IDENTIFYING HIGH PERFORMANCE CONSTRUCTION PROJECTS

Ingunn Myrtveit[1] and Erik Stensrud[2]

[1]*Department of Accounting, Auditing and Law, Norwegian School of Management, Pb 580, 1301 Sandvika,Norway*
[2]*Department of Computer Science, Buskerud University College, Pb. 235, 3601 Kongsberg, Norway*

Learning from high performance projects is crucial for construction improvement. Therefore, we need to identify outstanding projects or role models. A minimum prerequisite for identifying such projects is the ability to measure the performance. Unfortunately, two issues complicate the measuring task: i) diseconomies or economies of scale and ii) multidimensional inputs and outputs. We propose to use Data Envelopment Analysis (DEA) to measure the productivity of building projects. DEA fulfils the two requirements stated above, and to our knowledge, it is the only method complying with these two vital requirements. The presentation emphasises the strengths as well as the limitations of DEA, comparing it with regression analysis. The results from this empirical study of 58 projects extracted from a database in a large Norwegian construction company suggest that there is a 50% potential for productivity improvement by learning from and copying the role models that were identified. Also, we discuss economies and diseconomies of scale in construction projects. We recommend DEA as an appropriate method for identifying role models and for benchmarking projects. Used together with methods for hypothesis testing, DEA is a useful technique for assessing the effect of alleged process improvements.

Key Words: data envelopment analysis, productivity analysis, returns to scale, benchmarking.

INTRODUCTION

The Construction Users Roundtable (CURT) looked at specific problem areas in the construction industry. One problem area was productivity (CURT, 1982). This study found that there is no common definition of construction productivity. Even when definitions are consistent, approaches to measuring input and output vary so greatly that valid comparisons between projects are almost impossible. Furthermore, there is a need for better measurement approaches that apply more specifically to the work at the task level. The report recognized the importance of productivity measurement for determining trends and levels of productivity and for evaluating corrective actions.

Learning from high performance projects is crucial for construction process improvement. Therefore, we need to identify outstanding projects that may serve as role models. A minimum prerequisite for identifying these best practice projects is the ability to *measure* the performance. If you cannot measure it, you cannot possibly know which projects are best, and you cannot know whether you have improved. Also, if you are able to identify the best projects, they may serve as role models guiding you on *how* to improve. For practitioners, identifying and studying the best

[1] ingunn.myrtveit@bi.no

Myrtveit, I and Stensrud, E (2005) Identifying high performance construction projects. *In:* Khosrowshahi, F (Ed.), *21st Annual ARCOM Conference*, 7-9 September 2005, SOAS, University of London. Association of Researchers in Construction Management, Vol. 1, 149-58.

practice projects is an invaluable source of learning. Last, but not least, by measuring project performance, you create incentives that likely will yield higher performance.

In addition to identifying the best practice projects, several stakeholders are interested in the related problem of benchmarking the projects. (In this context, benchmarking means to measure the project performance against some established performance standard, or alternatively, against an observed best practice frontier.)

It is not trivial to correctly identify the outstanding, best performing construction projects. First, we need to establish criteria for what we actually mean by qualitative words like "outstanding", "high performance", "best", and so on, and then we must find appropriate quantifiable measures.

Next, it is vital that the comparisons of individual construction projects deal correctly with variable returns to scale and multivariate data because it is likely that construction projects exhibit variable returns to scale, in general. Finally, the input and output from construction projects could be multivariate.

In this paper, we measure the *productivity* and use it as a performance indicator. In other words, we use the productivity as the criterion to judge construction projects as "high performance" or "best". We use the non-parametric data envelopment analysis (DEA) on a dataset of 58 construction projects from a single firm. The results indicate a 50% potential for improvement.

MEASURING PRODUCTIVITY

Productivity (P) is defined as output (y) over input (x):

$$P = \frac{y}{x} \qquad (1)$$

In the building industry, a simple productivity measure would be the size of the building (m2) as output and labour (work hours) as input. A multivariate input could include additional variables such as material, equipment, energy, and capital. For productivity comparisons to be meaningful, the observations must be reasonably homogeneous, for example with respect to the quality and standard of the building. Comparing the number of square meters built per labour hour on Sagrada Familia and Snoopy’s dog shack does not give much insight from which to learn and generalise.

There is, however, one serious drawback with the productivity model in equation (1). The productivity model (1) assumes constant returns to scale (CRS). In other words, CRS assumes a linear relationship between input and output. This assumption is inconsistent with common belief in the industry. Many believe there are variable returns to scale (VRS). That is, they assume a non-linear relationship between input and output. If our assumptions are incorrect, we may be misled to draw invalid conclusions. We propose to use data envelopment analysis (DEA) because it can handle both CRS and VRS technology. If, in fact, the building industry exhibits VRS, DEA ensures that we compare large projects with large projects and small projects with small projects.

DATA ENVELOPMENT ANALYSIS (DEA)

The initial publication on DEA is credited to Charnes, Cooper and Rhodes (1978) handling CRS, only. Afriat (1972) laid the foundations for VRS, which later have been enhanced by several authors including Banker, Charnes and Cooper (1984) and Førsund and Hjalmarson (1979).

When performing DEA, the first step is to decide whether to use a CRS or a VRS model since DEA gives you the choice. For construction projects in general it is a topic of discussion. Next, we must decide to use an *input reducing* efficiency or alternatively an *output increasing* efficiency measure. (Using DEA terminology, we use the term *efficiency* instead of the term *productivity*. In the paper, they are used as synonyms.) These two measures are illustrated in Figure 1 for project C where AB/AC and EC/ED are the input decreasing and output increasing efficiencies, respectively. Both are reasonable approaches in the context of construction. We can either measure how much less input that could have been used to produce the same amount of output, or alternatively, we can measure how much more output that could have been produced with the same amount of input. In this paper we use the input reducing efficiency measure because for building construction projects usually the building size is fixed, and the objective would then be to investigate if the same building could be constructed with less labour, not the other way around.

Using project C as example, we attempt to find the *minimal* effort required to produce the same amount of output as C produces. That is, we ask how much effort it would take for a best practice project to produce just as much output as C. This minimal effort is the effort at the point B, which is a linear combination of the two frontier projects 21 and 22. These latter are termed *reference projects*. Thus, the idea is to move horizontally from C and towards the left until we hit the line segment at B. This is a minimisation problem, which can be solved using linear programming.

The formal problem is to minimise the objective function:

$$E_i = \min \theta_i \tag{2}$$

subject to the constraints:

$$\sum_j \lambda_{ij} Y_{kj} \geq Y_{ki}, \forall k \tag{2.1}$$

$$\theta_i X_{mi} \geq \sum_j \lambda_{ij} X_{mj}, \forall m \tag{2.2}$$

$$\sum_j \lambda_{ij} = 1 \tag{2.3}$$

$$\lambda_{ij} \geq 0, \forall j \tag{2.4}$$

The constraint in (2.3) is the VRS constraint, and furthermore:

E_i - is the efficiency score for observation i
θ_i – is the efficiency score variable to be determined for observation i
λ_i – are the weights to be determined for observation i
X_{mi}, Y_{ki} – are inputs and outputs of observation i - is the current observation
j - is all the other observations with which observation i is compared
m - is the number of inputs, in our case total cost, only
k - is the number of outputs, in our case m2, only.

RELATED WORK

Efficiency studies at project or firm level in the Scandinavian construction industry are very rare. There are, to our knowledge, only three Scandinavian studies. Jonsson (1996) investigated Swedish construction productivity at the project level. Albriktsen and Førsund (1991) investigated the efficiency of Norwegian construction firms, and Edvardsen (2004) investigated a cross section data set of Norwegian construction firms. The two Norwegian studies are at firm level. There are large organizational and technological differences between construction firms, indicating that an analysis at the project level would ensure a more homogenous dataset.

DATA

In this preliminary analysis, we gathered n=58 observations with two variables, building size (m2) and building cost (in the Norwegian currency: kr), at the project level from the Norwegian National Construction company, Statsbygg. Descriptive statistics are given in Table 1. Statsbygg acts on behalf of the Norwegian government as manager and advisor in construction and property affairs. Statsbygg offers governmental organisations premises suited to their needs, either in new or existing buildings. The present data set consists of all completed new projects from1998 to 2004 managed by Statsbygg. We believe that the data set is reasonably homogeneous. All observations are new public buildings. The majority are schools, but they also include prisons and other governmental buildings where there is reason to assume a similar quality.

The cost variable is the troublesome variable. Ideally, we should measure workhours, but we discovered it was more problematic than anticipated, for two reasons. Firstly, building projects typically involve a hierarchy of subcontractors, making it difficult to gather total work-hours. Secondly, many contracts are fixed price, and thus, there is no need or incentive to report work hours. The cost variable therefore is the contracted price and includes work hours as well as material costs, rent of equipment, etc. However, Statsbygg estimates that labour constitutes around 30% of total cost. If this figure is reasonably constant, it therefore still makes some sense to measure the productivity as m2/kr. Furthermore, for this preliminary analysis we have not adjusted the costs for inflation.

Table 1: Descriptive statistics of project data

	N	Minimum	Maximum	Mean	Std. Deviation
M2	58	523,00	132380,00	8794,2241	20034,07625
KR	58	10500,00	5015000,00	222205,3621	686196,78569

There are several ongoing initiatives in Norway to improve the data quality. Statsbygg is in the process of gathering more detailed data for later analysis. Also, there is an ongoing effort at Byggforsk (The Norwegian Building Research Center) to gather detailed data, also on labour, at the project level at the Norwegian construction industry.

RESULTS

We report two different productivity measures, DEA CRS efficiency (E_{CRS}) and DEA VRS efficiency (E_{VRS}). From Table 2, we observe that the mean efficiency is around 50-60%, depending on whether we apply DEA CRS or VRS, and that the least productive projects are around 25% as productive as the most efficient projects. The most efficient project using DEA CRS is project 8. This is the smallest building project (1000m2).

Table 2: Descriptive statistics of the Efficiency indicators

	N	Min	Max	Mean
E-CRS	58	0,26	1	0,53
E-VRS	58	0,28	1	0,65

When we examine the scatter plot (Figure 1) we observe that the relationship between m2 and kr is rather linear, suggesting CRS, except probably slightly decreasing return to scale (VRS) for the largest projects. The hypothesis of VRS is supported by the DEA VRS numbers. Eight projects are fully efficient (E_{VRS}=1). These eight projects constitute the VRS frontier, but the average efficiency has only increased to 0.65. Also, there are too few large observations to be confident on any conclusion on CRS vs. VRS.

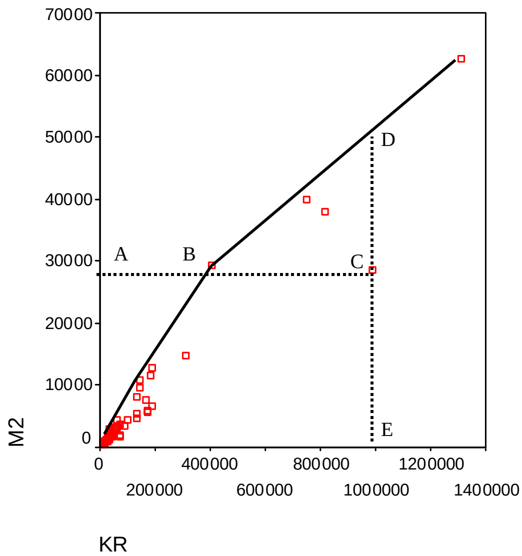

Figure 1: Scatter plot of all projects except the largest

Sensitivity analysis of the frontier

DEA identifies best practice rather than the average productivity. This makes the technique very sensitive to extreme observations. It is, therefore, necessary to do a sensitivity analysis of outliers. A simple method is the so-called superefficiency (Andersen and Petersen, 1993). Superefficiency means to leave out one frontier project at a time and calculate its efficiency relative to the new frontier. Thus its efficiency will exceed 1. The superefficiency analysis of both models suggests that the frontier is fairly stable and not very sensitive to removal of single frontier observations with the exception of the largest project, which is a true outlier. For the CRS model the one most productive project had a superefficiency of 1.08, whereas for the VRS model all the 6 projects one the front obtained values around 1.3. The average efficiency remained similar.

Another method is the analysis of reference units (Torgersen et al. 1996) that identifies the most influential reference units, i.e. those frontier observations, or *peers*, that are referenced most by the inefficient observations. This has a double purpose. First, it may be used to assess the robustness of the frontier (an efficient project that is not referenced at all must be in an area with few observations). Second, it may be used to identify the most worthy role models (by distinguishing the efficient projects that few or no projects reference from those efficient projects that many projects reference). The projects that are referenced most are more likely to be appropriate role models. One method to quantify the *degree of influence* of an efficient project is by computing the *peer index* (Torgersen et al. 1996). The larger the data set and the number of reference units, the more helpful this technique is as part of a sensitivity analysis. The peer index, ρ, is defined as follows.

$$\rho_j^m = \frac{\sum \lambda_{ij}\left(x_{mi} - x_{mi}^p\right)}{\sum \left(x_{mi} - x_{mi}^p\right)} \qquad (3)$$

where in our case with no slack, we have:

$$x_{mi}^p = x_{mi} E_{mi} \qquad (4)$$

$\rho^m{}_j$ – is the peer index for reference unit j and input m
λ_{ij} – is the determined weight for observation i with respect to reference unit j
X_{mi} – is the input m of observation i
$X_{mi}{}^p$ – is the *potential* input m of observation i, had it been efficient, i.e. on the DEA frontier
j - is the number of reference units
m - is the number of inputs
k - is the number of outputs

It is beyond the scope of this paper to discuss the technicalities of the general peer index formula. For a full account, see Torgersen et al (1996). One project (33) is not a reference unit for any project. This is the largest project and an outlier. Two projects are very important reference units, projects 7 and 23. These two projects are reference units for more than 80% of the projects in the dataset. They are not extreme outliers as superefficiency numbers where 1.2 and 1.3 respectively. These two projects may need to be examined more closely by Statsbygg, probably potential role models.

DISCUSSION

There are a number of assumptions underlying performance and productivity models like DEA and univariate CRS models. Most of the assumptions are general and apply to any performance and productivity model. Only a few assumptions are particular to DEA. We find it important to differentiate between general assumptions underlying all kinds of productivity models and assumptions that are specific to DEA. If one does not differentiate between the assumptions underlying all productivity measurements and the assumptions specific to DEA, one runs the risk of unjustly criticising DEA for making unrealistic assumptions and hence reject the use of DEA on false grounds.

Performance and productivity
Ideally, performance assessments should include productivity indicators as well as quality indicators, and it should also take other external factors into account such as schedule constraints. The pragmatic answer to this problem is to simplify the measuring task by measuring the productivity, only. To compare observations, this requires that the data are reasonably homogenous with respect to quality. To measure worker productivity, we need labour hours. This data is difficult to gather in the construction industry because of the typical project structure with a hierarchy of subcontractors and fixed price contracts.

Productivity measurements of individuals

When we compare the productivity of individual projects, rather then the average productivity of a firm or of the whole building industry, the data must be non-stochastic. DEA is unjustly criticised because it is a deterministic model as opposed to a stochastic model like regression analysis that allow for random errors. However, assume that project A delivered 1 m2/hour and project B delivered 2 m2/hour. If one infers that B is twice as productive as A, one has implicitly acknowledged a deterministic model, free of errors.

Regression analysis, on the other hand, assumes a stochastic error term. This is fine when one is interested in the average tendency. However, if you want to compare the productivity of projects A and B, the error term must be discarded.

DEA specific assumptions

Multivariate data. DEA does in fact not make many specific assumptions other than an assumption about how to handle multivariate cases. DEA proposes a method of obtaining a *single* productivity score for multivariate cases. The alternative would be to use a series of simple y/x ratios. DEA offers a solution to this problem by creating a single efficiency or productivity score. DEA makes one assumption in doing this. It assumes that all dimensions have equal weight in normalised space.

Other considerations of DEA

Multivariate DEA VRS. In the multivariate DEA *VRS* case, we require that the number of observations is much larger than the number of variables and that there are not too many specialised units. A considerable number of observations are characterised as efficient unless the sum of the number of inputs and outputs is small relative to the number of observations. Specialised units will typically be on the frontier.

Distribution of efficiency measures. The output from DEA, the efficiency measures, do not have a normal distribution that lends itself to simple statistical analysis since the distribution is truncated at 1. There are, however, more advanced techniques that may be used such as e.g. Tobit regression analysis (Tobin, 1958).

Weighting of the dimensions. As with most other multidimensional measures, DEA does not solve the problem of weighting the dimensions. All the dimensions are normalised, i.e. they have equal weights. As opposed to this crude approach, a statistical technique like, say, regression analysis is more sophisticated in that it provides a technique to weight the dimensions (by providing the sample regression coefficients).

On proving VRS. DEA VRS handles variable returns to scale but one should be cautious in using it to prove the data are VRS. DEA handles CRS as well as VRS cases. However, if there are large random errors, one may wrongly conclude that the data are VRS when in fact they are CRS (Myrtveit and Stensrud, 2005).

Benchmarking using regression analysis

It is possible to benchmark projects using regression analysis. As an example, in the univariate case, the regression model can handle both CRS and VRS, e.g. by using a multiplicative model of the form:

$$Y = \alpha X^{\beta} e^{u} \qquad (5)$$

where Y is labour, X is building size, and the u the error term. β is the scale factor (<1: economies of scale; =1: constant returns; >1: diseconomies of scale). The sign and magnitude of the residual is a measure of the productivity of an observation relative to the average (defined by the regression line). Equation (5) can be solved with linear regression using the loglinear form.

Productivity comparisons using regression analysis

Comparing the productivity of two individual observations may be straightforward if the residual has constant variance (homoscedastic). However, we suspect that building project data are heteroscedastic, with increasing spread. In this case, we would tend to identify mainly the largest projects as extreme performers (best or worst). The reason is that the largest projects in the data set would, in general, be farthest from the regression line on both sides. This problem could be handled by partitioning the data set into small, average and large projects. The challenge with this approach is to partition right to ensure we compare big with big and small with small. This would require some trial and error.

CONCLUSIONS

The conclusions in this paper are of two kinds: i) conclusions on the results of the empirical study and ii) conclusions on the usefulness of DEA.

As for the results, using DEA VRS, we identified six frontier projects. Two of them appear to be important role models, thus deserving to be studied as part of a process improvement initiative. The results further suggest that the average efficiency is 50-60%. Consequently, there seems to be a substantial improvement potential compared with the “best in class” projects. Thus, we recommend that one examines these projects, probably looking for differences with some of low productivity projects. However, due to a somewhat ambiguous data quality, we recommend caution in drawing firm conclusions with respect to productive and unproductive projects.

Regarding the usefulness of DEA, we conclude that it is a pragmatic, useful method for productivity assessments of *individual* projects. It is better than regression analysis in the presence of heteroscedasticity and multivariate inputs and outputs. Using regression on heteroscedastic data, we would tend to identify mainly the large projects as the most/least productive.

Also, DEA is appealing to a practitioner because it uses the *best practice frontier* as a benchmark rather than some theoretical baseline. As an alternative to DEA for *benchmarking*, multivariate regression might be used provided there is only one input. In this case, projects are assessed relative to the average rather than relative to the best.

There are no serious objections to be made against the DEA technique, in particular, but there definitely are objections to be made against productivity measurements and benchmarking, in general, be it with DEA or some other productivity measurement technique. The objections regarding its deterministic nature are general to productivity measurements and comparisons of *individual* observations and not particular to DEA. The problem of assessing an individual project rather than summary statistics of a group of projects is that we have to assume zero random errors in the former case. The reason is that it is impossible to know the exact size of model and measurement error of an individual project. We do not know of any techniques to differentiate between the random errors and true productivity differences caused by technical inefficiency. Therefore, whenever one performs productivity comparisons of individuals, one has to assume a perfect model and no measurement errors. So, the only advice is that one must ensure somehow that the random errors are small compared to productivity differences.

To summarise, productivity measurements are required to identify role models and best practice projects and to provide rough benchmarks and average efficiency scores. We recommend that productivity measurements be performed provided that certain guidelines are obeyed. First, the model must be valid, i.e. that the input and output indicators are carefully selected, and exogeneous factors are controlled for (to the extent possible). Second, we must have confidence that the model and measurement errors are small. Third, appropriate sensitivity analyses must be done. Fourth, productivity measurements should be used mainly to *assist* in identifying the best projects but not as the sole basis for compensation schemes or bombastic conclusions as to which project is deemed best.

Regarding future work, probably, the most useful research topic in order to add value from benchmarking exercises is to do more of the tedious stuff: improve data quality; strive to identify better productivity indicators; create benchmarking databases that practitioners can use.

ACKNOWLEDGEMENTS

This work was supported by the LINC centre at the Norwegian School of Management and Statsbygg Norway

REFERENCES

Afriat, S N (1972) Efficiency estimation of production functions. *International Economic Review*, **13**, 568-598.

Albriktsen, R and Førsund, F (1990) A productivity study of the Norwegian building industry. *J. Productivity Analysis*, **2**, 53-66.

Andersen, P and Petersen, N C (1993) A procedure for ranking efficient units in data envelopment analysis. *Man. Sci*., **39**(10), 1261-1264.

Banker, R D, Charnes, A, and Cooper, W W (1984) Some models for estimating technical and scale inefficiencies. *Man. Sci.*, **39**, 1261-1264.

Charnes, A, Cooper, W W, and Rhodes, E (1978) Measuring the efficiency of decision making units. *Eur. J. Oper. Res*., **2**, 429-444.

CURT (1982) Measuring productivity in construction, A construction industry cost effectiveness report, Report A-1, New York: The Construction Users Roundtable - CURT. Accessed on www.curt.org April 2005.

Edvardsen, D F (2004) *Four essays on the measurement of productive efficiency*, PhD thesis, Dept. Economics, School of Economics and Commercial Law, Göteborg University.

Førsund, F R and Hjalmarson, L (1979) Generalised Farrell measures of efficiency: an application to milk processing in Swedish dairy plants. *The Economic Journal*, **89**, 294-315.

Jonsson, J (1996) *Construction site productivity measurement: selection, application and evaluation of methods and measures*. PhD thesis, Luleå University of Technology.

Myrtveit, I and Stensrud, E (2005), On DEA and returns to scale in software development, Working paper.

Stensrud, E, Foss, T, Kitchenham, B, and Myrtveit, I (2002) An empirical validation of the relationship between the magnitude of relative error and project size. *Proc. METRICS'02,* IEEE Computer Society Press, Los Alamitos CA, 3-12.

Tobin, J (1958) Estimation of relationships for limited dependent variables. *Econometrica,* **26**, 24-36.

Torgersen, A M, Førsund, F R, and Kittelsen, S A C (1996) Slack adjusted efficiency measures and ranking of efficient units. *J. Productivity Analysis*, **7**, 379-398.