

# SAMPLING SIZE AND EFFICIENCY BIAS IN DATA ENVELOPMENT ANALYSIS

MOHAMMAD R. ALIREZAEI

*University of Calgary, and Teacher Training University*

MURRAY HOWLAND AND CORNELIS VAN DE PANNE

*University of Calgary*

**Abstract.** In Data Envelopment Analysis, when the number of decision making units is small, the number of units of the dominant or efficient set is relatively large and the average efficiency is generally high. The high average efficiency is the result of assuming that the units in the efficient set are 100% efficient. If this assumption is not valid, this results in an overestimation of the efficiencies, which will be larger for a smaller number of units. Samples of various sizes are used to find the related bias in the efficiency estimation. The samples are drawn from a large scale application of DEA to bank branch efficiency. The effects of different assumptions as to returns to scale and the number of inputs and outputs are investigated.

**Keywords:** Data Envelopment Analysis, Efficiency, Branch Banking, Sampling.

## 1. Introduction.

Data Envelopment Analysis (DEA) has become an important tool for the comparison of units in terms of efficiency and has been applied to many fields, see Charnes, Cooper, and Rhodes (1978), Banker, Charnes, and Cooper (1984) and Charnes, e.a. (1995). Its advantages are well known. Any number of inputs and outputs can be included in the comparison and no specific functional form of their relationship is assumed. Constant, variable, increasing and decreasing returns to scale can be accommodated. However, some difficulties related to the method have not been addressed so far.

In DEA the units of the dominant set, for which no combination of other units exists with lower inputs for the same outputs, are assigned efficiencies of 100%, and other units are expressed in terms of this dominant set. But these units are not necessarily efficient, they are merely dominant, which means that no other units were found that were more efficient. If the units of the dominant set are in reality less than 100% efficient, DEA overestimates their efficiency. The same is then true for the other, non-dominated units. This means that DEA efficiency scores overestimate efficiency and are biased. This has been recognized in theoretical work from Farrell (1957) to Banker (1993), but in applied work this is seldom mentioned.

The bias will depend on the relative size of the dominant set, because the smaller the relative size of this set, the larger the likelihood that its units will be 100% efficient. The size of the dominant set depends on many factors. Apart from

Table 1. Efficient DMU's in Bank Branch Efficiency Studies.

Study	DMU's	Number of inputs, outputs	Efficient DMU's	Percentage Efficient DMU's	Average Efficiency
Sherman & Gold	14	3, 4	8	57%	96%
Giokas	17	3, 3	5	29%	87%
Vassiloglou & Giokas	20	4, 4	9	45%	91%
Sherman & Ladino	33	5, 5	10	30%	80%
Parkan	35	6, 6	24	69%	98%
Schaffnit, c.s.	291	5, 6	153	49%	95%
Schaffnit, c.s.	291	5, 9	175	60%	97%
Tulkens, Public Bank	773	1, 7	13	2%	67%
Tulkens, Private Bank	911	1, 7	21	2%	62%
This study, 3I,3O	1282	3, 3	15	1%	50%
This study, 6I,12O	1282	6, 12	254	20%	83%

the distribution of efficiencies of the units, the most important seem to be the total number of units in the analysis, the number of inputs and outputs, and the assumption as to returns to scale. This study is an attempt to shed light on these relations by using sampling from the units of a large scale DEA application.

These units were 1282 branches of a major Canadian bank. The average efficiency found in this study was 50%, which differs from the results found in comparable bank branch studies, such as Sherman and Gold (1985), Parkan (1987), Oral and Yolalan (1990), Vassiloglou and Giokas (1990), Giokas (1991), Tulkens (1993), Sherman and Ladino (1995) and Schaffnit, c.s. (1995). Table 1 gives an overview of the characteristics of these studies as well as the average efficiencies found. Though these studies differ in many respects, there is a general tendency for the average efficiency to go down as the number of Decision Making Units (DMU's) increases.

In order to study the impact of the number of units on DEA efficiency measurement, this paper uses sampling from the 1282 branches for two different configurations of inputs and outputs and for two different assumptions as to returns to scale. The order of discussion is as follows. First some general theoretical background is given, followed by an explanation of the sampling. Then the data and the models of the bank branch study are described. The sampling experiments and their results are given in the next section, while the last section contains conclusions that are drawn from these results.

## 2. Data Envelopment Analysis and Efficiency.

Data Envelopment Analysis provides a measure of the efficiency of a decision making unit (DMU) relative to other such units, producing the same outputs with the same inputs. DEA, which was developed by Charnes and Cooper and Rhodes

(1978), is related to the concept of technical efficiency and can be considered as a generalization of the Farrell (1957) efficiency measure.

Consider a number of comparable units, represented by the index  $k$ , which have a number of inputs with index  $i$  and a number of outputs with index  $j$ . The quantity of input  $i$  for unit  $k$  is then given by  $x_{ki}$  and its quantity of output  $j$  by  $y_{kj}$ . The efficiency of the unit  $k = 0$  relative to all units is then determined by the following linear programming problem:

$$\begin{aligned} &\text{Minimize with respect to } \theta \text{ and all } \lambda_k \text{'s} \\ &g = \theta \\ &\text{subject to} \\ &\sum_k x_{ki} \lambda_k = x_{0i} \theta, \text{ for all } i, \\ &\sum_k y_{kj} \lambda_k = y_{0j}, \text{ for all } j, \\ &\lambda_k = 0, \text{ for all } k. \end{aligned}$$

This problem can be interpreted as that of finding a linear combination of all DMU's producing at least the same outputs as DMU 0 but using at most a fraction  $\theta$  of its inputs, with  $\theta$  to be minimized. For  $\lambda_0 = 1, \theta = 1$ , so that the efficiency has an upper bound of 1.

The formulation given above is input oriented. A similar output oriented formulation is possible, as well as equivalent formulations corresponding to dual linear programming problems.

Consider the following example for five DMU with the same output of 1:

k	DMU	Input 1	Input 2
1	A	5	0.5
2	B	2.5	1
3	C	1	2.5
4	D	0.5	5
5	E	3	3

Figure 1 gives the graphical representation. Points  $A, B, C$ , and  $D$  are situated on the efficiency frontier, while point  $E$  is not efficient as it uses more inputs than  $B$  and  $C$ .

The optimal solution is  $\lambda_2 = \lambda_3 = 0.5, \lambda_1 = \lambda_4 = \lambda_5 = 0$ , and  $\theta = 7/12 = 0.5833$ . This corresponds with point  $F$ , which is a linear combination of  $B$  and  $C$ , producing one unit but using only 0.5833 of the inputs of point  $E$ . Hence the efficiency of  $E$  is 0.5833 or 58.33%.

The formulation given above is the one given in Charnes, Cooper, and Rhodes (1978), which assumes that production functions have constant returns to scale. This is frequently not realistic, because a small unit may be not made comparable to a large one by simply reducing inputs and outputs by some factor. This can be avoided in a formulation given by Banker, Charnes, and Cooper (1984) allowing for

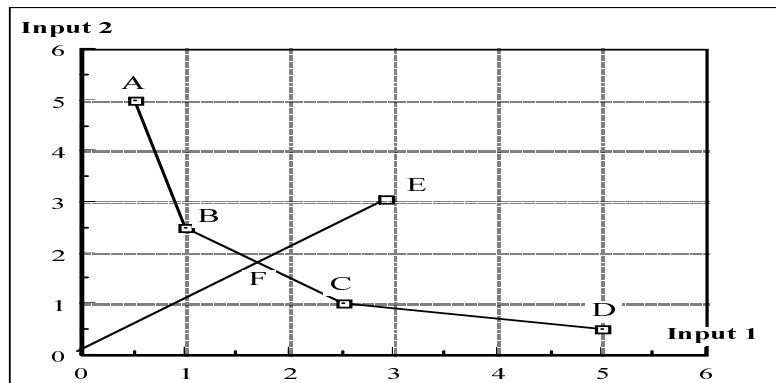


Figure 1. Numerical Example.

increasing and decreasing returns to scale. This is achieved by adding to the linear programming problem the convexity constraint

$$\sum_k \lambda_k = 1.$$

### 3. Data Envelopment Analysis and the Production Frontier.

Data Envelopment Analysis provides estimates of the production frontier. Banker (1993) has shown for the multiple input, one output case and variable returns to scale that under fairly general assumptions the DEA estimator of the production frontier can be interpreted as a Maximum Likelihood Estimator which is biased, but consistent. He indicated that similar results can be obtained for the multiple output case and for constant returns to scale.

If the number of DMU's is small, the dominant set resulting from any application of DEA does not exhaust all possible configurations of inputs and outputs and may contain units that are dominated by units that were not included. An overestimation of efficiency may result, which may be high for cases with few DMU's, but which will tend to zero in probability as the number of DMU's increase.

We may also wish to compare a model with another one in which inputs or outputs are aggregated. The linear programming nature of DEA implies that the latter model cannot have higher efficiencies than the former, and will in general have lower efficiencies. This leads to the general expectation that models with a higher number of inputs and outputs will have higher DEA efficiencies. In the following, models with aggregated and disaggregated inputs and output will be compared.

#### 4. The Sampling Framework.

For sampling from the units of a DEA application, a framework must be indicated. Consider an infinitely large set of decision making units for which data are available, and a DEA model with a given number of inputs and outputs and a returns to scale assumption. It is further assumed that the efficiencies of the units are given by an application of DEA to this infinite set, and that the units in the dominant set are 100% efficient. The DEA efficiencies of other units will vary from 100% downwards and will be considered as the real efficiencies.

If samples of various sizes are taken from this infinite set, and DEA is applied to these samples, efficiencies will be found that are generally different from the real efficiencies determined from the infinite set. Of particular importance are the efficiencies of the dominant set of the sample. If their real efficiencies are not 100%, the sample efficiencies of these units will be biased and overestimated. Since the efficiencies of the other units in the sample are based on those of the dominant set in the sample, they will be biased in the same direction and to a similar extent.

This can be proved as follows. For the input oriented formulation of DEA, the efficiency score compares the required inputs of the sample efficient set with those of the evaluated set. If the sample efficient set contains units that are inefficient in the infinite set, a lower input combination can be found using the efficient units of the infinite set.

We may consider taking a number of samples of a certain size, apply DEA to each of the samples, and analyze the results in terms of the number of units in the dominant set, their efficiencies, and the efficiencies of the other units, and compare these with the real efficiencies. By varying the size of the samples, information is obtained about the overestimation of efficiencies related to the sample size.

Data for an infinite number of decision making units are not available, unless they are simulated, which has its own difficulties. Instead, a large finite number of units for which data exist may be used, from which samples can be taken. In this case, data for 1282 bank branches were available. This number seemed large enough for practical purposes, but the sampling experiments may indicate to what extent this is true. The results of an application of DEA to these data are used as an approximation of the real efficiencies, with which sample efficiencies can be compared.

#### 5. Data and Model

The data originate from a major Canadian bank with a large branch network with 1282 branches. Branch size varies, with the largest branches having a size of 100 times that of the smallest. Figure 2 presents these data in terms one input, (salaries), and one output, (revenues), showing the range of the data.

Different assumptions as to returns to scale can be made (see Charnes, Cooper, and Rhodes (1978), Banker, Charnes and Cooper (1984), and Charnes, c.s. (1995)).

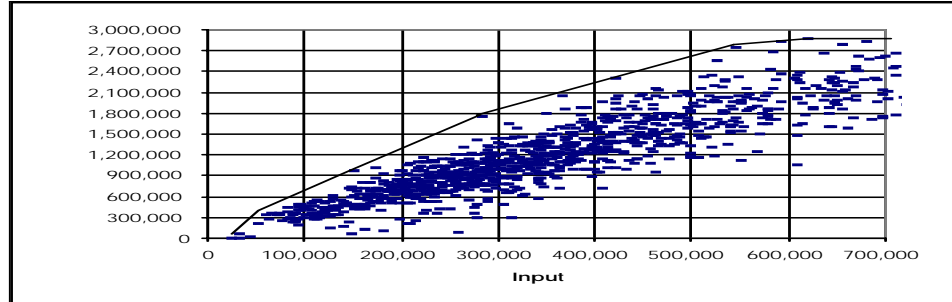


Figure 2. A Small Finite Number of Units in the Dominant Set.

Here the two most important possibilities will be chosen, namely Constant Returns to Scale (CRS) and Variable Returns to Scale (VRS).

The model is further determined by the choice for the inputs and outputs for DEA purposes. Here we shall use the so-called production approach (see Ferrier and Lovell (1990)) or value added approach, see Berg, Førsund, and Jansen (1990), where the volumes of the different kinds of deposits and loans are considered outputs. Inputs are defined in terms of various kinds of costs. Two models were considered, one with 3 inputs and 3 outputs, and one with 6 inputs and 12 outputs.

For the first model, the following inputs and outputs were used:

Inputs.	Outputs.
1. Total Salaries.	1. Deposits.
2. Supply Costs.	2. Retail Loans.
3. Number of Automatic Banking Machines.	3. Commercial Loans.

In the more disaggregated model, the inputs and outputs were:

<u>Inputs.</u>	<u>Outputs.</u>
1. Sales Salaries.	1. Retail Transaction Deposits.
2. Service Salaries.	2. Commercial Transaction Deposits.
3. Support Salaries.	3. Retail Investment Deposits.
4. Other Salaries.	4. Commercial Investment Deposits.
5. Supply Costs.	5. Retail Registered Plan Deposits.
6. Number of Automatic Banking Machines.	6. Retail Demand Loans.
7. Retail Personal Loans.	
8. Retail Other Loans.	
9. Retail Mortgage Loans.	
10. Commercial Loans, Variable Rate.	
11. Commercial Loans, Fixed Rate.	
12. Commercial Mortgage Loans.	

## 6. The Sampling Experiments.

Samples without replacements are taken from the 1282 units representing bank branches of a major Canadian bank. As the number 1282 is close to  $1280 = 10 \times 2^7$ , sample sizes of 640, 320, 160, 80, 40, and 20 are used. The number of samples taken for each sample size is 10. Two choices for inputs and outputs were made, with the first having 3 inputs and 3 outputs (case 3I,3O) and the second 6 inputs and 12 outputs (case 6I,12O). Both constant returns to scale (CRS) and variable returns to scale (VRS) were considered. Altogether DEA was applied  $4 + 10 \times 6 \times 2 \times 2 = 244$  times and 55,528 linear programming problems were solved. Calculations were performed with the General Algebraic Modeling System (GAMS) and spreadsheets.

For any particular sample, the efficiencies obtained when DEA is applied to these units only may be compared with the efficiencies of the same units when evaluated by DEA using all units. In Figure 3 these efficiencies are compared for a sample of 40 for the 3I,3O,CRS case. All points are on or above the 45-degree line, since the sample efficiencies cannot be lower than the full set efficiencies. In the sample evaluation there are 8 units that are 100% efficient, of which only one unit was 100% efficient in the full set evaluation. In the following, the average results for 10 samples of different sample sizes and model specifications are discussed.

### 6.1. Results for the 3I,3O,CRS Case.

First the case with three inputs and three outputs and constant returns to scale (3I,3O,CRS) is considered. For the full set of 1282 units, the dominant set contained 15 units, and the average efficiency was 50.3%. As the relative size of the dominant set is small, 1.17%, the units in this set are likely to be efficient or close to it.

Table 2. Results for the Dominant Set in the 3I,3O,CRS Case.

Sample Size	Average Number in Dominant Set	Percentage in Dominant Set	True Average Efficiency Score
20	8.6	43%	61%
40	10.1	25%	70%
80	14.5	18%	71%
160	16.3	10%	78%
320	16.6	5%	84%
640	17.2	3%	91%
1282	15.0	1%	100%

Table 2 gives an overview of the results for this case in terms of the dominant set. For a sample size of 640, the average number in the dominant set was 17.2. This is 3% of the number of units. For a sample size of 320, this percentage increases to 5%, and it increases to 43% for a sample of 20. The units of the dominant set are the best ones of the units considered, but they may be inefficient if all relevant units are included.

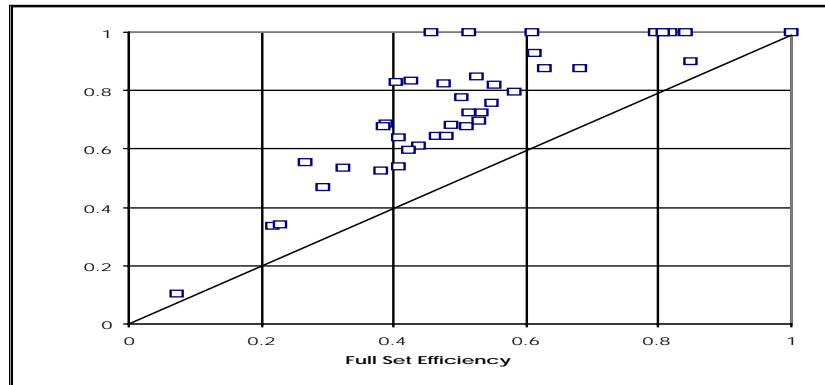


Figure 3. Sample and Full Set Efficiencies.

The “true” average efficiency score of the units in the dominant set, which is by assumption 100% for the sample evaluation, is lower for the evaluation based on 1282 units. This efficiency is given in the fourth column of Table 2. For samples of 320 units, this average efficiency is 84%, which is not very high. This means that samples of this size overestimate the efficiency of the dominant DMU’s by 19%.



These results confirm Banker’s proposition that the DEA estimates are consistent. The average error for the dominant DMU’s decreases from 39% for a sample of 20 to an error of 9% for a sample of 640.

It is possible to obtain one more observation by noting that the average efficiency of a sample of 1 equals the average efficiency of all 1282 units, which is 50.3%.

If the logarithm of the efficiencies of the dominant set are graphed against the logarithm of the sample size, a linear relation is obtained, see Figure 4. This relationship may be estimated using least squares. Since the observation for a sample of 1 is based on 1282 samples, whereas the other are base on 10 samples, a more efficient estimator is obtained by fitting a line through this point. A corresponding least squares estimation results in the following relationship:

$$E = 0.503SS^{0.089} \quad (R^2 = 0.97, t = 32.0).$$

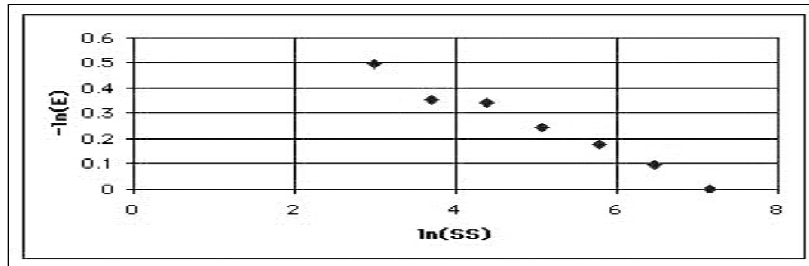


Figure 4. Efficiency of Dominant Set and Sample Size.

This implies that if a sample of size SS units is taken, the units that constitute the dominant set and that are therefore declared 100% efficient, can be expected to have in reality an efficiency of  $0.503SS^{0.089}$ . The other less efficient units of the sample may be given a similar efficiency correction, though this will result in some underestimation, as they are compared with the dominant units of the sample instead of the dominant units of the population.

A maximum percentage could be given for the number of units in the dominant set. If this percentage were set, somewhat arbitrarily, at 5%, then, according to Table 2, the sample size should be at least 320. This average efficiency of the dominant set at 5% depends, of course, on the real efficiencies of the units. Table 1 indicates that in this case the average efficiency of the maximum dominant set is 84%, which implies a bias of 19%.

Table 3. Results for the Dominant Set in the 3I,3O,VRS Case.

Sample Size	Average Number in Dominant Set	Percentage in Dominant Set	True Average Efficiency Score
20	10.9	55%	62%
40	15.6	39%	67%
80	22.8	29%	69%
160	25.5	16%	77%
320	31.5	10%	83%
640	31.0	5%	93%
1282	33.0	3%	100%

Consider now the number of units in the dominant set. For the 10 samples of size 640, the average was 17.2, the lowest number found was 9, and the highest 24, with the average minus and plus twice the standard deviation at 14.3 and 20.0. Even though the distribution cannot be normal, since the number of units is discrete, this gives a good idea of the variability.

For the full set, the number of units in the dominant set was 15, which can be considered as the result for one sample of 1282 units. For increasing sample sizes, the average number in the dominant set increases, but if upper and lower limits of twice the standard deviation are taken into account, it seems that for sample sizes above 100 the average does not change significantly. An asymptotic value between 16 and 18 may be assumed from the results given in Table 2.

## 6.2. Results for the 3I,3O,VRS Case.

Let us now consider the results for variable returns to scale. For the full set of 1282 units, the number of units in the dominant set was 33, and the average efficiency score was 54%, which is significantly higher than the 50% found for constant returns to scale. This is probably related to the increased size of the dominant set, which is now 3% of the total number of units. Note that in the CRS case, a 3% dominant set for 640 units has an average full set efficiency of 91%, which implies an overestimation of 11% if a sample of 640 is used. Table 3 gives the results for the dominant sets. From the second and the last two columns of this table, it may be concluded that the average number of units of the dominant set has stabilized for sample sizes above 320 to about 31, with 33 as just the sample value for 1282 units.

Also here an increased sample size leads to smaller errors in the estimation of the production frontier. The efficiency of the dominant set as a function of sample size can be estimated as indicated in the CRS case. The following result is found:

$$E = 0.54SS^{0.075} \quad (R^2 = 0.91, t = 16.3).$$

Table 4. Results for the Dominant Set in the 6I,12O,CRS Case.

Sample Size	Average Number in Dominant Set	Percentage in Dominant Set	True Average Efficiency Score
20	18.7	94%	76%
40	35.3	88%	79%
80	58.1	73%	81%
160	90.7	57%	85%
320	140.1	44%	88%
640	190.4	30%	92%
1282	254	20%	100%

If the 5% rule is used, a sample of 640 units is needed to reduce the dominant set to 5%. The average full set efficiency for such samples is 93%, which implies that the sample efficiency overestimates the full set efficiency by 7.5%. To this could be added the bias resulting from a larger dominant set for 1282 units. The assumption of variable returns to scale approximately doubles the size of the sample required to have the same accuracy as in constant returns to scale.

The size of the dominant set, which is 33 for the full set is about double that for CRS. It does not appreciably change for samples of 320 and higher. A larger dominant set must lead to higher average efficiency than in the CRS case. In accordance with this, we find an average efficiency of 54% versus 50.3% in the CRS case.

### 6.3. Results for the 6I,12O Cases.

If inputs and outputs are disaggregated into separate parts, the number of inputs and outputs increases. This has an impact on the results of DEA. Here we consider the case in which the three inputs are increased to six by splitting up the salaries in various groups, and instead of three outputs, twelve are taken by dividing up deposits, retail loans, and commercial loans according to the type. Table 4 gives the results for the constant returns case.

The average number in the dominant set now increases over the entire range of the sample size. If it has a finite asymptotic value, it is probably much larger than 1282. In practical cases such large samples can almost never be obtained. The average efficiency, which is 79%, is much higher than in the 3I,3O case, which can be explained by the substantial bias induced by the large relative size of the dominant set.

The double logarithmic relation indicated earlier gives in this case the result:

$$E = 0.79SS^{0.005} \quad (R^2 = 0.51, t = 4.1),$$

Table 5. Results for the Dominant Set in the 6I,12O,VRS Case.

Sample Size	Average Number in Dominant Set	Percentage in Dominant Set	True Average Efficiency Score
20	19.0	95%	85%
40	37.5	94%	86%
80	66.1	83%	88%
160	112.4	70%	91%
320	180.0	56%	94%
640	271.1	42%	98%
1282	384.0	30%	100%

which gives a fit less satisfactory than before. The main reason for this is that the average efficiency for the 1282 set of 79% is somewhat out of line with the other observations. This is probably related to the fact that the dominant set for 1282 is far from 100% efficiency.

The percentage in the dominant set is even for 1282 units equal to 20%. Only for samples larger than a few hundred units is the dominant set smaller than 50%. It is obvious that the results for this case are not realistic in the sense that the efficiency scores do not reflect real efficiency.

For the corresponding variable returns to scale case, the results are given in Table 5. For the full set of units, the number of units in the dominant set is now 384, which is 30% of the total. The average efficiency for the full set of units is 83%, which must include a large bias caused by the large relative size of the dominant set.

The relation between efficiency of the dominant set and sample size is now estimated as:

$$E = 0.83SS^{0.021} \quad (R^2 = 0.79, t = 9.3).$$

The percentage in the dominant set seems to go below 50% for about 500 units. It could be that also here a doubling in the number of units is required to obtain the same accuracy as in the constant returns case. There seems to be no realistic number of units that will reduce the dominant set to 5% or less.

## 7. Conclusions.

In applications of Data Envelopment Analysis it is implicitly assumed that the units of the dominant set are 100% efficient. This assumption leads to a biased efficiency evaluation, which tends to be larger when the number of units is smaller.

Sampling from a large number of units from the data of a bank branch study made it possible to analyze the impact of the number of units on the efficiency scores. This was done by studying the absolute and relative size of the dominant

set for varying numbers of units, for different returns to scale assumptions, and varying numbers of inputs and outputs.

It was found that for 3 inputs and 3 outputs and constant returns to scale, a reasonably accurate estimation of efficiency was possible if the number of units was at least a few hundred. For the corresponding variable returns to scale model, this number should be roughly doubled. For the 6 inputs, 12 outputs cases, the relative size of the dominant set is too large for any realistic number of units to yield reasonable accuracy in efficiency measurement.

The results obtained are valid for the bank branch data and model. This does not preclude that in other cases different results will be found. However, it must be expected that in most efficiency studies, the data will have a similar dispersion over the input and output spaces, which will give similar results.

As most DEA studies have at least three inputs and three outputs and less than 100 units, their efficiency scores must be severely biased in an upward direction.

However, this does not make these DEA results meaningless, as the scores of inefficient units may be interpreted as relative to the dominant set. Furthermore, these results may be used to propose improvements by noting the input or output combinations corresponding to the optimal DEA solution. But the caveat should be given that, because of the small number of units, many better solutions may go undetected.

## References

1. M.R. Alirezaee, M. Howland, and C. van de Panne (1995), *A Large-Scale Study of Bank Branch Efficiency*. Paper presented at the 37-th National Conference of the Canadian Operational Research Society, May 23 – 25, 1995 in Calgary.
2. R.D. Banker (1993), *Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation*. Management Science Vol. 39, No. 10, pp. 1265 – 1273.
3. R.D. Banker, A. Charnes, and W.W. Cooper (1984), *Some Methods for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis*. Management Science Vol. 30 (1984), No. 9, pp. 1078 – 1092.
4. R.D. Banker, A. Charnes, W.W. Cooper, and A. Maindiratta (1988), *A Comparison of DEA and Translog Estimates of Production Frontiers Using Simulated Observations from a known Technology*. In A. Dogramaci and R. Färe, Applications of Modern Production Theory. Boston, Kluwer Academic Publishers.
5. S.A. Berg, F.R. Førsund, and E.S. Jansen, *Technical Efficiency of Norwegian Banks: The Non-Parametric Approach to Efficiency Measurement*. The Journal of Productivity Analysis, Vol. 2 (1991), pp. 127 – 142.
6. S.A. Berg, F.R. Førsund, I. Hjalmarsson, and M. Suominen, *Banking Efficiency in the Nordic Countries*. Journal of Banking and Finance Vol. 17 (1993), pp. 371 – 388.
7. A. Charnes, W.W. Cooper, A.Y. Lewin, and L.M. Seiford (ed) (1995), *Data Envelopment Analysis: Theory, Methodology, and Application*. Boston, Kluwer Academic Publishers.
8. A. Charnes, W.W. Cooper, and E. Rhodes (1978), *Measuring the Efficiency of Decision Making Units*. European Journal of Operational Research. Vol. 2 (1978), pp. 429 – 444.
9. M.J. Farrell (1957), *The Measurement of Productive Efficiency*. Journal of the Royal Statistical Society, Series A, pp. 253 – 290.
10. G.D. Ferrier and C.A.K. Lovell (1990), *Measuring Cost Efficiency in Banking: Econometric and Linear Programming Evidence*. Journal of Econometrics, Vol. 46 (1990), pp. 229 – 245.

11. D. Giokas (1991), *Bank Branch Operating Efficiency: A Comparative Application of DEA and the Loglinear Model*. OMEGA Vol. 19 (1991), No. 8, pp. 549 – 567.
12. H. Oral and R. Yolalan (1990), *An Empirical Study on Measuring Operating Efficiency and Profitability of Bank Branches*. European Journal of Operational Research Vol. 46 (1990) pp. 282 – 294.
13. C. Parkan (1987), *Measuring the Efficiency of Service Operations: An Application to Bank Branches*. Engineering Costs and Production Economics, 12 (1987), pp. 237 – 242.
14. H.D. Sherman and F. Gold (1985), *Bank Branch Operating Efficiency: Evaluation with Data Envelopment Analysis*. Journal of Banking and Finance Vol. 9 (1985), pp. 297 – 315.
15. H.D. Sherman and G. Ladino (1995), *Managing Bank Productivity Using Data Envelopment Analysis*. Interfaces, Vol 25(1995), No. 2, pp. 60 – 73.
16. M. Vassiloglou and D. Giokas (1990), *A study of the Relative Efficiency of Bank Branches: An Application of Data Envelopment Analysis*. Journal of the Operational Research Society, Vol. 41 (1990) No. 7, pp. 591 – 597.