and proceeds according to the following steps:

1. Fix the sample size n, the standard deviation of the inefficiency σ_u and the standard deviation of the measurement errors σ_v.
2. Draw n values of K_i and L_i from a continuous uniform distribution $U(1, 20)$ and keep them fixed. As well draw n values of u_i from a half-normal distribution $|N(0, \sigma_u^2)|$ and transform them to efficiency levels by $A_i = \exp(-u_i)$ which are also kept fixed.
3. Generate the output levels by the production function $y_i = A_i \cdot F(K_i, L_i)$ and keep them also fixed.
4. For each of the $B = 1000$ replications draw n values of v_i from a normal distribution $N(0, \sigma_v^2)$ and use them as measurement errors disturbing the output $\widetilde{y}_i^{(b)} = y_i \cdot \exp(v_i)$, where $b = 1, ..., B$.[6]
5. Apply the efficiency measurement methods to the each of the B samples $(\widetilde{y}_i^{(b)}, K_i, L_i)$, $i = 1, ..., n$ and record the computed efficiency estimate $\hat{A}_i^{(b)}$.
6. Compute the mean absolute error and the Spearman rank correlation coefficient for each replication $b = 1, ..., B$ by $MAE^{(b)} = \frac{1}{n}\Sigma_{i=1}^{n}|\hat{A}_i^{(b)} - A_i|$ and $R_{SP}^{(b)} = 1 - 6 \cdot \Sigma_{i=1}^{n} d_i^2/n(n^2-1)$, respectively, with d_i as the rank difference of A_i and $\hat{A}_i^{(b)}$.
7. Finally, average over $b = 1, ..., B$ to reach the MAE and R_{SP} values reported in the tables below.[7]

In the baseline scenario a Cobb-Douglas production function $F(K_i, L_i) = K_i^{\alpha} L_i^{1-\alpha}$ with $\alpha = 0.5$ is used. Asymmetry is introduced by using the baseline scenario with $\alpha = 0.8$ as the single change. In the scenarios with variable returns to scale the production function is modified to $F(K_i, L_i) = (K_i^{\alpha} L_i^{1-\alpha})^{\rho}$ with $\rho = 0.8$ in the case of decreasing returns to scale (DRS) and $\rho = 1.2$ in the case of increasing returns to scale (IRS). The two scenarios with the CES production function use $F(K_i, L_i) = (K_i^{\alpha} + L_i^{\alpha})^{\rho/\alpha}$ where ρ is fixed to unity (meaning constant returns to scale) and the parameter α now governs the elasticity of substitution which is equal to $1/(1-\alpha)$. We consider two choices of α which lead to a relatively small elasticity of substitution of 2 ($\alpha = 0.5$) and a relatively large elasticity of substitution of 5 ($\alpha = 0.8$). Finally, the two outlier scenarios select a fraction η of observations for which the output levels in step 3 are each multiplied by a uniformly distributed random number from the interval $[1, 1.5]$. Two fractions of outliers are considered corresponding to a relatively small extent of outliers for $\eta = 0.05$ and a relatively larger extent of outliers for $\eta = 0.1$. All simulations are based on the same set of random draws for each scenario and each combination of parameters generated by fixing the seed of the random number generator.

[6] To avoid problems with wrong (positive) skewness for the SFA estimates we check the residuals of an OLS estimation of the production function. Only if these are negatively skewed this draw is used to generate a replication of the efficiency estimates with all methods. Otherwise a new sample of measurement errors is drawn.

[7] The usage of rank correlation coefficients is common in the literature and can be justified by results of Ondrich and Ruggiero (2001) who show that neither stochastic nor deterministic frontier methods are able to estimate absolute measures of efficiency but only relative ones.

5 Monte Carlo Results

In this section we present the summary tables for the different scenarios and discuss the main insights which can be gained from them. We consider the mean absolute error MAE (meaning the mean absolute deviation of the true efficiency level from the efficiency level measured by the respective method) and the Spearman rank correlation coefficient R_{SP} of the true and the measured efficiency levels. A lower MAE indicates that the estimate is closer to the true efficiency levels on average, whereas a higher R_{SP} shows that true and estimated efficiency levels are more strongly associated.

— insert table 1 about here —

Table 1 contains the results of the baseline scenario. Starting with the comparison of the parametric-stochastic and the nonparametric-deterministic methods we observe that SFA mostly has a smaller MAE (with some minor exceptions at the smallest sample size) and a similar R_{SP} compared to DEAc. Comparing DEAc and DEAv we see that DEAc has larger MAE values than DEAv in many cases, but also lower MAE values in other cases. DEAc shows a consistently larger R_{SP}. DEAv itself has much smaller MAE values compared to COLS, but also smaller R_{SP}. MAE of COLS is smaller than FDH only at small error variances, but R_{SP} is uniformly much larger in the case of COLS. FDH itself uniformly dominates order-m in terms of considerably smaller MAE values, but R_{SP} also tends to be smaller for FDH. Order-m has MAE values that are always considerably smaller compared to order-α for $n = 100$ and $n = 200$. Order-α, however, is slightly better in the cases with $n = 50$. In this comparison R_{SP} is uniformly larger for order-m.[8]

In addition, among the parametric methods the MAE is uniformly smaller for SFA compared to COLS, whereas R_{SP} is about equal. Comparing the nonparametric methods with variable returns to scale, DEAv and FDH, we see that MAE is smaller in the case of the former with exceptions at large error variances. R_{SP} is also uniformly substantially larger in the case of DEAv. Thus, for this smooth Cobb-Douglas production function DEAv relying on a convex technology set provides a better approximation than the non-convex FDH.

In summary, this pattern of results leads to a reasonably clear ranking of the efficiency measurement methods

$$\text{SFA} \succ \text{DEAc} \succcurlyeq \text{DEAv} \succ \text{COLS} \succ \text{FDH} \succ \text{order-}m \succ \text{order-}\alpha$$

with "$\succ$" indicating strictly better and "$\succcurlyeq$" indicating weakly better. It should be emphasized at this place that the baseline scenario of the Monte Carlo experiment is very advantageous for the parametric econometric approaches so that the good performance of SFA is not overly surprising. A bit more surprising is that COLS is on the one hand dominated by the two DEA variants and on the other hand is better than FDH except at large error variances. Also rather unexpected is that DEAv is frequently better than DEAc in terms of MAE although not in terms of R_{SP}. The performance of the order-m and order-α approaches shows that in the absence of outliers these robust nonparametric methods are not a very fortunate choice. In terms of MAE at the higher measurement error variances order-m and order-α are occasionally (at the largest error variance) better than COLS and DEA but they are never better than SFA or FDH. Compared to FDH, order-m and order-α tend to be better in terms of R_{SP} but they are not better than any of the other methods according to this criterion.[9] We will see in the following how this ranking changes when we depart from the baseline scenario in different directions.

— insert table 2 about here —

[8]Results for the mean error ME (not reported here) show that the order-m and order-α approaches have the largest negative biases (meaning the largest overestimation of A). Negative biases are also observed for FDH whereas we have positive biases throughout for COLS. The smallest biases are observed for SFA and the DEA variants. Interestingly, bias is reduced with larger sample size only for SFA and FDH.

[9]In the papers of Aragon, Daouia and Thomas-Agnan (2005), Daouia and Simar (2005, 2007) and Daraio and Simar (2005) in which the order-m and order-α approaches are developed, numerical illustrations are used to demonstrate the advantages of these methods. These numerical illustrations consist of a single draw of a data set and the introduction of a small number (2 to 5) of deliberately positioned outliers and are therefore not comparable to the Monte Carlo analyses we provide in this paper.

Table 2 reports the results of one of these deviations, the asymmetry scenario. We find no, not even numerical, changes for COLS and SFA. This is natural since the parameter change to induce the asymmetry can be perfectly accommodated by the translog production function. The MAE values of DEAc and DEAv increase only by small amounts relative to baseline. For FDH, order-m and order-α the MAE values are decreasing, sometimes substantially. R_{SP} changes in a reverse pattern. It is barely changing for DEAc and DEAv and it is increasing for FDH, order-m and order-α.

— insert tables 3 and 4 about here —

Tables 3 and 4 show the results for the DRS and IRS scenarios, respectively. Again and for the same reason as in the asymmetry scenario the results for COLS and SFA are identical to the baseline scenario. For the nonparametric methods interesting differences of the DRS and IRS scenarios relative to baseline CRS show up. We have strong increases of MAE for DEAc which are more pronounced for DRS rather than for IRS. These occur together with marked reductions of R_{SP}, which are again less pronounced for IRS. These changes are more clearly recognizable for small sample sizes and small standard deviations. In the case of DEAv MAE and R_{SP} hardly change under DRS, whereas MAE increases and R_{SP} decreases slightly under IRS (although much less than in the case of DEAc). The FDH results are analogous to DEAv with a slightly larger reduction of MAE under DRS and a smaller increase of MAE under IRS. Here, R_{SP} increases more than for DEAv compared to baseline under DRS, but also decreases more under IRS in general. Furthermore, for order-m and order-α the MAE is reduced under DRS. The extent of the reduction is larger than for DEAv and FDH. Under IRS MAE increases in a similar order of magnitude as in the cases of DEAv and FDH. For the order-m and order-α methods R_{SP} increases under DRS while it decreases under IRS in an order of magnitude comparable to FDH.

— insert tables 5 and 6 about here —

Tables 5 and 6 contain the results for the CES scenario with the small and large elasticity of substitution, respectively. Compared to the baseline scenario, the results for the parametric methods, COLS and SFA, are here not affected by very much. For the nonparametric methods DEAc experiences increasing MAE which is larger in the case of the larger elasticity of substitution. DEAv shows a similar pattern, but less pronounced. In the case of FDH nearly no changes of MAE can be observed in case of the smaller elasticity of substitution and minor reductions of MAE in case of the larger elasticity of substitution. We also recognize consistent MAE reductions for order-m and order-α. These reductions are small in magnitude and occur similarly for order-m and order-α. They are slightly larger for the larger elasticity of substitution. In general, we observe here only minor changes of R_{SP} which are largest in the case of FDH.

— insert tables 7 and 8 about here —

Tables 7 and 8 finally report the results for the outlier scenarios with small and large extents of outliers, respectively. Compared to baseline, the introduction of outliers raises MAE substantially for COLS and much less in the case of SFA. R_{SP} decreases in the majority of occasions, in about equal magnitude for COLS and SFA. Consistent with these observations a larger extent of outliers affects COLS much more than SFA. We also observe substantial increases of MAE for DEAc and DEAv with DEAc being more affected than DEAv, simultaneously R_{SP} decreases. FDH shows hardly changed MAE values with a slight tendency to decline whereas R_{SP} decreases somewhat more. Overall FDH appears surprisingly robust compared to the parametric methods and the DEA variants. The robust nonparametric and stochastic order-m and order-α methods are quite stable across all scenarios. MAE values even tend so be slightly smaller in the presence of outliers compared to the baseline scenario. Across methods, order-m and order-α are only on rare occasions better than DEAc (clustered around the cases $\sigma_u = 0.2$, $\sigma_v = 0.15$). Both order-m and order-α are much more frequently worse than DEAv and always worse than FDH and SFA. At least, they are much better than COLS. In most of the cases order-m is better than order-α in terms of MAE. R_{SP} is rather stable with only slight increases/decreases relative to the baseline scenario. The increases of R_{SP} tend to be larger for order-α especially for the smaller extent of outliers. These latter findings are somewhat puzzling since the comparison is relative to the baseline scenario without the additional outliers.

In general, for the nonparametric approaches it can be stated that a greater extent of outliers leads to increases of MAE for the DEA variants, but is associated with fewer changes and even decreases of MAE

in the cases of FDH, order-m and order-α. Moreover, little changes of R_{SP} are observed for order-m and order-α. In this respect they are robust. We recognize that the parametric methods COLS and SFA are not much affected by changes of the Monte Carlo setting except for the introduction of outliers. Even the case of the CES production function which is not a special case of the translog function (as the Cobb-Douglas function is) leads to only minor changes of MAE and R_{SP} compared to the baseline scenario. This means that the translog function has good approximating power also for the CES function.

6 Conclusion

Which general guidelines for the application of efficiency analysis methods can be derived from the results discussed above? Since we are faced here with results from a Monte Carlo experiment it is difficult to derive general guidelines. All Monte Carlo experiments are faced with the problem of "specificity" (Hendry 1984, p. 941) meaning that those experiments can cover only a small part of the space of possible parameters and functional forms. Thus we have generated only selective results pertaining to a limited set of points of the parameter space.

Keeping this caveat in mind we can summarize a few general patterns. SFA shows a good performance also in the CES scenarios and even in the presence of outliers. Although we have to be cautious with this statement since the experimental design is quite favorable for SFA, we can say that the criticism raised against this method (i.e. that it is not able to separate inefficiency and measurement error) can not be confirmed by the present study. The performance of DEA and FDH is also quite remarkable even for larger measurement error variances and in the presence of outliers. Some changes can be observed when increasing or decreasing returns to scale are induced. The results are, however, not much affected by the CES scenarios even in the case of the larger elasticity of substitution.

Overall, we have the ranking of SFA dominating DEA, which itself is dominating FDH and order-m and order-α are dominated in general. Nevertheless, order-m becomes better at larger measurement error variances and for larger extents of outliers. When outliers are induced, order-m and order-α are rather stable and sometimes even slightly better compared to their performance in the baseline scenario. Moreover, order-m performs better than the theoretically more robust order-α approach. This conclusion is, however, somewhat tentative. Maybe differently designed Monte Carlo experiments with less well behaved environments change this conclusion.

Compared to that the performance of the robust nonparametric-stochastic methods, the order-m and order-α approaches, appears not to be very advantageous in well behaved settings. At the bottom line it can be said that it is not appropriate to use order-m and order-α routinely in situations which are characterized by a smooth well behaved production environment with no outliers or limited extents of outliers. Even with the induction of additional outliers it depends much on the values of the other parameters whether these methods become better than more traditional methods of efficiency analysis. Thus, it is always wise to cross-check the results with other methods to asses their validity.

References

[1] Aigner, D.J., Lovell, C.A.K., Schmidt, P. (1977), Formulation and Estimation of Stochastic Frontier Production Function Models, Journal of Econometrics, vol. 6, pp. 21-37.

[2] Aragon, Y., Daouia, A., Thomas-Agnan, C. (2005), Nonparametric Frontier Estimation: A Conditional Quantile-Based Approach, Econometric Theory, vol. 21, pp. 358-389.

[3] Banker, R.D., Chang, H., Cooper, W.W. (1996), Simulation Studies of Efficiency, Returns to Scale and Misspecification with Nonlinear Functions in DEA, Annals of Operations Research, vol. 66, pp. 233-253.

[4] Banker, R.D., Charnes, A., Cooper, W.W. (1984), Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis, Management Science, vol. 30, pp. 1078-1092.

[5] Banker, R.D., Charnes, A., Cooper, W.W., Maindiratta, A. (1987), A Comparison of DEA and Translog Estimates of Production Functions Using Simulated Data from a Known Technology, in: A. Dogramaci, R. Färe (eds.), Applications of Modern Production Theory: Efficiency and Productivity, Boston (Mass.): Kluwer, pp. 33-58.

[6] Banker, R.D., Gadh, V.M., Gorr, W.L. (1993), A Monte Carlo Comparison of Two Production Frontier Estimation Methods: Corrected Ordinary Least Squares and Data Envelopment Analysis, European Journal of Operational Research, vol. 67, pp. 332-343.

[7] Bojanic, A.N., Caudill, S.B., Ford, J.M. (1998), Small-Sample Properties of ML, COLS, and DEA Estimators of Frontier Models in the Presence of Heteroskedasticity, European Journal of Operations Research, vol. 108, pp. 140-148.

[8] Caudill, S.B., Ford, J.M. (1993), Biases in Frontier Estimation due to Heteroscedasticity, Economics Letters, vol. 41, pp. 17-20.

[9] Cazals, C., Florens, J. P., Simar, L. (2002), Nonparametric Frontier Estimation: A Robust Approach, Journal of Econometrics, vol. 106, pp. 1-25.

[10] Charnes, A., Cooper, W.W., Rhodes, E. (1978), Measuring the Efficiency of Decision Making Units, European Journal of Operational Research, vol. 2, pp. 429-444.

[11] Coelli, T. (1996), A Guide to FRONTIER Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation, CEPA Working Paper 96/08, University of New England.

[12] Daouia, A., Simar, L. (2005), Robust Nonparametric Estimators of Monotone Boundaries, Journal of Multivariate Analysis, vol. 96, pp. 311-331.

[13] Daouia, A., Simar, L. (2007), Nonparametric Efficiency Analysis: A Multivariate Conditional Quantile Approach, Journal of Econometrics, vol. 140, pp. 375-400.

[14] Daraio, C., Simar, L. (2005), Introducing Environmental Variables in Nonparametric Frontier Models: A Probabilistic Approach, Journal of Productivity Analysis, vol. 24, pp. 93-121.

[15] Daraio, C., Simar, L. (2007a), Advanced Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications, Studies in Productivity and Efficiency, New York: Springer.

[16] Daraio, C., Simar, L. (2007b), Conditional Nonparametric Frontier Models for Convex and Nonconvex Technologies: A Unifying Approach, Journal of Productivity Analysis, vol. 28, pp. 13-32.

[17] Deprins, D., Simar, L., Tulkens, H. (1984), Measuring Labor Efficiency in Post Offices, in: M. Marchand, P. Pesiteau, H. Tulkens (eds.), The Performance of Public Enterprises: Concepts and Measurement, Amsterdam: North-Holland, 243-267.

[18] Färe, R., Primont, D. (1995), Multi-Output Production and Duality: Theory and Applications, Boston: Kluwer.

[19] Fried, H.O., Lovell, C.A.K., Schmidt, S.S. (2008), The Measurement of Productive Efficiency and Productivity Growth, Oxford: Oxford University Press.

[20] 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, vol. 51, pp. 259-284.

[21] Grosskopf, S. (1996), Statistical Inference and Nonparametric Efficiency: A Selective Survey, Journal of Productivity Analysis, vol. 7, pp. 161-176.

[22] Hampel, F.R. (1974), The Influence Curve and its Role in Robust Estimation, Annals of Statistics, vol. 69, pp. 383-393.

[23] Hendry, D.F. (1984), Monte Carlo Experimentation in Econometrics, in: Z. Griliches, M.D. Intriligator (eds.), Handbook of Econometrics, Amsterdam: North-Holland, vol. 2, ch. 16, pp. 937-976.

[24] Jondrow, J., Lovell, C.A.K., Materov, I.S., Schmidt, P. (1982), On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model, Journal of Econometrics, vol. 19, pp. 233- 238.

[25] Kumbhakar, S.C., Lovell, C.A.K. (2000), Stochastic Frontier Analysis, Cambridge (Mass.): Cambridge University Press.

[26] Meeusen, W., van den Broeck, J. (1977), Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error, International Economic Review, vol. 18, pp. 435-444.

[27] Ondrich, J., Ruggiero, J. (2001), Efficiency Measurement in the Stochastic Frontier Model, European Journal of Operational Research, vol. 129, pp. 434-442.

[28] Resti, A. (2000), Efficiency Measurement for Multi-Product Industries: A Comparison of Classic and Recent Techniques Based on Simulated Data, European Journal of Operational Research, vol. 121, pp. 559-578.

[29] Ruggiero, J. (1999), Efficiency Estimation and Error Decomposition in the Stochastic Frontier Model: A Monte Carlo Analysis, European Journal of Operational Research, vol. 115, pp. 555-563.

[30] Ruggiero, J. (2007), A Comparison of DEA and the Stochastic Frontier Model Using Panel Data, International Transactions in Operational Research, vol. 14, pp. 259-266.

[31] Tulkens, H. (1993), On FDH Efficiency Analysis: Some Methodological Issues and Applications to Retail Banking, Courts, and Urban Transit, Journal of Productivity Analysis, vol. 4, pp. 183-210.

[32] Van Biesebroeck, J. (2007), Robustness of Productivity Estimates, Journal of Industrial Economics, vol. 55, pp. 529-569.

[33] Wilson, P.W. (2008), *FEAR*: A Software Package for Frontier Efficiency Analysis with R, Socio-Economic Planning Sciences, vol. 42, pp. 247-254.

Table 1: Baseline Results

			COLS	SFA	DEAc	DEAv	FDH	order-m	order-α
Mean Absolute Error (MAE):									
$\sigma_u = 0.2$	$\sigma_v = 0.05$	$n = 50$	0.059	0.040	0.037	0.046	0.099	0.126	0.124
		$n = 100$	0.066	0.038	0.043	0.045	0.110	0.153	0.172
		$n = 200$	0.079	0.032	0.049	0.046	0.086	0.167	0.204
	$\sigma_v = 0.10$	$n = 50$	0.125	0.063	0.085	0.077	0.097	0.121	0.120
		$n = 100$	0.135	0.058	0.092	0.080	0.101	0.141	0.160
		$n = 200$	0.160	0.054	0.113	0.093	0.085	0.154	0.193
	$\sigma_v = 0.15$	$n = 50$	0.181	0.083	0.131	0.112	0.098	0.119	0.119
		$n = 100$	0.195	0.070	0.142	0.121	0.103	0.134	0.152
		$n = 200$	0.231	0.069	0.175	0.146	0.096	0.146	0.184
$\sigma_u = 0.3$	$\sigma_v = 0.05$	$n = 50$	0.055	0.042	0.035	0.055	0.132	0.161	0.158
		$n = 100$	0.059	0.041	0.040	0.047	0.141	0.187	0.205
		$n = 200$	0.073	0.033	0.043	0.046	0.106	0.191	0.230
	$\sigma_v = 0.10$	$n = 50$	0.111	0.068	0.075	0.078	0.127	0.155	0.154
		$n = 100$	0.119	0.063	0.082	0.076	0.131	0.175	0.194
		$n = 200$	0.143	0.056	0.098	0.085	0.102	0.179	0.220
	$\sigma_v = 0.15$	$n = 50$	0.169	0.088	0.120	0.110	0.127	0.150	0.149
		$n = 100$	0.179	0.081	0.128	0.112	0.126	0.163	0.182
		$n = 200$	0.211	0.074	0.156	0.131	0.108	0.167	0.209
Spearman Rank Correlation (R_{SP}):									
$\sigma_u = 0.2$	$\sigma_v = 0.05$	$n = 50$	0.796	0.803	0.864	0.690	0.414	0.515	0.450
		$n = 100$	0.888	0.884	0.836	0.787	0.580	0.579	0.510
		$n = 200$	0.893	0.897	0.875	0.794	0.675	0.667	0.570
	$\sigma_v = 0.10$	$n = 50$	0.631	0.632	0.685	0.560	0.331	0.449	0.395
		$n = 100$	0.710	0.707	0.657	0.622	0.498	0.526	0.483
		$n = 200$	0.735	0.736	0.708	0.639	0.561	0.578	0.509
	$\sigma_v = 0.15$	$n = 50$	0.493	0.495	0.552	0.459	0.280	0.381	0.339
		$n = 100$	0.559	0.558	0.522	0.498	0.415	0.449	0.425
		$n = 200$	0.594	0.594	0.574	0.516	0.461	0.480	0.434
$\sigma_u = 0.3$	$\sigma_v = 0.05$	$n = 50$	0.844	0.858	0.924	0.735	0.485	0.607	0.535
		$n = 100$	0.939	0.936	0.903	0.860	0.687	0.725	0.662
		$n = 200$	0.937	0.944	0.931	0.858	0.759	0.785	0.713
	$\sigma_v = 0.10$	$n = 50$	0.741	0.743	0.804	0.644	0.433	0.550	0.495
		$n = 100$	0.828	0.824	0.780	0.742	0.615	0.663	0.620
		$n = 200$	0.842	0.844	0.820	0.752	0.674	0.713	0.656
	$\sigma_v = 0.15$	$n = 50$	0.631	0.633	0.688	0.557	0.380	0.487	0.440
		$n = 100$	0.710	0.707	0.664	0.634	0.535	0.588	0.561
		$n = 200$	0.735	0.736	0.710	0.647	0.583	0.624	0.583

Note: all figures are based on $B = 1000$ replications.

Table 2: Asymmetry Scenario

			COLS	SFA	DEAc	DEAv	FDH	order-m	order-α
Mean Absolute Error (MAE):									
$\sigma_u = 0.2$	$\sigma_v = 0.05$	$n = 50$	0.059	0.040	0.038	0.047	0.094	0.115	0.115
		$n = 100$	0.066	0.038	0.044	0.043	0.102	0.139	0.152
		$n = 200$	0.079	0.032	0.051	0.047	0.079	0.142	0.172
	$\sigma_v = 0.10$	$n = 50$	0.125	0.063	0.089	0.080	0.092	0.109	0.110
		$n = 100$	0.135	0.058	0.097	0.084	0.096	0.127	0.142
		$n = 200$	0.160	0.054	0.118	0.098	0.080	0.129	0.161
	$\sigma_v = 0.15$	$n = 50$	0.181	0.083	0.137	0.118	0.096	0.110	0.111
		$n = 100$	0.195	0.070	0.149	0.128	0.101	0.123	0.137
		$n = 200$	0.231	0.069	0.181	0.153	0.097	0.126	0.154
$\sigma_u = 0.3$	$\sigma_v = 0.05$	$n = 50$	0.055	0.042	0.035	0.054	0.123	0.145	0.145
		$n = 100$	0.059	0.041	0.042	0.044	0.130	0.170	0.183
		$n = 200$	0.073	0.033	0.044	0.046	0.096	0.163	0.194
	$\sigma_v = 0.10$	$n = 50$	0.111	0.068	0.078	0.080	0.118	0.137	0.139
		$n = 100$	0.119	0.063	0.086	0.078	0.122	0.158	0.173
		$n = 200$	0.143	0.056	0.102	0.088	0.094	0.150	0.183
	$\sigma_v = 0.15$	$n = 50$	0.169	0.088	0.126	0.115	0.119	0.135	0.136
		$n = 100$	0.179	0.081	0.135	0.119	0.122	0.149	0.165
		$n = 200$	0.211	0.074	0.162	0.137	0.105	0.143	0.175
Spearman Rank Correlation (R_{SP}):									
$\sigma_u = 0.2$	$\sigma_v = 0.05$	$n = 50$	0.796	0.803	0.868	0.684	0.445	0.545	0.498
		$n = 100$	0.888	0.884	0.831	0.808	0.607	0.660	0.601
		$n = 200$	0.893	0.897	0.879	0.801	0.699	0.733	0.658
	$\sigma_v = 0.10$	$n = 50$	0.631	0.632	0.692	0.555	0.366	0.462	0.430
		$n = 100$	0.710	0.707	0.652	0.635	0.512	0.570	0.533
		$n = 200$	0.735	0.736	0.714	0.648	0.572	0.618	0.565
	$\sigma_v = 0.15$	$n = 50$	0.493	0.495	0.560	0.456	0.300	0.378	0.356
		$n = 100$	0.559	0.558	0.515	0.505	0.426	0.473	0.455
		$n = 200$	0.594	0.594	0.580	0.525	0.467	0.503	0.469
$\sigma_u = 0.3$	$\sigma_v = 0.05$	$n = 50$	0.844	0.858	0.926	0.730	0.570	0.658	0.612
		$n = 100$	0.939	0.936	0.895	0.872	0.699	0.765	0.724
		$n = 200$	0.937	0.944	0.934	0.861	0.765	0.820	0.772
	$\sigma_v = 0.10$	$n = 50$	0.741	0.743	0.808	0.639	0.491	0.579	0.546
		$n = 100$	0.828	0.824	0.773	0.751	0.624	0.693	0.663
		$n = 200$	0.842	0.844	0.826	0.758	0.680	0.739	0.703
	$\sigma_v = 0.15$	$n = 50$	0.631	0.632	0.695	0.553	0.416	0.496	0.470
		$n = 100$	0.710	0.707	0.656	0.638	0.540	0.606	0.588
		$n = 200$	0.735	0.736	0.717	0.655	0.586	0.641	0.615

Note: all figures are based on $B = 1000$ replications.

Table 3: Decreasing Returns to Scale Scenario

			COLS	SFA	DEAc	DEAv	FDH	order-m	order-α
Mean Absolute Error (MAE):									
$\sigma_u = 0.2$	$\sigma_v = 0.05$	$n = 50$	0.059	0.040	0.089	0.046	0.095	0.117	0.115
		$n = 100$	0.066	0.038	0.099	0.044	0.102	0.138	0.153
		$n = 200$	0.079	0.032	0.132	0.044	0.079	0.144	0.175
	$\sigma_v = 0.10$	$n = 50$	0.125	0.063	0.119	0.076	0.093	0.111	0.111
		$n = 100$	0.135	0.058	0.130	0.078	0.095	0.124	0.140
		$n = 200$	0.160	0.054	0.172	0.091	0.080	0.130	0.163
	$\sigma_v = 0.15$	$n = 50$	0.181	0.083	0.157	0.111	0.096	0.111	0.111
		$n = 100$	0.195	0.070	0.171	0.120	0.100	0.121	0.135
		$n = 200$	0.231	0.069	0.220	0.144	0.096	0.127	0.156
$\sigma_u = 0.3$	$\sigma_v = 0.05$	$n = 50$	0.055	0.042	0.076	0.055	0.125	0.149	0.147
		$n = 100$	0.059	0.041	0.085	0.047	0.129	0.167	0.181
		$n = 200$	0.073	0.033	0.115	0.044	0.096	0.165	0.198
	$\sigma_v = 0.10$	$n = 50$	0.111	0.068	0.104	0.077	0.120	0.141	0.142
		$n = 100$	0.119	0.063	0.113	0.074	0.119	0.154	0.170
		$n = 200$	0.143	0.056	0.154	0.083	0.094	0.152	0.186
	$\sigma_v = 0.15$	$n = 50$	0.169	0.088	0.140	0.109	0.121	0.137	0.137
		$n = 100$	0.179	0.081	0.152	0.111	0.118	0.144	0.160
		$n = 200$	0.211	0.074	0.195	0.130	0.104	0.143	0.177
Spearman Rank Correlation (R_{SP}):									
$\sigma_u = 0.2$	$\sigma_v = 0.05$	$n = 50$	0.796	0.803	0.592	0.689	0.438	0.553	0.487
		$n = 100$	0.888	0.884	0.696	0.799	0.621	0.652	0.588
		$n = 200$	0.893	0.897	0.699	0.806	0.702	0.721	0.640
	$\sigma_v = 0.10$	$n = 50$	0.631	0.632	0.508	0.551	0.356	0.470	0.421
		$n = 100$	0.710	0.707	0.591	0.630	0.520	0.565	0.531
		$n = 200$	0.735	0.736	0.622	0.648	0.574	0.608	0.555
	$\sigma_v = 0.15$	$n = 50$	0.493	0.495	0.432	0.449	0.299	0.391	0.354
		$n = 100$	0.559	0.558	0.491	0.502	0.427	0.468	0.451
		$n = 200$	0.594	0.594	0.538	0.522	0.468	0.495	0.461
$\sigma_u = 0.3$	$\sigma_v = 0.05$	$n = 50$	0.844	0.858	0.744	0.737	0.529	0.647	0.581
		$n = 100$	0.939	0.936	0.800	0.868	0.731	0.773	0.725
		$n = 200$	0.937	0.944	0.812	0.866	0.778	0.819	0.765
	$\sigma_v = 0.10$	$n = 50$	0.741	0.743	0.664	0.640	0.470	0.575	0.527
		$n = 100$	0.828	0.824	0.720	0.749	0.641	0.696	0.664
		$n = 200$	0.842	0.844	0.749	0.759	0.687	0.735	0.694
	$\sigma_v = 0.15$	$n = 50$	0.631	0.632	0.583	0.549	0.406	0.501	0.461
		$n = 100$	0.710	0.707	0.630	0.639	0.551	0.606	0.588
		$n = 200$	0.735	0.736	0.672	0.653	0.590	0.635	0.607

Note: all figures are based on $B = 1000$ replications.

Table 4: Increasing Returns to Scale Scenario

			COLS	SFA	DEAC	DEAV	FDH	order-m	order-α
Mean Absolute Error (MAE):									
$\sigma_u = 0.2$	$\sigma_v = 0.05$	$n = 50$	0.059	0.040	0.068	0.049	0.102	0.136	0.132
		$n = 100$	0.066	0.038	0.064	0.049	0.115	0.168	0.191
		$n = 200$	0.079	0.032	0.082	0.054	0.092	0.189	0.232
	$\sigma_v = 0.10$	$n = 50$	0.125	0.063	0.107	0.082	0.100	0.130	0.129
		$n = 100$	0.135	0.058	0.108	0.086	0.107	0.157	0.179
		$n = 200$	0.160	0.054	0.135	0.101	0.090	0.177	0.223
	$\sigma_v = 0.15$	$n = 50$	0.181	0.083	0.148	0.117	0.101	0.128	0.127
		$n = 100$	0.195	0.070	0.153	0.126	0.106	0.148	0.170
		$n = 200$	0.231	0.069	0.190	0.153	0.097	0.166	0.212
$\sigma_u = 0.3$	$\sigma_v = 0.05$	$n = 50$	0.055	0.042	0.060	0.054	0.137	0.172	0.167
		$n = 100$	0.059	0.041	0.059	0.050	0.150	0.205	0.227
		$n = 200$	0.073	0.033	0.073	0.052	0.114	0.216	0.261
	$\sigma_v = 0.10$	$n = 50$	0.111	0.068	0.095	0.081	0.133	0.166	0.164
		$n = 100$	0.119	0.063	0.096	0.081	0.140	0.194	0.217
		$n = 200$	0.143	0.056	0.117	0.091	0.109	0.204	0.252
	$\sigma_v = 0.15$	$n = 50$	0.169	0.088	0.135	0.114	0.132	0.161	0.160
		$n = 100$	0.179	0.081	0.138	0.117	0.134	0.181	0.204
		$n = 200$	0.211	0.074	0.169	0.138	0.112	0.192	0.241
Spearman Rank Correlation (R_{SP}):									
$\sigma_u = 0.2$	$\sigma_v = 0.05$	$n = 50$	0.796	0.803	0.804	0.690	0.389	0.482	0.417
		$n = 100$	0.888	0.884	0.732	0.761	0.543	0.509	0.447
		$n = 200$	0.893	0.897	0.697	0.760	0.652	0.614	0.507
	$\sigma_v = 0.10$	$n = 50$	0.631	0.633	0.697	0.571	0.312	0.431	0.371
		$n = 100$	0.710	0.707	0.599	0.609	0.477	0.486	0.438
		$n = 200$	0.735	0.736	0.595	0.619	0.546	0.543	0.462
	$\sigma_v = 0.15$	$n = 50$	0.493	0.495	0.591	0.472	0.262	0.370	0.322
		$n = 100$	0.559	0.558	0.490	0.490	0.403	0.427	0.398
		$n = 200$	0.594	0.594	0.502	0.503	0.454	0.461	0.406
$\sigma_u = 0.3$	$\sigma_v = 0.05$	$n = 50$	0.844	0.858	0.878	0.734	0.461	0.573	0.499
		$n = 100$	0.939	0.936	0.837	0.842	0.648	0.674	0.597
		$n = 200$	0.937	0.944	0.816	0.836	0.738	0.746	0.660
	$\sigma_v = 0.10$	$n = 50$	0.741	0.743	0.799	0.649	0.411	0.528	0.470
		$n = 100$	0.828	0.824	0.734	0.729	0.591	0.627	0.573
		$n = 200$	0.842	0.844	0.736	0.735	0.661	0.686	0.615
	$\sigma_v = 0.15$	$n = 50$	0.631	0.632	0.709	0.566	0.356	0.472	0.421
		$n = 100$	0.710	0.707	0.635	0.625	0.520	0.566	0.531
		$n = 200$	0.735	0.736	0.649	0.635	0.574	0.607	0.555

Note: all figures are based on $B = 1000$ replications.

Table 5: CES Scenario with Small Elasticity of Substitution

			COLS	SFA	DEAc	DEAv	FDH	order-m	order-α
Mean Absolute Error (MAE):									
$\sigma_u = 0.2$	$\sigma_v = 0.05$	$n = 50$	0.059	0.040	0.039	0.046	0.097	0.124	0.121
		$n = 100$	0.066	0.038	0.045	0.045	0.110	0.152	0.170
		$n = 200$	0.079	0.032	0.052	0.048	0.086	0.164	0.200
	$\sigma_v = 0.10$	$n = 50$	0.125	0.063	0.090	0.080	0.095	0.118	0.118
		$n = 100$	0.135	0.059	0.098	0.084	0.101	0.140	0.158
		$n = 200$	0.160	0.054	0.119	0.098	0.085	0.151	0.190
	$\sigma_v = 0.15$	$n = 50$	0.179	0.085	0.138	0.117	0.097	0.117	0.117
		$n = 100$	0.195	0.070	0.149	0.126	0.103	0.133	0.150
		$n = 200$	0.231	0.069	0.182	0.151	0.096	0.143	0.180
$\sigma_u = 0.3$	$\sigma_v = 0.05$	$n = 50$	0.055	0.042	0.035	0.053	0.129	0.157	0.155
		$n = 100$	0.059	0.041	0.042	0.047	0.142	0.186	0.203
		$n = 200$	0.073	0.033	0.045	0.047	0.106	0.189	0.227
	$\sigma_v = 0.10$	$n = 50$	0.111	0.067	0.079	0.080	0.124	0.150	0.150
		$n = 100$	0.119	0.064	0.087	0.079	0.130	0.173	0.192
		$n = 200$	0.143	0.056	0.104	0.088	0.102	0.176	0.216
	$\sigma_v = 0.15$	$n = 50$	0.168	0.088	0.127	0.114	0.124	0.146	0.146
		$n = 100$	0.178	0.081	0.135	0.117	0.126	0.161	0.180
		$n = 200$	0.211	0.074	0.163	0.136	0.108	0.165	0.205
Spearman Rank Correlation (R_{SP}):									
$\sigma_u = 0.2$	$\sigma_v = 0.05$	$n = 50$	0.796	0.802	0.868	0.687	0.430	0.517	0.461
		$n = 100$	0.888	0.884	0.832	0.781	0.570	0.577	0.514
		$n = 200$	0.893	0.897	0.877	0.790	0.671	0.669	0.574
	$\sigma_v = 0.10$	$n = 50$	0.630	0.632	0.690	0.559	0.351	0.454	0.408
		$n = 100$	0.708	0.705	0.651	0.613	0.487	0.521	0.482
		$n = 200$	0.735	0.736	0.709	0.636	0.556	0.578	0.512
	$\sigma_v = 0.15$	$n = 50$	0.502	0.503	0.565	0.469	0.298	0.389	0.354
		$n = 100$	0.559	0.558	0.517	0.490	0.404	0.441	0.421
		$n = 200$	0.594	0.594	0.574	0.513	0.457	0.479	0.436
$\sigma_u = 0.3$	$\sigma_v = 0.05$	$n = 50$	0.844	0.857	0.926	0.735	0.508	0.612	0.550
		$n = 100$	0.939	0.936	0.896	0.854	0.670	0.719	0.661
		$n = 200$	0.937	0.944	0.933	0.853	0.747	0.781	0.708
	$\sigma_v = 0.10$	$n = 50$	0.741	0.744	0.807	0.643	0.457	0.556	0.512
		$n = 100$	0.828	0.824	0.774	0.735	0.602	0.657	0.617
		$n = 200$	0.842	0.844	0.821	0.748	0.667	0.711	0.656
	$\sigma_v = 0.15$	$n = 50$	0.631	0.632	0.692	0.557	0.401	0.494	0.455
		$n = 100$	0.708	0.705	0.655	0.624	0.524	0.579	0.556
		$n = 200$	0.735	0.736	0.710	0.644	0.578	0.622	0.583

Note: all figures are based on $B = 1000$ replications.

Table 6: CES Scenario with Large Elasticity of Substitution

			COLS	SFA	DEAc	DEAv	FDH	order-m	order-α
Mean Absolute Error (MAE):									
$\sigma_u = 0.2$	$\sigma_v = 0.05$	$n = 50$	0.059	0.040	0.042	0.047	0.095	0.121	0.120
		$n = 100$	0.067	0.038	0.049	0.047	0.109	0.151	0.168
		$n = 200$	0.079	0.032	0.056	0.050	0.085	0.161	0.197
	$\sigma_v = 0.10$	$n = 50$	0.124	0.063	0.096	0.083	0.094	0.117	0.117
		$n = 100$	0.134	0.058	0.102	0.086	0.101	0.138	0.156
		$n = 200$	0.160	0.054	0.125	0.102	0.084	0.148	0.186
	$\sigma_v = 0.15$	$n = 50$	0.179	0.084	0.146	0.121	0.096	0.115	0.116
		$n = 100$	0.194	0.069	0.155	0.129	0.103	0.131	0.148
		$n = 200$	0.231	0.069	0.189	0.156	0.097	0.141	0.177
$\sigma_u = 0.3$	$\sigma_v = 0.05$	$n = 50$	0.056	0.042	0.037	0.052	0.126	0.153	0.152
		$n = 100$	0.060	0.041	0.045	0.048	0.141	0.184	0.201
		$n = 200$	0.073	0.033	0.049	0.048	0.105	0.186	0.223
	$\sigma_v = 0.10$	$n = 50$	0.111	0.068	0.085	0.082	0.122	0.147	0.148
		$n = 100$	0.120	0.064	0.092	0.081	0.129	0.170	0.189
		$n = 200$	0.143	0.056	0.109	0.092	0.101	0.173	0.212
	$\sigma_v = 0.15$	$n = 50$	0.167	0.087	0.133	0.117	0.123	0.144	0.145
		$n = 100$	0.177	0.080	0.140	0.119	0.125	0.159	0.178
		$n = 200$	0.211	0.074	0.168	0.140	0.108	0.162	0.201
Spearman Rank Correlation (R_{SP}):									
$\sigma_u = 0.2$	$\sigma_v = 0.05$	$n = 50$	0.794	0.799	0.871	0.689	0.442	0.516	0.470
		$n = 100$	0.888	0.884	0.830	0.774	0.565	0.578	0.518
		$n = 200$	0.893	0.896	0.878	0.787	0.672	0.672	0.582
	$\sigma_v = 0.10$	$n = 50$	0.632	0.633	0.695	0.561	0.361	0.452	0.415
		$n = 100$	0.711	0.708	0.650	0.611	0.483	0.518	0.481
		$n = 200$	0.735	0.736	0.709	0.633	0.554	0.579	0.515
	$\sigma_v = 0.15$	$n = 50$	0.502	0.504	0.567	0.469	0.307	0.388	0.359
		$n = 100$	0.561	0.560	0.514	0.485	0.396	0.435	0.418
		$n = 200$	0.594	0.594	0.574	0.510	0.456	0.479	0.438
$\sigma_u = 0.3$	$\sigma_v = 0.05$	$n = 50$	0.843	0.854	0.927	0.734	0.528	0.614	0.564
		$n = 100$	0.939	0.937	0.892	0.848	0.663	0.717	0.662
		$n = 200$	0.937	0.944	0.933	0.849	0.745	0.781	0.711
	$\sigma_v = 0.10$	$n = 50$	0.741	0.742	0.810	0.644	0.472	0.558	0.521
		$n = 100$	0.828	0.825	0.770	0.729	0.596	0.654	0.616
		$n = 200$	0.842	0.844	0.821	0.744	0.664	0.710	0.658
	$\sigma_v = 0.15$	$n = 50$	0.632	0.633	0.696	0.560	0.408	0.492	0.460
		$n = 100$	0.710	0.707	0.652	0.621	0.521	0.577	0.553
		$n = 200$	0.735	0.736	0.710	0.641	0.576	0.621	0.584

Note: all figures are based on $B = 1000$ replications.

Table 7: Outlier Scenario with Small Extent of Outliers

			COLS	SFA	DEAc	DEAv	FDH	order-m	order-α
Mean Absolute Error (MAE):									
$\sigma_u = 0.2$	$\sigma_v = 0.05$	$n = 50$	0.156	0.042	0.075	0.063	0.096	0.120	0.119
		$n = 100$	0.136	0.057	0.074	0.065	0.106	0.150	0.170
		$n = 200$	0.260	0.057	0.172	0.105	0.088	0.149	0.194
	$\sigma_v = 0.10$	$n = 50$	0.157	0.059	0.100	0.085	0.096	0.117	0.117
		$n = 100$	0.160	0.064	0.110	0.094	0.101	0.140	0.160
		$n = 200$	0.256	0.062	0.182	0.127	0.091	0.142	0.185
	$\sigma_v = 0.15$	$n = 50$	0.196	0.081	0.138	0.115	0.100	0.117	0.116
		$n = 100$	0.207	0.073	0.153	0.130	0.104	0.135	0.153
		$n = 200$	0.279	0.069	0.211	0.164	0.104	0.139	0.177
$\sigma_u = 0.3$	$\sigma_v = 0.05$	$n = 50$	0.146	0.044	0.069	0.067	0.127	0.152	0.153
		$n = 100$	0.088	0.051	0.053	0.054	0.140	0.188	0.208
		$n = 200$	0.237	0.045	0.147	0.092	0.107	0.174	0.222
	$\sigma_v = 0.10$	$n = 50$	0.155	0.065	0.095	0.087	0.126	0.149	0.149
		$n = 100$	0.135	0.072	0.093	0.085	0.131	0.176	0.197
		$n = 200$	0.237	0.062	0.163	0.116	0.106	0.165	0.211
	$\sigma_v = 0.15$	$n = 50$	0.190	0.085	0.130	0.115	0.127	0.146	0.146
		$n = 100$	0.182	0.085	0.134	0.118	0.128	0.165	0.186
		$n = 200$	0.261	0.079	0.195	0.151	0.115	0.158	0.202
Spearman Rank Correlation (R_{SP}):									
$\sigma_u = 0.2$	$\sigma_v = 0.05$	$n = 50$	0.793	0.802	0.776	0.641	0.373	0.525	0.498
		$n = 100$	0.796	0.797	0.744	0.676	0.543	0.579	0.509
		$n = 200$	0.850	0.850	0.620	0.594	0.554	0.631	0.563
	$\sigma_v = 0.10$	$n = 50$	0.643	0.646	0.675	0.564	0.323	0.458	0.423
		$n = 100$	0.633	0.633	0.606	0.562	0.475	0.517	0.480
		$n = 200$	0.703	0.703	0.597	0.551	0.498	0.548	0.499
	$\sigma_v = 0.15$	$n = 50$	0.515	0.518	0.564	0.477	0.284	0.396	0.360
		$n = 100$	0.498	0.498	0.478	0.451	0.392	0.433	0.413
		$n = 200$	0.574	0.574	0.528	0.479	0.433	0.464	0.430
$\sigma_u = 0.3$	$\sigma_v = 0.05$	$n = 50$	0.844	0.859	0.868	0.696	0.430	0.593	0.550
		$n = 100$	0.885	0.888	0.871	0.808	0.648	0.703	0.644
		$n = 200$	0.901	0.902	0.743	0.711	0.655	0.747	0.706
	$\sigma_v = 0.10$	$n = 50$	0.747	0.754	0.786	0.638	0.408	0.545	0.508
		$n = 100$	0.780	0.781	0.752	0.702	0.578	0.642	0.602
		$n = 200$	0.817	0.817	0.720	0.672	0.605	0.682	0.648
	$\sigma_v = 0.15$	$n = 50$	0.644	0.646	0.690	0.564	0.374	0.492	0.455
		$n = 100$	0.666	0.666	0.637	0.601	0.508	0.567	0.543
		$n = 200$	0.716	0.716	0.660	0.606	0.541	0.600	0.573

Note: all figures are based on $B = 1000$ replications.

Table 8: Outlier Scenario with Large Extent of Outliers

			COLS	SFA	DEAC	DEAV	FDH	order-m	order-α
Mean Absolute Error (MAE):									
$\sigma_u = 0.2$	$\sigma_v = 0.05$	$n = 50$	0.162	0.044	0.092	0.067	0.096	0.119	0.119
		$n = 100$	0.189	0.076	0.131	0.101	0.099	0.140	0.162
		$n = 200$	0.256	0.075	0.196	0.157	0.090	0.152	0.200
	$\sigma_v = 0.10$	$n = 50$	0.174	0.061	0.109	0.087	0.096	0.116	0.116
		$n = 100$	0.203	0.068	0.146	0.117	0.098	0.133	0.153
		$n = 200$	0.277	0.070	0.211	0.173	0.094	0.145	0.190
	$\sigma_v = 0.15$	$n = 50$	0.206	0.081	0.145	0.119	0.100	0.116	0.116
		$n = 100$	0.234	0.072	0.175	0.145	0.104	0.131	0.149
		$n = 200$	0.306	0.071	0.237	0.200	0.107	0.143	0.183
$\sigma_u = 0.3$	$\sigma_v = 0.05$	$n = 50$	0.152	0.045	0.083	0.069	0.127	0.152	0.153
		$n = 100$	0.176	0.081	0.117	0.087	0.129	0.175	0.198
		$n = 200$	0.267	0.085	0.196	0.159	0.107	0.174	0.227
	$\sigma_v = 0.10$	$n = 50$	0.169	0.066	0.104	0.090	0.125	0.147	0.148
		$n = 100$	0.188	0.084	0.135	0.107	0.124	0.166	0.188
		$n = 200$	0.283	0.086	0.207	0.172	0.109	0.167	0.218
	$\sigma_v = 0.15$	$n = 50$	0.194	0.086	0.134	0.116	0.126	0.146	0.146
		$n = 100$	0.214	0.089	0.159	0.133	0.125	0.159	0.180
		$n = 200$	0.302	0.089	0.226	0.193	0.117	0.162	0.208
Spearman Rank Correlation (R_{SP}):									
$\sigma_u = 0.2$	$\sigma_v = 0.05$	$n = 50$	0.783	0.799	0.779	0.648	0.372	0.526	0.495
		$n = 100$	0.751	0.752	0.608	0.614	0.545	0.586	0.537
		$n = 200$	0.828	0.830	0.779	0.682	0.556	0.645	0.561
	$\sigma_v = 0.10$	$n = 50$	0.645	0.655	0.663	0.564	0.325	0.461	0.423
		$n = 100$	0.614	0.615	0.544	0.536	0.485	0.519	0.497
		$n = 200$	0.701	0.702	0.666	0.599	0.509	0.567	0.508
	$\sigma_v = 0.15$	$n = 50$	0.529	0.538	0.561	0.480	0.289	0.402	0.362
		$n = 100$	0.494	0.495	0.456	0.445	0.406	0.438	0.428
		$n = 200$	0.582	0.583	0.561	0.511	0.447	0.483	0.442
$\sigma_u = 0.3$	$\sigma_v = 0.05$	$n = 50$	0.835	0.862	0.861	0.703	0.426	0.590	0.549
		$n = 100$	0.843	0.845	0.718	0.728	0.638	0.698	0.655
		$n = 200$	0.896	0.899	0.844	0.758	0.639	0.754	0.705
	$\sigma_v = 0.10$	$n = 50$	0.744	0.759	0.777	0.642	0.412	0.551	0.510
		$n = 100$	0.756	0.757	0.674	0.666	0.580	0.639	0.611
		$n = 200$	0.814	0.817	0.771	0.700	0.602	0.692	0.651
	$\sigma_v = 0.15$	$n = 50$	0.647	0.655	0.685	0.566	0.379	0.497	0.457
		$n = 100$	0.652	0.653	0.599	0.582	0.513	0.563	0.547
		$n = 200$	0.719	0.720	0.690	0.630	0.552	0.616	0.582

Note: all figures are based on $B = 1000$ replications.