JACKSTRAPPING DEA SCORES FOR ROBUST EFFICIENCY MEASUREMENT

Borko D. Stošić

Departamento de Física e Matemática, Universidade Federal Rural de Pernambuco
Rua Dom Manoel de Medeiros s/n, Dois Irmãos
CEP: 52171-900, Recife-PE, Brasil
Tel: 081-34781366; e-mail: borko@ufpe.br

Maria da Conceição Sampaio de Sousa

Departamento de Economia, Universidade de Brasília
Campus Universitário Darcy Ribeiro, Instituto Central de Ciências - Asa Norte
CEP: 70910-900, Brasília - DF – Brasil
Tel:061-347-5304; e-mail: mcss@unb.br

ÁREA: ECONOMETRIA

JACKSTRAPPING DEA SCORES FOR ROBUST EFFICIENCY MEASUREMENT

ABSTRACT

A new approach for robust DEA technical efficiency measurement is presented, based on a combination of Jackknife and Bootstrap resampling schemes. First, an algorithm implementing Jackknife is used to extract leverage of all data points, that is, the impact of removal of the observed point on DEA calculations performed on the rest of the data set. Next, Bootstrap stochastic resampling is implemented, taking into account leverage information. It is demonstrated that this approach proves robust to the presence of outliers and/or errors in the data set, and as it is completely automatic, it is suitable for implementation on very large data sets.

KEYWORDS
Data Envelopment Analysis, Bootstrap, Jackknife

1. Introduction

Data Envelopment Analysis (DEA), originally introduced by Farell (1957), is a non-parametric method for estimating technical efficiency of Decision Making Units (DMUs), by application of linear programming for comparative analysis of input and output variables. With the exponential advent of easily accessible computing power, over the past decades DEA has been steadily gaining momentum in different research areas, ranging from economics to social sciences (see e.g. Seiford (1996) and references therein). The DEA method is extremely attractive because of the fact that it does not require a priori knowledge of the functional relation between the input and output variables (technology), nor does it impose arbitrary statistical weights on variables. On the other hand, the method is based on the (extreme) concept of production frontier, so that a single error in the data set (or an exceptionally well performing unit – an outlier) may seriously compromise the analysis. Generally speaking, DEA yields successful results mainly in situations where the DMUs are "well behaved", and the input and output variables have balanced, non-dispersed values. On the other hand, if the data set contains DMUs that perform extremely well (which may stem from some outstanding practice, or may simply be the result of an error in the data), the results for the remaining DMUs become shifted towards lower efficiency values, the efficiency frequency distribution becomes highly asymmetric, and the overall efficiency scale becomes non-linear. Several approaches have been proposed to deal with this effect (Seaver and Triantis 1989; Wilson 1993, 1995; Banker and Gifford 2000) but they largely depend on manual inspection of data, which becomes virtually impossible for large data sets.

In this paper we present a new, computationally intensive approach for automatic detection of outliers, based on a combination of Bootstrap and Jackknife resampling schemes. The essence of this approach is to stochastically reduce the impact of the (usually few) most influential DMUs on the final efficiency scores, using the concept of leverage (Cribari and Zarkos 2003), that is, the effect produced on the outcome of DEA efficiencies of all the other DMUs, when the observed DMU is removed from the data set. This is achieved through a two-phase process. First, the leverage of each DMU on the overall DEA analysis is evaluated, by examining a series of stochastic (bootstrapped) Jackknife replications. Then, Bootstrap resampling (Efron and Tibshirani 1993) is used, taking into account the observed leverage, to yield information on technical efficiency. For this combination of methods we propose the term "Jackstrap" (as it seems somewhat more appropriate then "bootknife").

The paper is organized as follows. In section 2 we first describe the adopted approach of calculating the leverage, and then the procedure of stochastic Bootstrap resampling which takes into consideration the leverage information to reduce de impact of outliers. Section 3 presents numerical results on two large test data sets, and conclusions are drawn in Section 4.

2. DEA and the "Jackstrap" procedure

The "Jackstrap" procedure combines Jackknife (deterministic) and Bootstrap (stochastic) resampling schemes, as follows. First, an algorithm implementing Jackknife is used to extract leverage of all data points, that is, the impact of removal of the observed point on DEA calculations performed on the rest of the data set. The underlying idea is that outliers are expected to show higher leverage then mean, and should be selected with lower probability than the other DMUs. Next, Bootstrap stochastic resampling is implemented, taking into account leverage information. In the rest of this section we describe this approach in some detail.

Let us consider a set of $\boldsymbol{K}$ Decision Making Units, where $\boldsymbol{k}$'th DMU ($\boldsymbol{k=1,\ldots,K}$) uses $\boldsymbol{N}$ nonnegative inputs $\boldsymbol{x^k} = (\boldsymbol{x_{k1}, ..., x_{kN}}) \in \Re_+^N$ to produce $\boldsymbol{M}$ nonnegative outputs $\boldsymbol{y^k} = (\boldsymbol{y_{k1},..., y_{kM}}) \in \Re_+^{\boldsymbol{M}}$. As is well known, the DEA method (Charles, Cooper, and Rhodes 1978) implements linear programming to estimate technical efficiency $\boldsymbol{\theta_k}$ of $\boldsymbol{k}$'th DMU as the minimum positive value that satisfies inequalities

$$\begin{aligned} &\sum_{j=1}^{K} \lambda_{kj} x_{jn} \le \theta_k x_{kn} \qquad && n = 1,\ldots,N \\ &\sum_{j=1}^{K} \lambda_{kj} x_{jm} \ge y_{km} \qquad && m = 1,\ldots,M \end{aligned} \tag{1}$$

where λ_{kj} are adjustable positive coefficients. In fact, this is only one version of DEA (input oriented, with constant returns to scale), as here we are mainly concerned with dealing with the effect of outliers and errors on the overall calculations, the reader is referred to the abundant literature (see e.g. Charles, Cooper, and Rhodes 1978, Banker, Charnes, and Cooper 1984, and Färe, Grosskpof and Lovell 1985) for more details on this and other DEA variants.

Leverage of a single observed DMU may be understood as the quantity that measures the impact of removal of the DMU from the data set, on the efficiency scores of all the other DMUs. Formally, it may be defined as the standard deviation of the efficiency measures before and after the removal. The most straightforward possibility is to perform Jackknife resampling technique as follows. One first applies DEA for each of the DMUs using the unaltered original data set, to obtain the set of efficiencies $\{\theta_k \mid k = 1,\ldots,K\}$. Then, one by one DMU is successively removed, and each time the set of efficiencies $\{\theta^*_{kj} \mid k = 1,\ldots,K\,; k \neq j\}$ is recalculated, where index $j = 1,\ldots,K$ represents the removed DMU. The leverage of j-th DMU may then be defined as standard deviation

$$\ell_j = \sqrt{\frac{\sum_{k=1\,;\,k\neq j}^{K}\left(\theta^*_{kj} - \theta_k\right)^2}{K-1}} \quad . \tag{2}$$

While rather straightforward, this direct approach is extremely computationally intensive, and may turn unfeasible for very large data sets with the available computer resources. More precisely, removing each of the K DMUs from the data set and then performing (K-1) DEA calculations requires solving K(K-1) linear programs, and may become prohibitively computationally expensive for large K. We therefore propose a more efficient stochastic procedure, which combines Boostrap resampling with the above Jackknife scheme as follows:

1. Select randomly a subset of L DMUs (typically 10% of K) and perform the above procedure to obtain subset leverages $\tilde{\ell}_{k1}$, where index k takes on L (randomly selected) values from the set $\{1,\ldots,K\}$.
2. Repeat the above step B times, accumulating the subset leverage information $\tilde{\ell}_{kb}$ for all randomly selected DMUs (for B large enough, each DMU should be selected roughly $n_k \approx BL/K$ times).
3. Calculate mean leverage for each DMU as

$$\tilde{\ell}_k = \frac{\sum_{b=1}^{n_k} \tilde{\ell}_{kb}}{n_k}$$

and the global mean leverage as

$$\tilde{\ell} = \frac{\sum_{k=1}^{K} \tilde{\ell}_k}{K}$$

This completes the first phase of the proposed approach. In the second phase, one can either use the leverage information to detect and simply eliminate outliers from the data set, or one can implement Bootstrap method (Simar and Wilson 1997) to produce confidence intervals and bias information, using leverage information to reduce the probability of selecting the outliers in the stochastic resampling process. In either case, some probability function favoring the low leverage DMUs needs to be adopted, and here we test linear, inverse, exponential and the Heaviside step function. The linear probability distribution is given by

$$P\left(\tilde{\ell}_k\right)=\frac{\ell_{\max}-\tilde{\ell}_k}{\ell_{\max}-\ell_{\min}} \tag{3}$$

where $\ell_{\max}$ and $\ell_{\min}$ are the maximum and the minimum leverage of the set $\left\{\tilde{\ell}_k\ ; k=1,\ldots,K\right\}$, respectively. The probability of retaining/choosing a DMU with leverage $\ell_{\min}$ is therefore unity, while probability of retaining DMUs with leverage $\ell_{\max}$ is zero. The inverse probability distribution may be represented by

$$P\left(\tilde{\ell}_k\right)=\begin{cases} 1 & , \quad \tilde{\ell}_k<\ell_0 \\ \left(\dfrac{\ell_0}{\tilde{\ell}_k}\right)\left(\dfrac{\ell_{\max}-\tilde{\ell}_k}{\ell_{\max}-\ell_0}\right) & , \quad \tilde{\ell}_k\geq\ell_0 \end{cases} \tag{4}$$

where $\ell_0>0$ is now the lower bound for $\tilde{\ell}_k$, that is, an independent parameter (leverage threshold) below which DMU's are retained with probability one. The lower bound ℓ_0 is introduced here primarily to deal with the case of zero leverage values $\tilde{\ell}_k=0$, and may be arbitrarily small. The exponential distribution is given by

$$P\left(\tilde{\ell}_k\right)=\frac{e^{-\tilde{\ell}_k}-e^{-\ell_{\max}}}{e^{-\ell_{\min}}-e^{-\ell_{\max}}} \tag{5}$$

and finally, the Heaviside step function is given by

$$P\left(\tilde{\ell}_k\right)=\begin{cases} 1 & , \quad \tilde{\ell}_k<\tilde{\ell}\log K \\ 0 & , \quad \tilde{\ell}_k\geq\tilde{\ell}\log K \end{cases} \tag{6}$$

Here the threshold level $\tilde{\ell}\log K$ was chosen in order to take into account the sample size, so that for e.g. K=1000 a DMU with leverage greater than three times the global mean is rejected.

The probability functions (3), (4), (5) and (6) are shown schematically on Figure 1, where the scaled leverage variable $\left(\tilde{\ell}_k - \ell_{\max}\right)/\left(\ell_{\max} - \ell_{\min}\right)$ has been used for the abscissa.

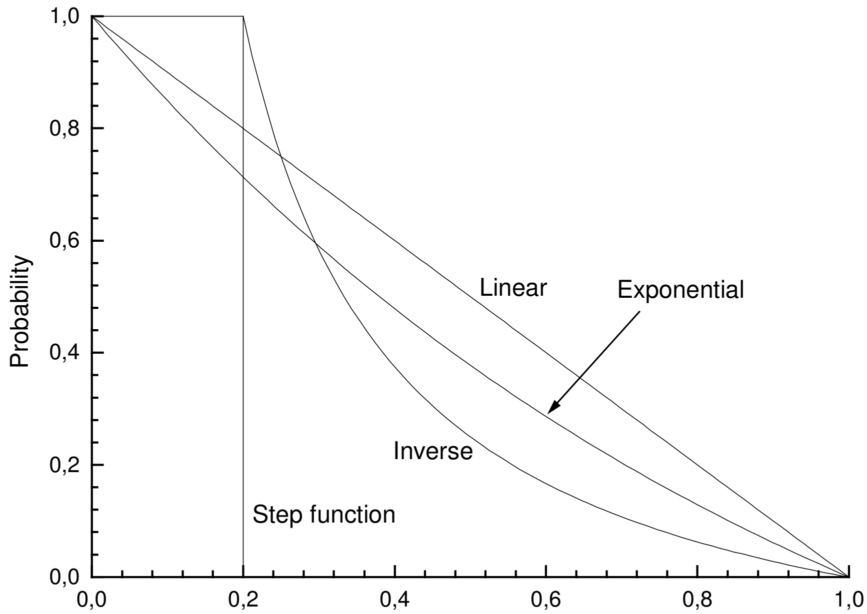
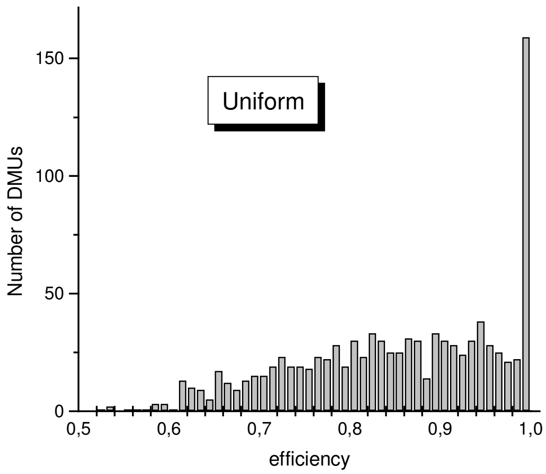


Figure 1. Schematic representation of the probability functions (3), (4), (5) and (6), where the threshold for the Inverse and the Heaviside step function have been set to 0.2.

4. Numerical Results

We have tested the above procedure on two datasets with 1000 DMUs each, with three input and seven output variables. The variables for the first set were generated uniformly in the interval $(0,1)$, and the variables for the second set were generated using the normal distribution with mean $\mu = 1.5$ and standard deviation $\sigma = 1/6$, truncated at points 1.0 and 2.0. Applying DEA on the two datasets we obtain efficiency distributions as shown in Fig.2.

For the Uniform sample, there are 141 efficient DMUs, and the inefficient DMU's are distributed with mean 0.8289 and standard deviation 0.1053, while for the Normal sample there are 69 efficient DMUs, the inefficient having mean 0.8638 with standard deviation 0.0681. The Normal sample displays a more symmetric distribution of efficiencies, and may be considered somewhat more realistic.



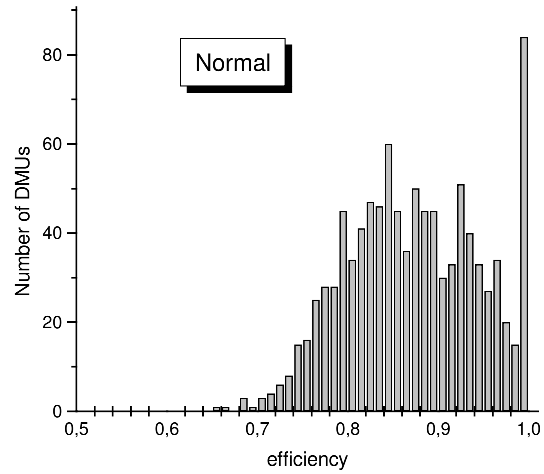
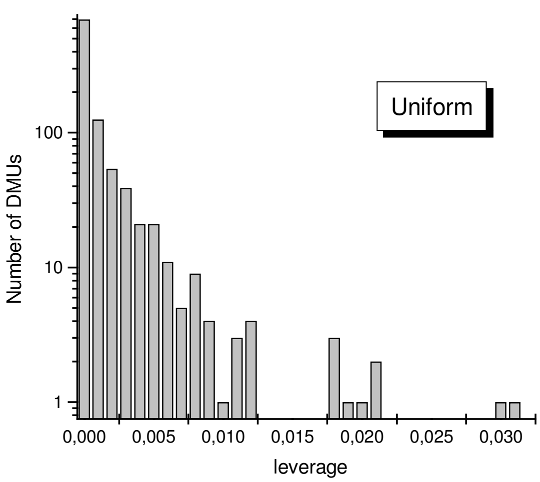


Figure 2. Frequency distributions of efficiency obtained applying DEA, for the two test samples generated using Uniform and Normal distributions, respectively.

Next, we have calculated the leverage for the datasets using the procedure described above, with B=1000 "Jackstrap" passes on subsets of L=100 DMUs (which represents $n_k \approx BL/K = 100$ leverage tests per DMU), and the resulting leverage distributions are shown in Fig.3.



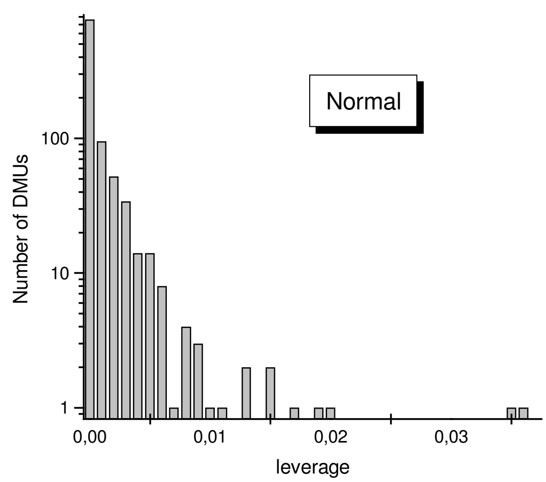


Figure 3. Frequency distributions of leverage obtained through Jackstrap procedure, for the Uniform and Normal test sample, respectively.

From the observed curvature on the semi-logarithmic plot, it follows that DMU frequency decays with leverage faster than exponential, with only two of the generated DMUs (in each dataset) demonstrating leverage of the order 0.03. The mean leverage values are found to be 0.001055 and 0.000777 for the two sets, respectively.

In order to check the effect of the subset size L used in the Jackstrap procedure, we have also performed additional runs for the Uniform sample, for L from 25 to 250, with $B = 100000 / L$ (in order to retain constant ratio $BL / K = 100$ among the individual runs). As may be expected, the absolute leverage values decrease with increasing L, however, the successive sets of mean leverage results $\left\{ \tilde{\ell}_k \; ; k = 1, \dots, K \right\}$ are highly correlated, as may be seen from Table 1.

L	25	50	75	100	125	150	175	200	225	250
25	1,0000									
50	0,9883	1,0000								
75	0,9756	0,9940	1,0000							
100	0,9636	0,9896	0,9962	1,0000						
125	0,9520	0,9816	0,9929	0,9966	1,0000					
150	0,9406	0,9738	0,9889	0,9939	0,9967	1,0000				
175	0,9305	0,9662	0,9836	0,9902	0,9952	0,9975	1,0000			
200	0,9229	0,9611	0,9790	0,9878	0,9928	0,9964	0,9979	1,0000		
225	0,9094	0,9490	0,9710	0,9782	0,9871	0,9922	0,9949	0,9950	1,0000	
250	0,9037	0,9462	0,9669	0,9774	0,9858	0,9902	0,9936	0,9955	0,9956	1,0000

Table 1. Correlation coefficients for successive sets of mean leverage results for the Uniform sample, for subsample size L ranging from 25 to 250.

The dependence of mean leverage on L is shown in Figure 4 for the first 30 DMUs with highest mean leverage. Keeping in mind that this is a stochastic procedure, it is seen that the leverage ranking is generally preserved across the spectrum of different values of L.

Another test was performed in order to check the effect of the number of Jackstrap passes for the Uniform sample, using L=100 and ranging B from 100 to 5000. From the results shown in Figure 5, it is seen that after initial fluctuations the mean leverage attains a stable value that does not change with further increase of B. We now introduce outliers (or “errors”) in the samples, by adding 10 new DMUs, where the first outlier ERR1 has all the outputs higher and all the inputs lower than the DMUs in the original samples, outlier ERR2 has a single output one order of magnitude higher then the samples average, and outlier ERR3 has a single input one order of magnitude lower then average. The rest of the outliers have varying degrees of deviation from average, as can be seen from Table.2.

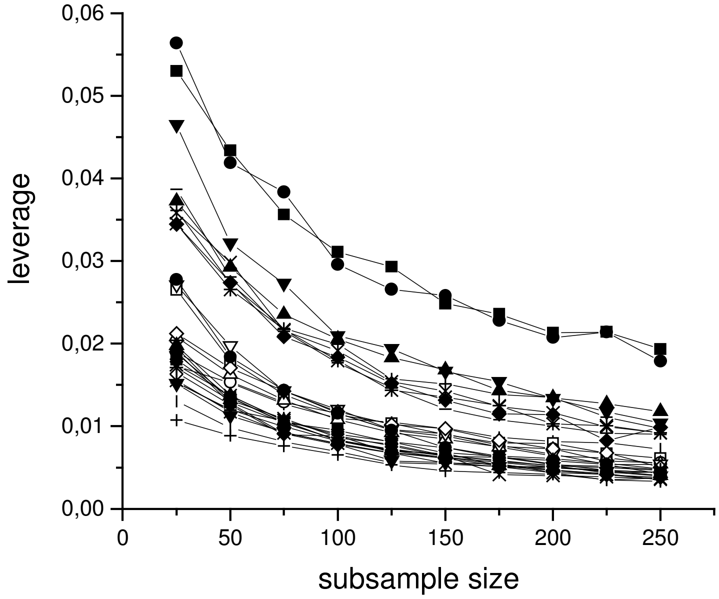


Figure 4. Mean leverage dependence on subsample size L, for the 30 DMUs with highest leverage values in the Uniform sample.

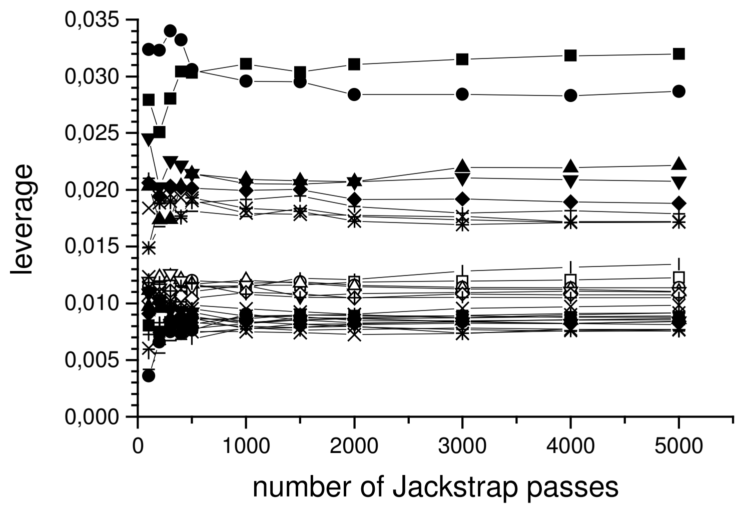
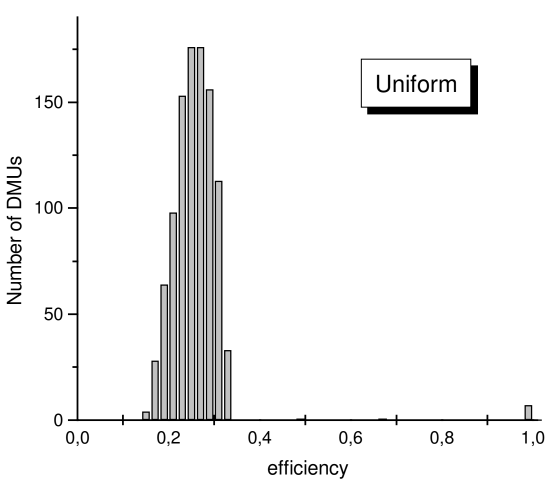


Figure 5. Mean leverage dependence on number of Jackstrap passes B, for the 30 DMUs with highest leverage values in the Uniform sample.

DMU	OUT1	OUT2	OUT 3	OUT 4	OUT 5	OUT 6	OUT 7	INP1	INP2	INP3
ERR1	3,00	3,00	3,00	3,00	3,00	3,00	3,00	0,50	0,50	0,50
ERR2	15,00	1,50	1,50	1,50	1,50	1,50	1,50	1,50	1,50	1,50
ERR3	1,50	1,50	1,50	1,50	1,50	1,50	1,50	0,15	1,50	1,50
ERR4	3,00	3,00	3,00	3,00	3,00	3,00	3,00	1,50	1,50	1,50
ERR5	1,50	1,50	1,50	1,50	1,50	1,50	1,50	0,50	0,50	0,50
ERR6	3,00	1,50	3,00	1,50	3,00	3,00	3,00	1,50	0,50	0,50
ERR7	3,00	3,00	3,00	1,50	1,50	1,50	3,00	0,50	1,50	0,50
ERR8	3,00	3,00	1,50	3,00	1,50	3,00	1,50	1,50	0,50	1,50
ERR9	1,50	1,50	1,50	1,50	1,50	3,00	3,00	0,50	0,50	1,50
ERR10	3,00	1,50	1,50	1,50	1,50	1,50	1,50	0,50	1,50	1,50

Table 2. Outliers added to the test samples, with varying degrees of high output and low input variable values.

Applying DEA on the two datasets with outliers, we obtain efficiency distributions as shown in Fig.6, where it is seen that presence of outliers drastically affects the efficiency distributions, shifting the mean for both samples close to 0.2.



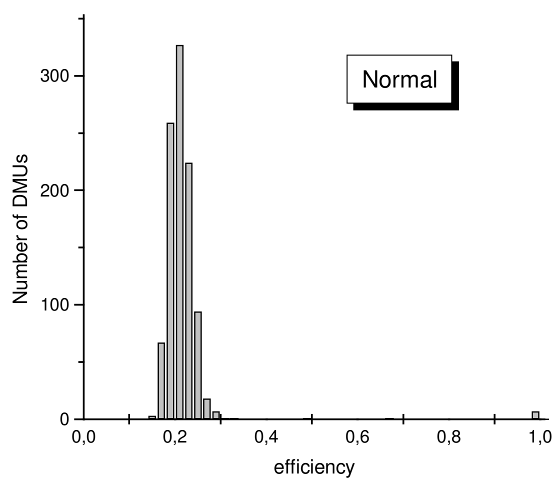


Figure 6. DEA efficiency frequency distributions after adding outliers, for the Uniform and Normal sample, respectively.

The efficiencies obtained for the outliers are the same for both samples as shown in Table 3, showing that the rest of the dataset does not affect these values, while two outliers lie outside of the frontier.

DMU	ERR1	ERR2	ERR3	ERR4	ERR5	ERR6	ERR7	ERR8	ERR9	ERR10
Efficiency	1.0000	1.0000	1.0000	0.3333	1.0000	1.0000	1.0000	1.0000	1.0000	0.6774

Table 3. Outlier efficiencies.

subsample size L=100 with B=1000 Jackstrap passes, the data for the thirty DMUs with highest leverage are shown in Table 4 for both samples.

The second column for each sample in Table 4 lists the obtained leverage values, while columns “Leverage hits” and “Total hits” correspond to the number of times that the removal of the given DMU produced nonzero leverage, and the total number of times that the given DMU was chosen within the Jackstrap procedure, respectively. While all the DMUs have the same chance to be chosen for leverage testing (values in column 4 are similar for all the DMUs), the additionally introduced outliers typically present high number of leverage hits. In comparison with the leverage distribution of the original samples shown in Figure 4, the maximum leverage has risen from 0.03 to 0.5 for both samples, the mean leverage has risen from 0.001055 to 0.002033 while the leverage of the original DMUs has fallen to 0.000687 for the Uniform sample, and for the Normal sample mean leverage has risen from 0.000777 to 0.002004 while mean leverage of the original DMUs has fallen to 0.000491.

Uniform sample				Normal sample			
DMU	Leverage	Leverage hits	Total Hits	DMU	leverage	Leverage hits	Total Hits
ERR1	0,531081	82	88	ERR1	0,553371	82	88
ERR5	0,299649	81	99	ERR5	0,376291	81	99
ERR6	0,114094	76	99	ERR4	0,210663	71	96
ERR4	0,108820	71	96	ERR6	0,112622	76	99
ERR7	0,106558	97	111	ERR7	0,096731	97	111
ERR9	0,073234	95	106	ERR9	0,068987	95	106
ERR8	0,059825	83	95	ERR8	0,059602	83	95
ERR3	0,032835	95	100	ERR3	0,024172	95	100
ERR10	0,022309	63	91	DMU641	0,023936	67	104
DMU846	0,021295	69	102	DMU685	0,023528	77	110
ERR2	0,017747	101	105	ERR10	0,018656	65	91
DMU381	0,016687	83	113	DMU940	0,013378	61	89
DMU52	0,016150	65	92	DMU914	0,012295	57	92
DMU155	0,016110	61	100	ERR2	0,012068	99	105
DMU743	0,013850	86	121	DMU676	0,010943	77	122
DMU333	0,011994	63	91	DMU437	0,010549	55	96
DMU45	0,011726	62	101	DMU966	0,010132	64	102

DMU518	0,011503	65	95	DMU737	0,009635	74	115
DMU868	0,010283	93	132	DMU799	0,007268	67	112
DMU295	0,010198	77	112	DMU663	0,006874	73	112
DMU792	0,009716	63	101	DMU351	0,006847	73	115
DMU10	0,008371	69	99	DMU839	0,006831	63	89
DMU653	0,007816	66	100	DMU837	0,006584	61	100
DMU413	0,007272	60	101	DMU40	0,006251	49	82
DMU789	0,006952	56	93	DMU950	0,005907	61	92
DMU33	0,006675	66	99	DMU135	0,005054	58	110
DMU186	0,006628	62	100	DMU588	0,004730	70	110
DMU318	0,006588	68	105	DMU689	0,004699	54	100
DMU349	0,006526	61	99	DMU279	0,004309	61	99
DMU201	0,006201	62	98	DMU61	0,004293	57	93

Table 4. Thirty DMUs with highest leverage values for the two samples, after adding outliers.

Table 4 shows that the introduced outliers generally have higher leverage than the original DMUs, with the exception of ERR2 which has lower leverage then some of the original DMUs in both cases, and DMU10 in the case of the Uniform sample. From Table 2 it is seen that ERR2 has a single output value one order of magnitude larger than average, while ERR10 has one output three times the average, and one input equal to a third of the average. While they both remain on the top of the leverage list, these deviations are not sufficient to dominate (in terms of leverage) all of the well performing units from the original sets. In contrast to ERR2, it is interesting to note that ERR3, which has a single input one order of magnitude lower than average, maintains higher leverage than all of the original DMUs (although barely). This may be explained by the fact that there are altogether seven output and only three input variables, leading to the conclusion that a single error becomes more influential if a smaller number of variables of the same kind (input or output) is used.

Finally, we have performed Bootstrap on both the datasets with outliers, using leverage information to reduce the probability of selecting the outliers in the stochastic resampling process, according to probability functions (3), (4), (5) and (6). Results for the Uniform and Normal test samples are shown in Figure 7 and 8, respectively.

Comparing the results on Figure 7 and 8 with the original distribution shown on Figure 2 and the distribution after adding the outliers shown on Figure 6, it is seen that in all cases there has been considerable improvement in recovering the original distributions. This is most pronounced in the case of Heaviside step function, where applying expression (6) to find the cutoff leverage values 0.006109 and 0.006022 for the two samples, it follows from Table 4 that the step function identifies (and eliminates) all of the introduced outliers for both samples with a large margin, together with roughly twenty outliers from the original sets.

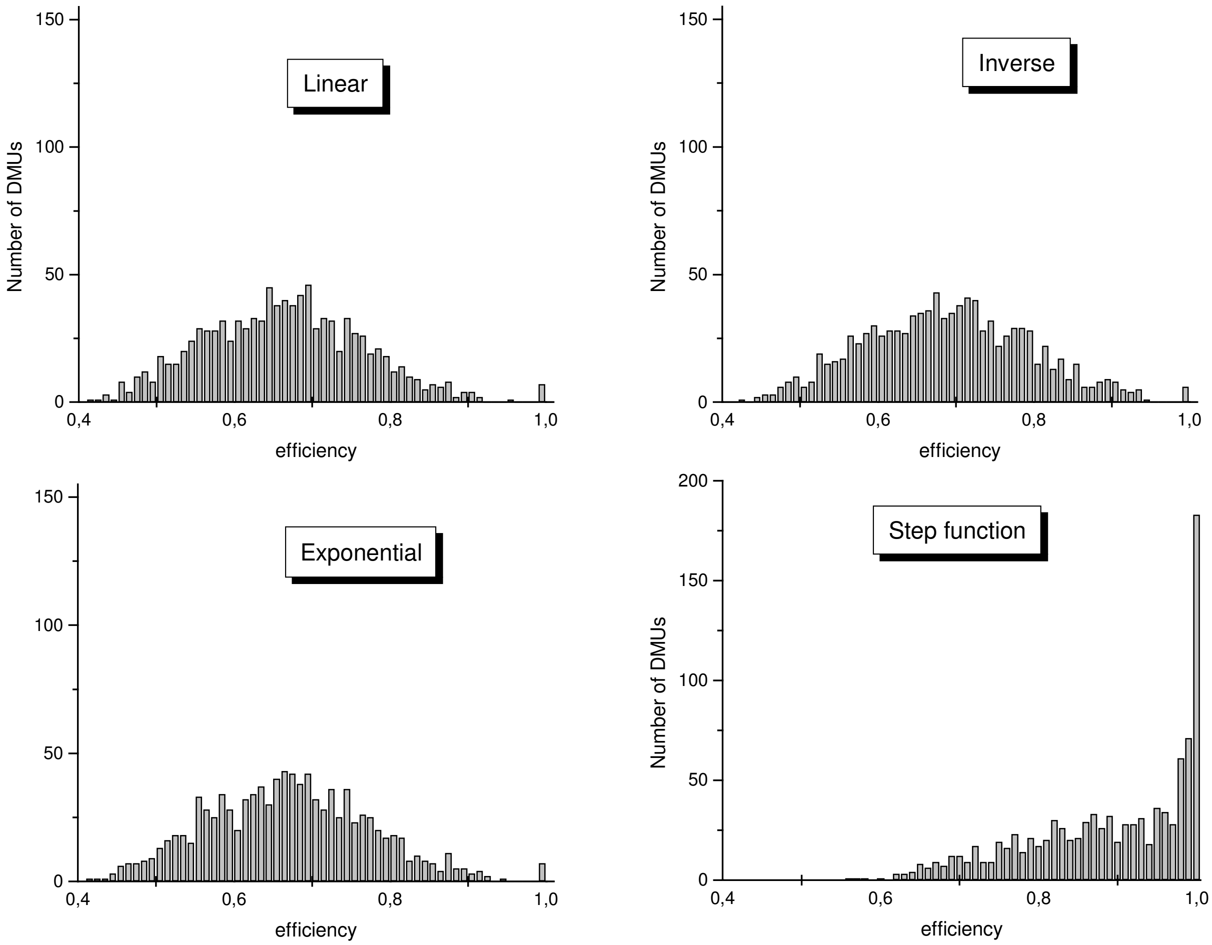


Figure 7. Frequency distributions of efficiency obtained applying Bootstrap with DEA, using probability functions (3), (4), (5) and (6) in the Bootstrap resampling process, for the Uniform test sample.

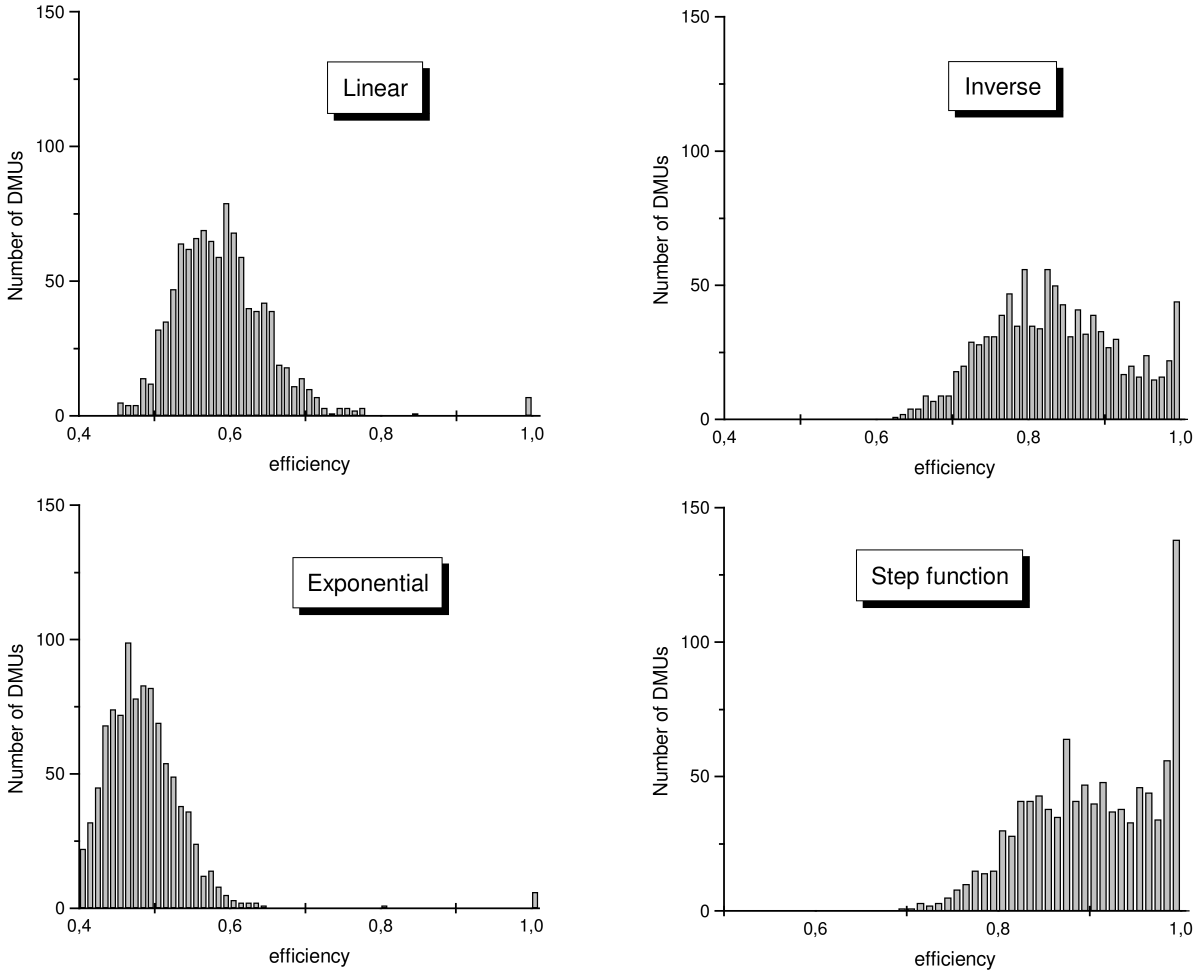


Figure 8. Frequency distributions of efficiency obtained applying Bootstrap with DEA, using probability functions (3), (4), (5) and (6) in the Bootstrap resampling process, for the Normal test sample.

5. Conclusion

In this paper we introduce a new approach for application of DEA on large datasets, that reduces the effect of outliers and/or errors by using a combination of Jackknife and Bootstrap resampling schemes to determine leverage of the DMUs (the impact of removal of the observed point on DEA calculations performed on the rest of the data set). Extensive case study is presented on two random test samples, generated using uniform and normal distributions. The impact of the choice of subsample size and number of passes of the Jackstrap procedure on the leverage results is also investigated. It is also shown how leverage information may be used within Bootstrap resampling scheme to obtain robust results, insensitive to presence of outliers and/or errors.

Moreover, this approach proves to be robust to the presence of outliers and/or errors in the data set, and it has the ability to detect those atypical observations even if they are not on the efficiency frontier. Finally, as it is completely automatic, this method does not require manual inspection of the data set, and is thus particularly suitable for implementation on very large data sets. Indeed, successful application of this method for estimating the DEA technical efficiency for close to five thousand Brazilian municipalities, with nine output and four input variables with huge diversity and a considerable number of errors in the data, is to be published elsewhere.

REFERENCES

Banker, R. D and Gifford, J.L. (1988) "A Relative Efficiency Method for the Evaluation of Public Health Nurse Productivity." Mimeo.

Banker, R. D., Charnes, A and Cooper, W.W. (1984) "Some Models for Estimating Technical and Scale Efficiencies in Data Envelopment Analysis," *Management Science* 30: 1078-1092.

Charles, A., Cooper, W.W. and Rhodes, E. (1978) "Measuring Efficiency of Decision Making Units", *European Journal of Operational Research* 1: 429-44.

Cribari, F and Zarkos, S. "That Voodoo That You Do So Well: Leverage Adjusted Weighted Bootstrap Methods", preprint.

Efron, B. and Tibshirani, R.J.. (1993) "An introduction to Bootstrap", Monographs on Statistics and Applied Probability, Chapman and Hall, New York.

Farrell, M. (1957) "The Measurement of Productive Efficiency." *Journal of the Royal Statistic Society Series A* 120: 253-281.

Färe, R., Grosskpof, S. and Lovell, C. K. (1985) *The Measurement of Efficiency of Production.* Boston: Kluwer-Nijhoff Publishing.

Seaver, B. L. and Triantis, K. P. (1989). "The Implications of Using Messy Data to Estimate Production-frontier-based Technical Efficiency Measures." *Journal of Business and Economic Statistics* 7: 49-59.

Seiford, L.M. (1996) “Data Envelopment Analysis: The evolution of the state-of-the-art (1978-1995)”. Journal of Productivity Analysis 7, 99-138.

Simar, L. and Wilson, P.W. (1997) “Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models”. Managements Science 44, 1, 49-61.

Wilson, P. (1993) “Detecting Influential Observations in Data Envelopment Analysis.” *Journal of Productivity Analysis* 6: 27-45.

Wilson, P. (1995) “Detecting Influential Observations in Deterministic Non-Parametric frontiers models.” *Journal of Business and Economic Statistics* 11: 319-323.