More is better than one: the impact of different numbers of input aggregators in technical efficiency estimation

Ana M. Aldanondo and Valero L. Casasnovas

Dept. of Business Administration. Public University of Navarre.

April 2015

More is better than one: the impact of different numbers of input aggregators in technical efficiency estimation

Ana M. Aldanondo [a] & Valero L. Casasnovas [a,*]

[a] Department of Business Administration, Public University of Navarre,
Campus Arrosadia, Pamplona, 31006, Spain

* Corresponding author. E-mail: valero.casasnovas@unavarra.es

WORKING PAPER

April 2015

Abstract

The results of an experiment with simulated data show that combining inputs with different criteria (as cost, material inputs aggregates and other) increases the accuracy of the Data Envelopment Analysis (DEA) technical efficiency estimator in data sets with dimensionality problems. The positive impact of this approach surpasses that of reducing the number of variables, since replacement of the original inputs with an equal number of aggregates improves DEA performance in a wide range of cases.

Key words: Technical efficiency, Aggregation bias, Monte Carlo, DEA Estimator accuracy

I. Introduction

Data Envelopment Analysis (DEA) is one of the most widely-used nonparametric frontier models for evaluating the technical efficiency of Decision Making Units (DMUs) in a multiple input/output scenario. The DEA radial technical efficiency estimator is statistically consistent; that is, it converges towards true efficiency with growing sample size (see Simar and Wilson, 2015, for a summary of DEA properties). Nevertheless, performing a DEA on real data with an inappropriate choice of inputs and/or outputs will generate a biased efficiency estimate (Smith, 1997). Also, even if the model is correctly specified, the DEA estimator, like many other non-parametric estimators, is prey to the curse of dimensionality; that is, its rate of convergence to true efficiency diminishes as more inputs and outputs are added.

While input (output) aggregation is standard practice in the specification of variables for use in DEA, its validity has been the object of academic research. Several authors (Primont, 1993; Tauer, 2001; Färe and Zelenyuk, 2002; Färe *et al*., 2004) show that radial technical efficiency measures specifying composite cost inputs (or revenue outputs) can be biased downwards by allocative inefficiency. The same authors use the term "aggregation bias" to refer to the gap between input- oriented technical efficiency scores obtained using aggregate versus multiple inputs/outputs. Simar and Wilson (2001) suggest various tests for additive inputs or outputs with difference-based statistics. Despite this bias, Podinovski and Thanassoulis (2007) consider that the use of composite inputs (and/or outputs) it is a practical remedy for reducing the number of variables and thus enhancing the discriminatory power of DEA. Ultimately, therefore, the results of these studies suggest that whether to use aggregate data in DEA technical efficiency estimation is an empirical question that depends on the dimensionality of the problem and the possibility of aggregation bias.

All the above-mentioned literature on the impact of input (output) aggregation on the performance of the DEA technical efficiency estimator focuses on the implications of collapsing several groups of inputs into a single composite. Very little research, however, has yet gone into assessing and comparing DEA performance with different numbers of aggregates of the same inputs. [1] Aldanondo and Casasnovas (2015) extend the analysis of technical efficiency input aggregation bias to the use of multiple aggregators. They conclude, firstly, that, in complex production problems, the use of different aggregates of the same inputs guarantees coherence between technical efficiency and the various, conflicting criteria (cost minimization, optimisation of management resources, the reduction of pollution by inputs, etc.) of overall efficiency. Secondly, they show that input aggregation bias diminishes with the number of aggregates of the same inputs. The greater the number of different linear aggregators of the same inputs, the closer the outcome to that of the original DEA technical efficiency estimator. Aldanondo and Casasnovas (2015) make no specific claims with respect to the empirical accuracy of the various estimators, however.

The purpose of this study is to build on previous research by exploring the implications for the radial DEA technical efficiency estimator when multiple composite inputs are used. In particular, we incorporate a measurement of the error in the comparison of the various estimators by applying a Monte Carlo simulation and generate decision rules for the use of multiple aggregators based on the particular conditions in each case.

[1] One exception to this is the combined use of Principal Component Analysis and DEA (PCA-DEA) (Adler and Yazhemsky, 2010), where multiple linear aggregates of the same inputs and outputs are used. One drawback of PCA-DEA is that it hinders interpretation of the results: the coefficients of the input and output aggregates can be negative (Yap *et al*., 2013).

The paper is organised as follows: section one describes the Monte Carlo design and the methodology used to analyse the DEA model; section two presents the results of the analysis of aggregation bias in technical efficiency when using multiple aggregate criteria. The paper ends with some conclusions from the research.

II. Experimental design

We use a Monte Carlo experiment to compare aggregation bias and the accuracy of the DEA estimator for several linear aggregates of the same inputs. Aggregation bias is approximated by estimating DEA efficiency scores for the baseline model with fully disaggregated data and comparing them with the efficiency scores obtained when some of the inputs in the DEA are linearly aggregated into several composite inputs. The accuracy of the models, including the baseline model, is determined by comparing the simulated true efficiency value with the DEA efficiency estimates. All comparisons are carried out for different numbers of observations *n* Є (10, 50, 100, 500, 1 000, 2 000, and 5 000) and the degree of inefficiency is defined by the standard deviation of inefficiency term σ_u (0.2, 0.3).

For ease of comparison with other studies, we conduct a variation on an experiment used by Tauer (2001), where we assume that technology is characterised by a Cobb-Douglas production function, $y_i^e = \prod_k x_{ik}{}^{\alpha_k}$,with constant economies of scale, $\sum_k \alpha_k = 1$, one single efficient output y_i^e and five inputs x_{ik} (k=1,…,5) for each observation *i*.. Our choice of ranges of variation of inputs, outputs and efficiency are guided by the values used in Simar and Wilson (2001) and Banker *et al.* (1993). The experiment consists of 1 000 replications of the following procedure.

1) Five parameters α_k are generated from a uniform distribution [0.1, 1] and each α_k is divided by the sum of the five selected α_k, such that the coefficients add up to 1.

2) A uniform distribution [0.1, 100] generates the single efficient output y_i^e and the five w_{ik} factor prices are drawn from independent random variables with uniform distribution [0.1, 5]. The quantity of inputs is computed by means of the factor demand function: $x_{ik} = \left(\prod \alpha_k^{\alpha_k}\right)^{-1} \alpha_k y_i^e w_{ik}^{-1} \left(\prod w_{ik}^{\alpha_k}\right)$.

3) Inefficiency is simulated by multiplying the output of each unit y_i^e by the technical inefficiency coefficient $A_i = \exp(-u_i)$, where u_i is a random value drawn from a normal distribution $N(0, \sigma_u)$. Then, the observed output value of each unit *i* is computed as $y_i = A_i y_i^e$

4) Four inputs are linearly aggregated four times using as weights the corresponding prices of these inputs for the first four units of the sample: respectively, $C_i^j = \sum_{k=1}^{4} w_{jk} x_{ik}$ (*i*=1,…,*n* ; *j*=1,…,4; *k*=1,…,4) where C_i^j denotes the aggregate of the *k* inputs of unit *i*, weighted by the w_{jk} input prices of unit *j*.

5) From this initial 5 000-unit population, we take subsamples of the first 10, 50, 100, 500, 1 000, and 2 000 observations in order to obtain smaller samples. Thus, this study simulates change in sample size as successive enlargements up to population size, thereby maintaining the same technology and the same aggregate weights for different-sized samples in each replication.

6) The linear programming in Equation (1) is used to compute radial technical input efficiency $\hat{A}_i^h$ with constant returns to scale (Charnes et al., 1978) for unit *i*, with models with different number of aggregates (*h*=0,1,2,3,4). The baseline model $\hat{A}_i^0$ computes the efficiency scores obtained with the five original inputs. The other models include one or several aggregates of the first four inputs and the fifth original input. [2]

$$
\begin{aligned}
&\hat{A}_i^h(y_i, C_i^j, x_{ik}) = \min \beta \\
&\text{subject to} \\
&\sum_i z_i y_i \geq y_i \\
&\sum_i z_i C_i^j \delta^h \leq \beta C_i^j \delta^h \quad j = 1,...,h;\ \delta^h = 0 \text{ if } \mathrm{h} = 0 \text{ and } \delta^h = 1 \text{ if } \mathrm{h} > 0 \\
&\sum_i z_i x_{ik} \leq \beta x_{ik} \quad k = 1,...,5 \text{ if } h = 0 \text{ and } k = 5 \text{ if } h > 0 \\
&z_i \geq 0 \quad i = 1,..., n
\end{aligned}
\tag{1}
$$

All the efficiency scores are computed using FEAR software (Wilson, 2008) for platform R.

7) For every replication, we compute average technical efficiency scores, mean aggregation bias (difference between the baseline DEA estimation model and the estimators of models with aggregates), $MAB = \frac{1}{n}\sum_i^n \hat{A}_i^0 - \hat{A}_i^h$ and mean absolute error (absolute difference between estimated efficiency and true or simulated efficiency), $MAE = \frac{1}{n}\sum_{i=1}^n \left|\hat{A}_i^h - A_i\right|$. [3] It should be noted that, in applied research using real data, *MAB* is the only possible statistic for measuring the goodness-of-fit of the efficiency

[2]. Thus, the aggregate models have different numbers of composite inputs and one original input. We have replicated this experiment with different numbers of original and aggregated inputs. The results, which are similar to the case presented here, are available from the authors upon request.

[3] Spearman rank correlation coefficients estimated to measure the accuracy of the estimators (which have no impact on the findings) are omitted for lack of space and because our aim is to compare *MAB* and *MAE*.

estimators of the various models, while, in experimental scenarios, there is an appreciable difference between the *MAB* and the *MAE,* with the latter showing the true accuracy of each estimator.

8) Finally, the results presented in the next section are the average over the 1 000 replications of the average technical efficiency, the *MAE* and *MAB.*

III. Results and discussion

Table 1 and Table 2 give the estimates for standard deviations of inefficiency of 0.2 and 0.3, respectively. For the sake of clarity, both tables include average technical efficiency, mean aggregation bias *(MAB*) and mean absolute error (*MAE*). The results will be discussed in blocks, starting with the average technical efficiency scores, which are the indicators most widely discussed in the literature cited above. This will be followed by an analysis of the *MAE* in each model. Lastly, the advantages and disadvantages of aggregating data will be discussed and the *MAB* and *MAE* analysed in order to test the capacity of the former as a goodness-of-fit estimator.

The average technical efficiency scores uphold some known theoretical and experimental findings reported by Fare *et al*. (2004) and Tauer (2001). Firstly, the DEA average efficiency score for the baseline model, using the five original inputs, is well above the true average efficiency for small sample sizes, converging towards true efficiency with growing sample size. This can be checked by looking at the average efficiency trends displayed in Table 1 and Table 2. In Table 1, for example, the DEA baseline model average efficiency diminishes from 0.993 (*n*=10) to 0.890 (*n*=5 000) for a true average efficiency of 0.858. Secondly, the average efficiency score obtained using models with input aggregators is biased downwards relative to that given by the model with fully disaggregated inputs. This bias decreases as more aggregates of the same inputs are added. Thirdly, the average efficiency score for models with composite inputs falls below true efficiency as the sample size increases. Again, Table 1 shows that the average efficiency score given by the model with one aggregator drops to 0.706 (n=5 000) and to 0.817 (n=5 000) for the model with four aggregators. This has been reported by Tauer (2001) as evidence of the inconsistency of DEA technical efficiency estimation using aggregates, since the computed average efficiency score does not converge towards true average efficiency.

The true error values, that is, the *MAE* scores, confirm some of the above observations while also providing new findings. With respect to aggregation bias, the trend of the *MAE* as a function of sample size confirms the inconsistency of DEA efficiency estimators when using aggregates for non-additive inputs. As can be seen from both tables, the *MAE* of the models using aggregates does not converge towards zero with larger sample size in any of the models, while the *MAE* of the baseline model DEA estimator without aggregates always decreases with larger sample size. For example, with four aggregates of four inputs and σ_u=0.2, the *MAE* score decreases gradually from 0.104 to 0.049 as sample size grows from 10 to 1 000 units, and increases slightly to 0.051 when sample size reaches 5 000 units. This stabilisation effect or increase in *MAE* appears in all the models using aggregates reported in Tables 1 and 2, highlighting the fact that an increase in sample size does not correct the true error or bias in a misspecification of the variables.

The *MAE* performance, however, suggests that it is better to use aggregates when faced with dimensionality problems. Indeed, although the aggregate models contain some bias and do not converge towards true efficiency, they may have greater estimation accuracy in certain empirical contexts than the baseline model. As can be seen in Table 2 for the case of n=10-unit, for example, the *MAE* for the one-aggregate model is 0.111, which is lower than in the baseline model (*MAE*=0.211) and all the other models. Conversely, using the same table, the lowest *MAE* for n=50 is found in the model with 3 aggregators and, for n=100, in the model with four aggregators. Generally speaking, the results show that models with fewer aggregates produce better estimates with small sample size and high standard deviation of efficiency; while the accuracy of estimators using a larger number of linear aggregates improves as sample size grows. When the sample size is large enough to eliminate dimensionality problems, the basic DEA estimator without aggregates gives the best performance.

As far as we are aware, it has never been reported in experimental studies that the DEA efficiency estimator using a number of aggregates equal to the number of replaced original inputs could give a better result than the model using fully disaggregated inputs, despite that both programs have the same dimension. [4] It can be seen, for example, that the model with four aggregates has a lower *MAE* than the model with the five original inputs when applied to samples of 1 000 units or less for a σ_u=0.2 (Table 1) and to samples of 2 000 units or less for a σ_u=0.3 (Table 2). This holds even for a production function without additive inputs, like the one specified herein.

Thus, aggregation of the inputs achieves more than the mere reduction of the number of variables in the DEA program. Indeed, one drawback of the DEA radial technical efficiency measure is that it does not capture all sources of inefficiency, because this measure fails to take into account the non cero slacks in inputs and outputs. The slack trouble increases with the dimensionality of the problem: the greater the number of

[4] In contrast to PCA-DEA studies, when the number of aggregates is equal to the number of inputs or outputs (explaining the total variance of the sample), the DEA efficiency estimator gives the same result with or without aggregates (Adler and Golany, 2001).

inputs and outputs, the less likely the efficiency score incorporates all excesses in inputs and shortfalls in outputs. Allen *et al.* (1997) propose weight constraints in (dual) DEA as a way to solve slack allocation problems. Charnes *et al.* (1990) and Podinovski and Thanassoulis (2007) show that, in very particular circumstances, [5] replacement of the inputs with linear aggregates in the primal DEA program is equivalent to constraining input weights in the corresponding multiplicative DEA model. [6] Our results suggest that the impact of the multiple aggregation of the same inputs in the DEA model works in two ways: firstly, the higher the degree of aggregation, the fewer slacks (or zero weights) will appear (Olesen and Petersen, 1996; Førsund, 2013); and, secondly, the more linear aggregators of the same inputs are included, the closer the estimate to the true frontier of efficiency (Varian, 1984; Banker and Maindiratta, 1988).

Finally, the experiment allows comparison between *MAB* and *MAE*. The *MAB* statistic is usually used to compare the performance of various DEA models or measure input and output additivity, since it is the only available statistic when estimating efficiency using real data. In our case, we are able to identify the conditions under which the *MAB* is the appropriate indicator for establishing superiority criteria for input aggregation in DEA by analyzing the differences between *MAB* and *MAE*. We focus on comparing models with different numbers of aggregates, since additivity tests based on the *MAB* can be found in Simar and Wilson (2001). As can be seen from the tables, the *MAB* is

[5] Charnes *et al*. (1990) demonstrate the mathematical equivalence of constraining relative weights (assurance region) and aggregating inputs and outputs, for the very simple case in which one constrains the variation of the relative weights of the only two inputs of the input set within an interval of R^+. Podinovski and Thanassoulis (2007) suggest the equivalence of aggregating inputs or outputs in the baseline DEA and constraining their relative weights in the dual form of the DEA model. This, however, would hold whenever the constraint on the respective aggregate input (output) in the DEA model holds.

[6] The relation between the relative weight constraints and the use of input or output aggregates for multiple constraints, has, to our knowledge, never been described. See Førsund (2013) for a recent review of the research on weight constraints.

lower in the model using four aggregates of the same inputs than in those using fewer than four. With small samples, however, the *MAE* in models with fewer aggregates is lower than in those with larger numbers of aggregates. Thus, our results suggest that the *MAB* discriminates well between models that are free of dimensionality problems. This finding is line with that of Simar and Wilson (2001) and consistent with the functional form of the estimated aggregation bias. [7]

IV. Conclusions

The main conclusion from this research is that the use of multiple linear aggregates of the same inputs has a positive impact on the performance of the radial DEA efficiency estimator in the presence of dimensionality problems. Our results show that this positive effect outweighs the known effect of reducing the number of variables in the DEA program. Indeed, in several cases the mean absolute error (*MAE*) of the model with four linear aggregates of the same four inputs is lower than the *MAE* of the program using the original inputs. With no dimensionality problems and no additive inputs, DEA technical efficiency models with fully disaggregated inputs are the most appropriate method.

Estimators with multiple aggregates of the same inputs perform better overall than those with a single aggregate, except when applied to very small samples with high standard deviation of inefficiency. These results have major implications for DEA efficiency

---

[7] The estimated aggregation bias AB_i^h which can be broken down as follows:

$$AB_i^h = \hat{A}_i^0 - \hat{A}_i^h = (\hat{A}_i^0 - A_i) - (\hat{A}_i^h - A_i) \qquad h = 1, 2, 3, 4$$

Thus, AB_i^h is a better estimator of true error, $\hat{A}_i^h - A_i$, when the estimated efficiency of unit $\hat{A}_i^0$ converges towards true efficiency A_i, with growing sample size. If $\hat{A}_i^h < A_i$, AB_i^h is the sum of the two errors.

estimation. The use of multiple rather than a single linear aggregate of the same inputs can improve the performance of the radial DEA efficiency estimator while also ensuring coherence between technical efficiency measures and multiple criteria of overall efficiency.

References

Adler, N. and Golany, B. (2001) Evaluation of deregulated airline networks using data envelopment analysis combined with principal component analysis with an application to Western Europe, *European Journal of Operational Research*,**132**, 260–73.

Adler, N. and Yazhemsky, E. (2010) Improving discrimination in data envelopment analysis: PCA–DEA or variable reduction, *European Journal of Operational Research* **202**, 273–84.

Aldanondo, A. M. and Casasnovas; V. L. (2015) Input aggregation bias in technical efficiency with multiple criteria analysis, *Applied Economic Letters*, 22, 430-35.

Allen, R., Athanassopoulos, A., Dyson, R. G. and Thanassoulis, E. (1997) Weights restrictions and value judgements in data envelopment analysis: Evolution, development and future directions, *Annals of Operations Research*, **73**, 13–34.

Allen, R. and Thanassoulis, E. (2004) Improving envelopment in data envelopment analysis, *European Journal of Operational Research*, **154**, 363–79

Banker, R. D., Gadh, V. M. and 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*, **67**, 332-43.

Banker, R. D. and Maindiratta, A. (1988) Nonparametric analysis of technical and allocative efficiencies in production, *Econometrica*, **56**, 1315–32.

Charnes, A., Cooper, W. W. and Huang, Z. M. (1990) Polyhedral cone-ratio DEA models with an illustrative application to large commercial banks, *Journal of Econometrics*, **46**, 73–91.

Charnes, A., Cooper, W. W. and Rhodes, E. (1978) Measuring the efficiency of decision-making units, *European Journal of Operational Research*, **2**, 429–44.

Färe, R., Grosskopf, S. and Zelenyuk, V. (2004) Aggregation bias and its bounds in measuring technical efficiency, *Applied Economics Letters*, **11**, 657–60.

Färe, R. and Zelenyuk, V. (2002) Input aggregation and technical efficiency, *Applied Economics Letters*, **9**, 635–6.

Førsund, F. R. (2013) Weight restrictions in DEA: misplaced emphasis?, *Journal of Productivity Analysis*, **40**, 271–83.

Olesen, O. B. and Petersen, N. C. (1996) Indicators of ill-conditioned data sets and model misspecification in data envelopment analysis: an extended facet approach, *Management Science*, **42**, 205–19.

Podinovski, V. V and Thanassoulis, E. (2007). Improving discrimination in data envelopment analysis: some practical suggestions, *Journal of Productivity Analysis*, **28**, 117–26.

Primont, D. (1993) Efficiency measures and input aggregation, in *Mathematical Modelling in Economics. Essays in Honor of Wolfgang Eichhorn*, Diewert, W. E., Spremann, K. and Stehling, F. (Eds), Springer, Berlin, pp. 288–94.

Simar, L. and Wilson, P. W. (2001) Testing restrictions in nonparametric efficiency models, *Communications in Statistics: Simulation and Computation*, **30**, 159–84.

Simar, L. and Wilson, P. W. (2015) Statistical approaches for nonparametric frontier models: a guided tour, *International Statistical Review*, **83**, 77–110.

Smith, P., (1997) Model misspecification in data envelopment analysis, *Annals of Operations Research*, **73**, 233–52.

Tauer, L. W. (2001) Input aggregation and computed technical efficiency, *Applied Economics Letters*, **8**, 295–7.

Varian, H. R. (1984) The nonparametric approach to production analysis, *Econometrica*, **52**, 579-97.

Wilson, P. W. (2008). FEAR: a software package for frontier efficiency analysis with R. *Socio-Economic Planning Sciences*, **42**, 247–54.

Yap, G. L. C., Ismail, W. R., and Isa, Z. (2013) An alternative approach to reduce dimensionality in data envelopment analysis, *Journal of Modern Applied Statistical Methods*, **12**, 128-47.

Table 1. Computed average technical efficiency, mean aggregation bias and mean absolute error (from random production data and σ_u=0.2)

	DEA Basic	1 agg.	2 agg.	3 agg.	4 agg.
			Efficiency scores		
n=10	0.993	0.864	0.920	0,942	0.953
	(0.020)*	(0.140)	(0.108)	(0.089)	(0.077)
50	0.967	0.788	0.852	0.878	0.893
	(0.066)	(0.147)	(0.129)	(0.118)	(0.111)
100	0.953	0.765	0.831	0.857	0.873
	(0.077)	(0.145)	(0.130)	(0.121)	(0.115)
500	0.922	0.729	0.797	0.824	0.841
	(0.093)	(0.141)	(0.129)	(0.122)	(0.117)
1 000	0.910	0.719	0.787	0.815	0.832
	(0.096)	(0.140)	(0.128)	(0.122)	(0.117)
2 000	0.901	0.710	0.779	0.807	0.824
	(0.097)	(0.139)	(0.128)	(0.121)	(0.116)
5 000	0.890	0.706	0.772	0.801	0.817
	(0.099)	(0.137)	(0.127)	(0.120)	(0.116)
			Mean Aggregation Bias (*MAB*)		
n=10		0.128	0.073	0.051	0.039
50		0.178	0.115	0.089	0.074
100		0.187	0.122	0.096	0.080
500		0.193	0.125	0.097	0.081
1 000		0.192	0.124	0.096	0.079
2 000		0.191	0.122	0.094	0.077
5 000		0.184	0.118	0.089	0.072
			Mean Absolute Error (*MAE*)		
n=10	0.135	0.104	0.096	0.100	0.104
50	0.108	0.108	0.074	0.068	0.067
100	0.094	0.116	0.073	0.062	0.058
500	0.063	0.136	0.078	0.059	0.050
1 000	0.052	0.144	0.082	0.061	0.049
2 000	0.042	0.151	0.086	0.063	0.050
5 000	0.031	0.153	0.090	0.064	0.051

Notes: Standard deviation of computed efficiency in parentheses. True mean efficiency is 0.858 with standard deviation 0.097 for all sample sizes.

Table 2. Computed average technical efficiency, mean aggregation bias and mean absolute error (from random production data and σ_u=0.3)

	DEA Basic	1 agg.	2 agg.	3 agg.	4 agg.
			Efficiency scores		
n=10	0.979	0.837	0.892	0.913	0.924
	(0.049)*	(0.163)	(0.135)	(0.121)	(0.111)
50	0.935	0.751	0.810	0.835	0.850
	(0.107)	(0.166)	(0.154)	(0.148)	(0.143)
100	0.914	0.726	0.786	0.811	0.827
	(0.119)	(0.164)	(0.155)	(0.149)	(0.146)
500	0.873	0.687	0.749	0.775	0.790
	(0.132)	(0.157)	(0.152)	(0.148)	(0.145)
1 000	0.859	0.675	0.738	0.764	0.780
	(0.135)	(0.155)	(0.150)	(0.147)	(0.145)
2 000	0.847	0.666	0.730	0.756	0.772
	(0.136)	(0.154)	(0.149)	(0.146)	(0.144)
5 000	0.835	0.661	0.722	0.749	0.764
	(0.136)	(0.152)	(0.148)	(0.145)	(0.143)
			Mean Aggregation Bias (*MAB*)		
n=10		0.142	0.087	0.067	0.055
50		0.184	0.124	0.100	0.085
100		0.188	0.128	0.103	0.088
500		0.186	0.124	0.098	0.083
1 000		0.184	0.121	0.095	0.079
2 000		0.181	0.118	0.091	0.076
5 000		0.174	0.113	0.086	0.070
			Mean Absolute Error (*MAE*)		
n=10	0.211	0.111	0.116	0.124	0.130
50	0.136	0.099	0.076	0.075	0.076
100	0.115	0.104	0.071	0.064	0.063
500	0.074	0.122	0.072	0.056	0.049
1 000	0.060	0.130	0.075	0.056	0.047
2 000	0.048	0.137	0.078	0.057	0.046
5 000	0.035	0.140	0.082	0.058	0.047

Notes: Standard deviation of computed efficiency in parentheses. True mean efficiency is 0.799 with standard deviation 0.133 for all sample sizes.