Target and Technical Efficiency in DEA – Controlling for Environmental Characteristics

by

Ole Bent Olesen

and

Niels Chr. Petersen

Discussion Papers on Business and Economics
No. 11/2008

FURTHER INFORMATION
Department of Business and Economics
Faculty of Social Sciences
University of Southern Denmark
Campusvej 55
DK-5230 Odense M
Denmark

Tel.: +45 6550 3271
Fax: +45 6550 3237
E-mail: lho@sam.sdu.dk
http://www.sdu.dk/osbec

ISBN 978-87-91657-27-6

Target and Technical Efficiency in DEA - Controlling for Environmental Characteristics

O. B. Olesen and N. C. Petersen

The University of Southern Denmark

September 26, 2008

Abstract

In this paper we propose *a target efficiency DEA model* that allows for the inclusion of environmental variables in a one stage model while maintaining a high degree of discrimination power. The model estimates the impact of managerial and environmental factors on efficiency simultaneously. A decomposition of the overall technical efficiency into two components, target efficiency and environmental efficiency, is derived.

Estimation of target efficiency scores requires the solution of a single large non-linear optimization problem and provides both a joint estimation of target efficiency scores from all DMUs and an estimation of a common scalar expressing the environmental impact on efficiency for each environmental factor.

We argue that if the indices on environmental conditions are constructed as the percentage of output with certain attributes present, then it is reasonable to let all reference DMUs characterized by a composed fraction lower than the fraction of output possessing the attribute of the evaluated DMU enter as potential dominators. It is shown that this requirement transforms the cone-ratio constraints on intensity variables in the BM-model (Banker and Morey 1986) into *endogenous handicap functions* on outputs. Furthermore, a priori information or general agreements on allowable handicap values can be incorporated into the model along the same lines as specifications of assurance regions in standard DEA.

1 Introduction

Nonparametric efficiency evaluation in the tradition of Data Envelopment Analysis (DEA) focuses on the distance from an observed input-output combination to the efficient frontier either in input or in output space. When all inputs are discretionary the inverse distance in input space has a nice interpretation as the possible radial contraction of input consumption. However, already in the discussion of Farrell's seminal paper (Farrell 1957) Winsten argued in a critical remark (Winsten 1957) on the Farrell indices that "*it is necessary to specify the range of variation (of choice perhaps) open to the firm before one can set up a standard of efficiency*" (p. 283). Inputs can in various analyses occur in two roles. Sometimes they are controlled by the managers and sometimes they are terms of production, typically denoted environmental conditions and not controlled by the managers. Winsten continues, "*we should expect an index of efficiency to distinguish these two roles*". In fact, quite often the production process is characterized by fixed or quasi-fixed inputs, and the radial distance in the full input space is not a meaningful concept to the managers. Farrell's suggestion of how to handle this case is simply to divide the observations into groups that are homogeneous with regard to such "quasi-factors".

The distinction between controllable and non-controllable inputs is further discussed by (Hall and Winsten 1959), who suggest a distinction between **target efficiency** and **technical efficiency**. If environmental conditions differ, then a particular environment defines a set of choices for management. When measuring target efficiency we "*are interested in how much a particular firm has to increase its output to reach the best in its particular environment . [...] If only we could change the management of this firm for the best of its kind, how much more output could we obtain?*" (p. 74). Target efficiency focuses on groups of firms being comparable with each other. Measuring technical efficiency, on the other hand, requires that the analyst marks out precisely which changes are caused by changes of technique. Notice that the two notions coincide, if there is no difference in environment. In this paper we will propose a method for estimating target efficiency and decomposing an overall technical efficiency index into a target and an environmental efficiency index. Our approach is an attractive alternative to the frequently used two stage models, where DEA scores are regressed on a set of environmental variables in the second stage. Like the two stage model our model provides only **one** estimated common scalar expressing the environmental impact on the overall technical efficiency for each environmental effect. Hence including environmental information implies a very limited loss of discrimination power.

Clearly, the distinction between short run and long run is important in relation to target and technical efficiency. The presence of fixed inputs in the short run problem of allocating variable inputs is well known. But the point to emphasize here is that even in a long run analysis certain inputs can be considered controlled or not controlled by the managers depending on what type of analysis is wanted. For example, if focus in a comparison of bank efficiency is on the performance including the siting of branches in various rural or urban areas then siting is part of the technique and consequently a controllable input. However, if focus is on the performance of each branch compared to other branches within the same bank or from other banks, then the actual siting of a specific branch is a non-controllable input and is part of the environment.

The presence of (quasi-)fixed inputs is quite common, especially of course in short run analyses. Examples are climate, location, level of advertising, median income in service area, number of competitors, age of production equipment, capacity, etc. (Quasi-)fixed inputs need of course not be productive inputs in the traditional sense and are in fact often so-called environmental variables.

Public production is an area where the concept of fixed inputs is very important. The provision of public services often requires significantly different effort in different areas because of heterogeneities among the households these areas. Analysis of the efficiency of education is a prominent example of a public production, where fixed inputs play an important role and we will use production of education below for motivating the approach proposed in this paper.

Significant impact from environmental variables on production implies that it is necessary to modify the "standard" DEA models (Charnes, Cooper and Rhodes 1978), (Banker, Charnes and

Cooper 1984) to control for non-discretionary inputs. (Banker and Morey 1986) recognize the inappropriateness of treating fixed input factors as discretionary in DEA and modify for that reason the constraints on the non-discretionary inputs (the BM-model) such that reference DMUs must have a level of each non-discretionary input less than or equal to the DMU under evaluation. There is no attempt to constrain the environmental effect to be of a common magnitude over the sample of DMUs, since the environmental variables enter a DEA for each DMU separately, which implies that the BM-model is sensitive to the "curse of dimensionality". Adding environmental dimensions implies that the discrimination power decreases.

(Ray 1988) proposes an alternative two stage model (for short 2SR-DEA) combining a first stage DEA with a second stage regression, where the environmental impact is estimated as a common magnitude over the sample of DMUs. A set of DEA efficiency scores is obtained in the first stage while ignoring differences in environment. This set of DEA scores is subsequently regressed on a number of environmental variables in the second stage.[1] 2SR-DEA involves a limited loss of discrimination power, since **a single** estimated common scalar expressing the environmental impact on the overall technical efficiency for each environmental effect is provided, and the approach has for that reason been used extensively in the litterature. However, it is well known that 2SR-DEA causes several well known problems of its own (see (Ray 2004), page 105).

In this paper we propose a target efficiency DEA model (TE-DEA model) that allows for the inclusion of environmental variables in a one stage model without a decrease in discrimination power caused by the curse of dimensionality as in the BM-model. Furthermore, combining the TE-DEA model with the basic BCC-model we propose a decomposition of the overall technical efficiency into two components: target efficiency and environmental efficiency. The proposed target efficiency model does not suffer from the problems of the 2SR-DEA. Solving one large optimization provides both a joint estimation of target efficiency scores from all DMUs and an estimation of a common

[1] (Simar and Wilson 2007) have recently proposed a coherent Data Generating Process (DGP) that can serve as a rationale for this approach.

environmental impact on efficiency. Hence, the proposed TE-DEA model and the commonly used 2SR-DEA share the same structure with regard to how the environmental effects are parameterized. The model will be developed in three steps. A so-called Value Weighted Banker-Morey model, the VWBM-model, designed to control for the volume of output is first derived for the multiple inputs and one output case. The single output VWBM-model is next generalized to the multiple inputs and multiple outputs case, which in turn provides the foundation for the development of the TE-DEA-model.

Recently, (Yang and Paradi 2003), (Paradi, Vela and Yang 2004) have introduced the so-called "**a handicap function**", which is used to adjust inputs and/or outputs in order to allow for a more fair comparison of DMUs operating in very different environments. DMUs operating in a more favourable environment are penalized by a higher input handicap (inputs are increased) and/or a lower output handicap (outputs are decreased). An important aspect of the VWBM-model as well as the TE-DEA-model is that the framework can be shown to involve an **endogenously** determined handicap function resembling the structure suggested by (Paradi et al. 2004). Furthermore, a priori information or general agreements on allowable handicap values can be incorporated into the model along the same lines as assurance regions in standard DEA.

The paper unfolds as follows. Section 2 formally introduces the BM-model, the R-model and our new VWBM-model, and it is discussed how these models relate to each other. The VWBM-model is derived based upon the simplifying assumption of one output to allow a simple presentation of the basic idea. The general multi input and multi output VWBM-model is presented in section 3. The target efficiency model is introduced in section 4. Its relation to the two stage model is analyzed, and a decomposition of the overall technical efficiency into target and environmental efficiency is derived. Section 5 presents an application of the general model on a data set comparing the production of education in a number of OECD countries and section 6 finally provides some concluding remarks.

2 Description of the VWBM-model

Consider n DMUs each consuming m different inputs $X_j \in \mathbb{R}_+^m, \forall j$, and producing only one output[2] $Y_j \in R_+, \forall j$, and with each DMU operating in an environment characterized by p indices, $Z_j \in R_+^p, \forall j$. For convenience, let the indices be defined such that $Z_{lj_1} < Z_{lj_2}$ implies that the DMU_{j_1} is operating in a more harsh environment in the l'th dimension compared to DMU_{j_2}. The model proposed by (Banker and Morey 1986) (the BM-model, output oriented) includes p constraints involving the environmental indices which force a reference DMU to have a level of the non-discretionary environmental index less than or equal to the DMU under evaluation:

$$\begin{array}{llcll} \max & \theta^{BM} & & & (1.a) \\ s.t. & \sum_{j=1}^{n} \lambda_j X_j & \leq & X_0 & (1.b) \\ & \sum_{j=1}^{n} \lambda_j Z_j & \leq & Z_0 & (1.c) \\ & -\sum_{j=1}^{n} \lambda_j Y_j + \theta^{BM} Y_0 & \leq & 0 & (1.d) \\ & \sum_{j=1}^{n} \lambda_j & = & 1 & (1.e) \\ & \lambda \geq 0 & & & \end{array} \tag{1}$$

(Ruggiero 1996) (the R-model) argues that the BM-model may lead to reference points that are not feasible. Hence, Ruggiero suggests a correction of the problem by excluding all DMUs with a more favorable environment from the analysis of the DMU under evaluation. Hence, no DMU with a higher environmental index in any dimension than the evaluated DMU should be allowed in a (virtual) dominating combination. Replacing (1.c) with $Z_j > Z_0 \Longrightarrow \lambda_j = 0, \forall j,$ and θ^{BM} with θ^R gives us the R-model. The BM-model is easily seen to be less restrictive compared to the R-model so that $\theta^{BM} \geq \theta^R$. We will argue that the R-model is overly conservative when environmental indices are reflecting the likelihood of the presence of such attributes in relation to the output being produced

[2] The simplifying assumption of one output is made to accommodate the need for a presentation of the basic idea behind the VWBM-model that is as simple as possible. The general case is covered in the next section.

In this paper we focus on on the output oriented model, since this orientation is used in the application in section 5. It is straight forward to change the various models to an input orientation.

The BM-model does not account for the fact that reference DMUs in general are composed of DMUs of a highly different size. Thus, a 'small' λ_k ('large' λ_l) may correspond to a large DMU$_k$ (small DMU$_l$), which in turn implies that the above weighting is not appropriate, since the optimal λ-values do not reflect the influence of respective DMUs in the construction of the reference DMU. We denote this problem "Lack of control for volume of production"[3]. To be more specific consider a convex combination of two schools, a very large school A and a very small school C operating in a very friendly and a very harsh environment, respectively. Let a third school B operating in an environment characterized by $Z_B = \frac{1}{2}(Z_A + Z_C)$ be the one under evaluation. The BM-model allows the virtual combination $\frac{1}{2}A + \frac{1}{2}C$ based on discretionary inputs and outputs to dominate B since this virtual combination by assumption operates in an environment corresponding to Z_B. The part of output produced under very friendly/very harsh conditions by the dominating DMU is $\frac{1}{2}Y_A$ and $\frac{1}{2}Y_C$, respectively. But $\frac{1}{2}Y_A >> \frac{1}{2}Y_C$. Hence, *an unreasonably large part of the dominating output is produced under friendly conditions* compared to the relative size of Z_A and Z_C. To prevent this situation without going to the extreme represented by the R-model, the VWBM-model (2) below insists that the share of output produced under friendly conditions by the dominating DMU is used as a weight on the DMU's environmental index.[4] In other words, the combined school is only allowed to dominate B if the part of the combined school's output produced under harsh conditions is at least as large as the part of B's output being produced under such harsh conditions. Hence, with environmental indices defined such that smaller values implies a more harsh environment, the weighted sum of the environmental indices with weights set equal to respective shares of total production of output for the reference DMU must not exceed the socioeconomic index for the DMU under evaluation. The suggested VWBM-model (output oriented) designed to control for the volume of output has the following structure:[5]

[3] In this paper we focus on produced output as a measure of size. It is left for future research to consider measures of size based on input consumption.

[4] Clearly, the weighting of environmental variables is only meaningful in the case of cardinal data. The treatment of ordinal data is not considered in this paper.

[5] It is of interest to observe that the VWBM- and the BM-model will approach each other for increasing sample size, since facets become smaller which in turn implies a smaller deviation between volume of produced outputs among the DMUs spanning each facet.

$$
\begin{array}{llcclr}
\max & \theta^{VWBM} & & & \text{duals} & (2.a) \\
s.t. & \sum_{j=1}^{n} \lambda_j X_j & \leq & X_0 & v \in \mathbb{R}_+^m & (2.b) \\
 & \sum_{j=1}^{n} \lambda_j Y_j \left(Z_j - Z_0\right) & \leq & 0 & g \in \mathbb{R}_+^p & (2.c) \\
 & -\sum_{j=1}^{n} \lambda_j Y_j + \theta^{VWBM} Y_0 & \leq & 0 & u \in \mathbb{R}_+ & (2.d) \\
 & \sum_{j=1}^{n} \lambda_j & = & 1 & v_0 \in \mathbb{R} & (2.e) \\
 & \lambda \geq 0 & & & &
\end{array}
\tag{2}
$$

The key point in this model compared to the R- and BM-model is that the environmental indices Z_j are weighted with the respective shares of total production Y_j. The model generates for that reason in general different results compared to the BM-model, when the DMUs differ in size. The approach is highly appropriate in instances with environmental conditions defined by the percentage of the output with certain attributes present.[6] For example percentage of parents with college education, percentage of brick houses in relation to efficiency of fire stations, different banking strategies on the composition of output imposed on branches from headquarter, demographic characteristics of households, income distribution, occupational characteristics, educational background either related to individuals or to the population in different areas, such as school districts or municipalities, competitive environments, etc.[7]

The p constraints related to the environmental indices can be rewritten as $\sum_{j=1}^{n} \left[\sum_{k=1}^{n} \lambda_k Y_k\right]^{-1} \lambda_j Y_j Z_j \leq$

[6] The collection of additional data allowing for a transformation of indices to quantities is a more straightforward approach in applications where environmental conditions are not defined in terms of output with certain attributes.

[7] For illustration, consider an example from (Ruggiero 2004) with three DMUs A, B & C producing one output Y using one discretionary input X and one non-discretionary input Z measuring harshness of environment. The $[Y, X, Z]-$vectors from A,B& C are $[10, 10, 20]$, $[10, 8, 30]$ and $[10, 1, 40]$. Let Y be number of (100) students, let X be number of (10) teachers, and let Z be the percentage of parents with a college education. Using the model proposed by Ruggeiero all DMUs are efficient. B is efficient, since A cannot dominate B on its own. B is inefficient in the BM-model and is dominated by the convex combinations $\frac{1}{2}A + \frac{1}{2}C$. In the context of an efficiency evaluation of schools, a precise interpretation of the convex combination $\frac{1}{2}A + \frac{1}{2}C$ is available and the expected Z-index at this virtual school equals the harshness index of B. The expected number of students at the virtual school $\frac{1}{2}A + \frac{1}{2}C$ with a parent with a college education is $100 + 200$ and the corresponding Z-index is $\frac{300}{1000} \times 100 = 30$. Hence $\frac{1}{2}A + \frac{1}{2}C$ should be allowed to dominate B and this is indeed the outcome of the volume weighted BM-model (VWBM-model) to be suggested in next section.

Z_0 with the following interpretation: i) Y_j measures volume of production at DMU$_j$, ii) $\lambda_j Y_j$ measures volume of production at DMU$_j$ in the virtual DMU serving as reference DMU, and iii) $\left[\sum_{k=1}^{n} \lambda_k Y_k\right]^{-1} \lambda_j Y_j$ measures the share of volume of production for the virtual DMU produced at DMU$_j$. Hence, the virtual DMU is required to have a level of each non-discretionary factor that is less than or equal to the DMU under evaluation with the shares above serving as intensity factors for Z_j.

It was mentioned in the introduction that the idea of a handicap function relates to the structure of the VWBM-model. To see this, consider the VWBM-model (2) where a weighted combination of the p environmental constraints (2.c) with non negative weights $w = [w_1, \ldots, w_p]^T$ is added to the traditional output constraint (2.d):

$$\sum_{j=1}^{n} \lambda_j \left[1 - w^T (Z_j - Z_0)\right] Y_j \geq \theta^{VWBM} Y_0 \tag{3}$$

Clearly, any feasible solution in (2) satisfies (3). Output Y_j from DMU$_j$, being part of a dominating virtual DMU, is multiplied by the factor $\left[1 - w^T (Z_j - Z_0)\right]$ which is greater than one if $Z_j < Z_0$ and smaller than one if $Z_j > Z_0$. Let us identify an "optimal" w by considering the dual program to (2):

$$\begin{array}{llcll} \min & v^T X_0 + v_0 & & & \text{duals} \\ s.t. & v^T X_j + g^T (Z_j - Z_0) Y_j - u Y_j + v_0 & \geq & 0, \forall j & \lambda \in \mathbb{R}_+^n \\ & 1 - u Y_0 & = & 0 & \theta \in \mathbb{R} \\ & u \geq 0, v \geq 0, v_0 \in \mathbb{R}, g \geq 0 & & & \end{array} \tag{4}$$

Let the handicapped output be $\widehat{Y}_j = \widehat{h}_j Y_j = \left[1 - u^{-1} g^T (Z_j - Z_0)\right] Y_j$, $\forall j$ (notice $\widehat{Y}_0 = Y_0$)[8]. Hence, "the optimal" $w = u^{-1} g$ and the VWBM-model (4) can be interpreted as a standard BCC-model on the handicapped input-output vectors $\left(X_j, \widehat{Y}_j\right), \forall j$. The handicap function is seen to be

[8] For ease of exposition we have omitted non-Archimedians as lower bound on the virtual multipliers. A stringent approach would require $u > 0$. Hence u^{-1} always exists.

linear[9] in outputs. Notice that $\widehat{h}_j$ is determined from two sources: i) the exogenously given vector $(Z_j - Z_0)$, and ii) the endogenously determined vector of relative multipliers $u^{-1}g \in \mathbb{R}^p_+$. Hence, the VWBM-model is compatible with the use of **an endogenous handicap function**. Consider e.g. a handicapped output from the k'th DMU, $\widehat{Y}_k$, where $(Z_k - Z_0) > 0$. In this case $\widehat{h}_k \leq 1$[10] reflecting that the $k'th$ DMU is operating in a more favorable environment than the evaluated DMU and handicapped output $\widehat{Y}_k$ is for that reason given a value strictly less than Y_k.

3 The VWBM-model in the case of many inputs and many outputs

Let each DMU$_j, \forall j$, produce s different outputs $Y_j \in \mathbb{R}^s_+, \forall j$, using m inputs $X_j \in \mathbb{R}^m_+, \forall j$, in an environment characterized by p indices, $Z_j \in R^p_+, \forall j$, and assume that the p environmental indices relate to the s outputs[11]. The proposed use of the level of the outputs as weights on these indices extends to the case with multiple inputs and multiple outputs as follows:

$$\begin{array}{lllll} \max & \theta^{VWBM} & & & \textit{duals} \\ s.t. & \sum_{j=1}^n \lambda_j X_j & \leq & X_0 & & v \in \mathbb{R}^m_+ \\ & \sum_{j=1}^n \lambda_j Y_j \left(Z_{lj} - Z_{l0}\right) & \leq & 0 & l = 1, \dots, p & w^l \in \mathbb{R}^s_+ \\ & -\sum_{j=1}^n \lambda_j Y_j + \theta^{VWBM} Y_0 & \leq & 0 & & u \in \mathbb{R}^s_+ \\ & \sum_{j=1}^n \lambda_j & = & 1 & & u_0 \in \mathbb{R} \\ & \lambda \geq 0, \theta^{VWBM} \text{ free} & & & & \end{array} \tag{5}$$

The ps constraints on environmental characteristics are homogeneous inequalities defining a

[9] The challenging task of incorporating non-linear handicap functions is left for future research.

[10] Analysis of assurance regions such as $\widehat{h}_k = \left[1 - u^{-1} g^T \left(Z_j - Z_0\right)\right] \geq 0$ and more generally in the next section $g^T \left(Z_j - Z_0\right) \leq 1 = u^T Y_0$ is left for future research.

[11] The set of environmental indices may well relate to a subset of the outputs, and two different indices may not relate to the same subset. One must in cases like that select which combinations of the p indices and the s outputs it is meaningful to include in model (5). To simplify we have included all ps constraints in the presentation of the model (5).

cone-ratio in λ-space. The dual to (5) is

$$\begin{array}{lllll} \min & v^T X_0 + v_0 & & & \\ s.t. & v^T X_j + \sum_{l=1}^{p} \left(w^l\right)^T Y_j \left(Z_{lj} - Z_{l0}\right) - u^T Y_j + v_0 & \geq & 0, \forall j & \\ & 1 - u^T Y_0 & = & 0 & \\ & u \geq 0, v \geq 0, v_0 \in \mathbb{R}, w^l \geq 0, \forall l & & & \end{array} \tag{6}$$

The corresponding model (2) in the previous section with only one output has p constraints related to the control of environmental characteristics. In the general model (5) we enforce such constraints to be satisfied *for each of the outputs* so that we have p constraints $\sum_{j=1}^{n} \lambda_j y_{jk} \left(Z_j - Z_0\right) \leq 0$ for each index k. The model can be given an interpretation in terms of the motivating example on education in schools. Let the set of multiple outputs be defined by the number of students in various grades. According to (5) the fraction of parents with college education at the reference school must not exceed this fraction at the evaluated school *for any subgroup of students as defined by grades.*

(5) maintains the hypothesis that the p environment indices relate to the s outputs, which in turn implies a dual structure (6) with handicapped outputs, see Section 2. However, (5) allows in the case of multiple outputs the s outputs to be handicapped differently. There is no structure in the model restricting how different outputs are handicapped, which means that the model includes s independent handicap functions $h_k\left(Y_{kj}\right) \equiv \widehat{h}_k Y_{kj} = \left[1 - \sum_{l=1}^{p} u_k^{-1} w_k^l \left(Z_{lj} - Z_{l0}\right)\right] Y_{kj}$, $\forall k$, specifying how each output k is handicapped.

Two outputs are used in the application below, namely the sum of scores obtained in two (sets of) tests of students enrolled in secondary education programs in a number of OECD countries in combination with an environment index on parental background. It will be argued that it is not meaningful to handicap the two sets of test scores differently based on the variation in the background of the parents. Imposing the condition that any handicapping should have an equal impact on the two test-scores requires the additional constraints: $w_{k_1}^l / w_{k_2}^l = u_{k_1} / u_{k_2}, \forall k_1, k_2$, or equivalently that $w^l = g_l u, g_l \in \mathbb{R}_+, l = 1, \ldots, p$. Including these constraints in (6) the model can be rewritten as follows:[12]

$$
\begin{array}{llcll}
\min & v^T X_0 + v_0 & & & \text{Lagrange multipliers} \\
s.t. & u^T Y_j - g^T (Z_j - Z_0) u^T Y_j - v^T X_j - v_0 & \leq & 0, \forall j & \lambda \in \mathbb{R}^n_+ \\
 & 1 - u^T Y_0 & = & 0 & \theta \in \mathbb{R} \\
 & u \geq 0, v \geq 0, v_0 \in \mathbb{R}, g = [g_1, \ldots, g_p]^T \geq 0 & & &
\end{array}
\qquad (7)
$$

The solution of model (7) requires a non-linear optimization solver, but an equivalent formulation in terms of a Mixed Integer Linear Program is available, see Appendix. A comparison of (7) with the corresponding model (4) from section 2 reveals that the endogenous handicap function with one output $\left[1 - u^{-1} g^T (Z_j - Z_0)\right]$ in this more general model is replaced by $\left[1 - g^T (Z_j - Z_0)\right]$. Hence, the interpretation of the VWBM-model (7) as a standard BCC-model on a set of endogenously handicapped input-output vectors is still valid.

Let us trace the primal structure from the Lagrange function for (7). The following characteristics for an optimal solution to (7) can be derived by the Kuhn-Tucker conditions:

$$
\begin{array}{lll}
X_0 - \sum_{j=1}^n \lambda_j X_j & \geq 0 & (8.a) \\
\sum_{j=1}^n \lambda_j \left[1 - g^T (Z_j - Z_0)\right] Y_j - \theta Y_0 & \geq 0 & (8.b) \\
- \sum_{j=1}^n \lambda_j u^T Y_j (Z_j - Z_0) & \geq 0 & (8.c) \\
1 - \sum_{j=1}^n \lambda_j & = 0 & (8.d)
\end{array}
\qquad (8)
$$

These conditions require some remarks. Firstly, (8) defines a feasible set in (λ, θ)-space contained in the feasible set in (5) for any fixed vector g. Secondly, the cone ratio structure in (5) (and in (2)) with the traditional output constraints included but with the intensity variables constrained by

[12] Another way to argue for this model is to consider facet inducing dual constraints $F(x, y, \Delta z) = u^T y - g^t \Delta z u^T y - v^T x = 0$ and to argue that the marginal rate of substitution between observed outputs must equal the marginal rate of substitution between handicapped outputs along the facet, i.e. $\frac{dy_2}{dy_1} = -\frac{u_1 - g^t \Delta z u_1}{u_2 - g^t \Delta z u_2} = -\frac{\left(1 - g^t \Delta z\right)}{(1 - g^t \Delta z)} \frac{u_1}{u_2}$ must be equal to $\frac{d\left(1 - g^t \Delta z\right) y_2}{d(1 - g^t \Delta z) y_1} = -\frac{u_1}{u_2}$. This links nicely to the data generating process proposed by (Simar and Wilson 2007).

environmental cone ratio constraints is in (8) replaced by an endogenous handicap function on the outputs (8.b). Hence, we do not insist that the reference DMU must produce at least as much as the evaluated DMU in all output dimensions. We "only" require that the reference DMU after the handicapping of outputs produces at least as much as the evaluated DMU. Thirdly, the intensity variables λ_j are still constrained by environmental cone ratio constraints (8.c).

4 A target efficiency model and its relation to the 2SR-DEA model

As mentioned in the introduction it is an extensively used practise in the literature to apply the two stage approach (2SR-DEA) (Ray 1988)(Ray 1991)(Ray 2004) to control for environmental differences[13]. Ignoring the differences in environment in the first stage a set of DEA efficiency scores $\theta_j, j = 1, \ldots, n$, is obtained. The resulting indices capture both the environmental and the managerial impact on performance. These scores are in the second stage regressed on a number of environmental variables $Z_1, \ldots, Z_n, Z_j \in \mathbb{R}^p_+$. Using a linear regression we maintain a model

$$\theta_j = \beta_0 + \beta_1^T Z_j + \varepsilon_j, \quad j = 1, \ldots, n \tag{9}$$

where $\varepsilon_j, j = 1, \ldots, n$, are the random disturbances, so that $E(\theta|Z_j) = \beta_0 + \beta_1^T Z_j$ defines a conditional mean of the radial expansion factor θ_j. (Simar and Wilson 2007) have recently proposed a coherent Data Generating Process (DGP) that serves as a rationale for this approach and argue that the correct regression approach based on this DGP is a truncated regression as opposed to the more commonly used Tobit regression[14].

[13] An extensive list of references to applications of such two stage approaches can be found in (Simar and Wilson 2007).

[14] An OLS approach is of course problematic since the efficiency scores obtained have a bounded support, either (0,1] or [1,∞).(Simar and Wilson 2007) are also concerned with the problems that DEA scores are serially correlated, a phenomenon that creates problems for the use of traditional inference in the regression stage and that the obtained scores are biased.

In this section we show that the 2SR-DEA model relates closely to a model termed the *target efficiency DEA model* (TE-DEA model), which involves a joint estimation of all n efficiency scores from the VWBM-model (7) with one additional requirement, namely **equal handicapping for all DMUs**. We will formulate one large optimization problem and estimate the efficiency scores for all DMUs simultanously by minimizing a sum of residuals towards a traditional piecewise linear envelopment. We will add the requirement that the endogenous parts of the handicap functions[15], i.e. the decision-vectors g of multipliers in (7), must be identical for all DMUs. Hence, the approach will become an appropriate alternative to the 2SR-DEA with a number of advantages to be discussed below[16].

Both the TE-DEA-model and the 2SR-DEA-model are coherent with the DGP proposed by (Simar and Wilson 2007)[17], since the VWBM-model works through a set of handicap functions depending on environment. Hence, for any given DMU under evaluation all potential DMUs contributing to the virtual dominating DMU will have their output vectors adjusted through a radial contraction or expansion depending on the difference in harshness of environment.

Consider an output vector Y_j from DMU$_j$. The second stage 2SR-DEA model (9) predicts, using the conditional mean of the radial expansion factor, that this output vector will be projected to the frontier by a factor $\left(\beta_0 + \beta_1^T Z_j\right) Y_j$. Clearly, if DMU$_j$ was operating in the same environment as DMU$_l$, $l \neq j$, we would expect to see this output vector projected to the frontier as $\left(\beta_0 + \beta_1^T Z_l\right) Y_j$. On the other hand, with $Z_j \neq Z_l$ the difference between the two projections can be interpreted as

[15] Without this additional requirement this large optimization model is simply n separable optimization problems, but adding the requirement of one common g implies a non-separable nonparametric estimation of this common handicap function and all scores.

[16] The same idea can be applied to the classical BM-model (Banker and Morey 1986). Solve an estimation of all BM-scores using a simultaneous estimation of n dual programs to the BM-model in (1) with a set of linking constraints. These constraints must force each environmental virtual multiplier relative to all other virtual multipliers to be equal across the set of DMUs.

[17] An important question is of course whether or not the DGP described in (Simar and Wilson 2007) is reasonable. In this formulation, the environmental variables influence the mean and variance of the inefficiency process, but not the boundary of its support. A number of different alternatives are discussed in ((Coelli, Rao and Battese 1998), pp. 166-171). Among these alternatives are the BM- and the R-model which are coherent with a DGP where the boundary of the production possibility set is dependent on the environment.

how much the output vector of DMU_j should be handicapped in order to enter in a fair (environment being equal) relative efficiency analysis of DMU_l. Hence, the 2SR-DEA is consistent with a handicap function given by $h(Y_j) \equiv \widehat{h}Y_j = \left[1 + \beta_1^T (Z_j - Z_l)\right] Y_j$. Consider the following example with all DMUs using the same amount of input. Consider two output vectors $Y_j = (1,4)$ and $Y_l = (2,1)$ from the j'th and the l'th DMU working under very different environment with $Z_j = 80$ and $Z_l = 30$. Let the estimates from the second-stage regression be $\widehat{\beta}_0 = 2.4, \widehat{\beta}_1 = -0.01$ and assume that a BCC-model ignoring the differences in environment based on the full sample of DMUs provides output oriented efficiency scores $\theta_j = 2$ and $\theta_l = 3$. The expected difference between the efficiency of DMU_j working in environments Z_j and Z_l respectively is $E(\theta|Z_j) - E(\theta|Z_l) = \beta_1 (Z_j - Z_l)$ and DMU_j'th output vector $(1,4)$ should be handicapped by a factor $1 + \beta_1 (Z_j - Z_l) = 0.5$.

Let us formally specify the *target efficiency DEA model* (TE-DEA model) as the *joint estimation of all n efficiency scores* from the VWBM-model (7) with one additional requirement, namely **equal handicapping for all DMUs.** The following program provides a joint estimation of all scores from the VWBM-model:

$$\begin{array}{llll} \min & \sum_{l \in N} v_l^T X_l + v_{0l} \quad (\equiv \sum_{j \in N} (1 + s_{jj})) & & \\ s.t. & u_l^T \left[1 - g_l^T (Z_j - Z_l)\right] Y_j - v_l^T X_j - v_{0l} + s_{jl} & = & 0, j \in N, l \in N \\ & 1 - u_l^T Y_l & = & 0, l \in N \\ & u_l \geq 0, v_l \geq 0, v_{0l} \in \mathbb{R}, g_l = [g_{1l}, \ldots, g_{pl}]^T \geq 0, \forall l & & s_{jl} \geq 0, \forall j, l \end{array} \tag{10}$$

where $N = \{1, \ldots, n\}$. A comparison of the handicap functions in (10) vs. 2SR-DEA reveals that a requirement of equal size of the endogenous part of the handicap function is imposed in 2SR-DEA but not in (10), where each DMU "can pick" its own vector $g_l, l \in N$. Hence, TE-DEA can be made more comparable with 2SR-DEA by imposing the following restrictions in (10) enforcing equal decision vectors g_l for all DMUs:

$$g_l = g, l = 1, \ldots, n \tag{11}$$

The resulting model [(10),(11)] is the TE-DEA-model. This model provides a joint DEA estimation of target (or managerial) efficiency score θ_j^{Target} and a within DEA estimation of a common environment impact vector g consistent with the DGP above. Obviously, θ_j^{Target} is the value of the j'th component in the optimal objective function in the TE-DEA model, i.e. $\theta_j^{\text{Target}} \equiv \left(v_j^*\right)^T X_j + v_{0j}^*, j = 1, \ldots, n$ where (v_j^*, v_{0j}^*) are the optimal decision vectors from [(10),(11)].

The TE-DEA model is an attractive alternative to the 2SR-DEA-model, since there are several well known problems related to the use of 2SR-DEA that can be avoided by using [(10),(11)]. (Ray 2004) and (Coelli et al. 1998) summarize some of these problems as i) a choice of functional form for the regression is involved, ii) the need for a Tobit/Truncated regression since the scores have a bounded support, either [0,1] or $[1,\infty)$, iii) difficulties getting a coherent managerial inefficiency measure from the estimated residuals, since the residuals have a zero mean, and iv) the results are likely to be biased if the variables used in the first stage are highly correlated with the second-stage variables.

The TE-DEA model does not rely on a specific functional form and a target efficiency score θ_j^{Target} is easily provided. In addition, an environment efficiency score θ_j^{Env} is available by a decomposition of the overall technical efficiency into target and environmental efficiency defining θ_j^{Env} from: $\theta_j^{\text{Env}} \times \theta_j^{\text{Target}} \equiv \theta_j^{\text{BCC}}$, where θ_j^{BCC} is the score obtained from the BCC-model ignoring the environmental information.

Comparing the BCC-model and the BM-model (where environmental characteristics are included in the second but not in the first) implies by structure (in an output oriented model) that the BM-scores are at least as large as the BCC-scores. Hence, controlling for environmental differences using (1) implies loss of discrimination power. However, comparing the BCC-model and the TE-DEA-model (with environment included in the second but not in the first) does not by structure provide any ordering on the obtained scores. BCC-efficiency does *not* imply target efficiency, and it is indeed possible that a BCC-undominated DMU in an output-oriented model is given an environmental index above one reflecting inefficient performance in a rather friendly environment. This phenomenon is illustrated in the application in the next section.

((Ray 2004), page 105) argues that the requirement of a priori decisions on whether an environmental variable has a favorable or an unfavorable impact on production is an inherent weakness of models characterized by an inclusion of environmental variables as non-discretionary inputs or outputs. One has to choose to whether to include the non-discretionaries as inputs or outputs. By contrast, the direction of impact is determined endogenously in 2SR-DEA. However, as observed by ((Coelli et al. 1998), page 168) the problem can be remedied by an inclusion of non-discretionaries in terms of equality constraints, which allows for an endogenous determination of direction of impact . This idea translates in relation to the TE-DEA- and the VWBM-model into using the vector g as decision variables unconstrained in sign and letting the estimation procedure determine the direction of influence.

5 A cross country comparison of the OECD educational efficiency of the lower and upper secondary educations

Primary and secondary education have received much of attention in recent years. In response to the need for cross-country comparisons on student performance OECD launched in 1997 the so-called PISA program (PISA is the acronym for Program for International Student Assessment). Performance of students from 40 countries was compared by an evaluation using the same questionaires. In relation to the models discussed in this paper we will apply the various models to data from the PISA 2003 project to perform a cross country comparison of the OECD educational efficiency of the lower and upper secondary educations. The performance of 15-year-olds on the PISA reading, mathematics, problem solving, and science literacy scales in 2003 will be at focus.

We will use two outputs: the average of the three scores for mathematics, problem solving, and science literacy scales and the score for the reading literacy scale. The data describing the performance of students (the output) in the form of a mean score on the PISA reading, mathematics, problem solving and science literacy scales is extracted from (OECD 2004b)(OECD 2004c). The three different scores relating to mathematics, science and problem solving are averaged into one index. The scores reported are mean scores and are not quite applicable for an efficiency analysis comparing the performance across countries of very different size. We "de-normalize" these mean scores by multiplying scores with the number of students enrolled in lower and upper secondary education, all educational programmes, full-time and part-time. We use an average enrollment from 2000-2002 extracted from the educational database at www.oecd.org.

Two inputs are used in the application, namely "Total intended instruction time in hours per year for 12 to 14-years-old students" and "Students per teaching staff in public and private institutions, secondary educations", (see Table D1.1 and Table D2.2 in (OECD 2002-2004a)). An average over the three years 2000-2002 prior to the PISA analysis of these inputs are used after being converted to absolute amounts of teachers and instruction hours by multiplying with the average enrollment in the lower and upper secondary educations in the period 2000-2002[18]. As environmental variable we use Parental Education Attainment 2001-2002 defined as the population that has attained at least upper secondary education, aged 35-44 average for 2001-2002 (OECD 2002-2004a), Tables A1.2 and A2.2.

The data and some summary statistics are included in the appendix. Table 1 presents the efficiency scores from this application. Columns 2-5 contain for each of 27 countries the scores from the R-, the VWBM-, the BM-model and finally from a BCC-model without the environmental variable.

[18] The PISA scores from 2003 were given to 15 year old students. We use the student-teacher ratio from the previous 3 years as a measure of the possible teacher impact on the performance measured by these PISA scores. We aim at an efficiency analysis of the lower and upper secondary educations. Hence, we multiply all the PISA scores (means) and the teacher-student ratio by the average (2000-2002) enrollment in lower and upper secondary educations. This procedure requires an assumption, that the performance in 2003 of the 15 year old students are representative for the performance of all 5 grades in the secondary education, i.e. for similar performance in the previous 5 years.

Insert Table 1 here.

Table A3 in appendix presents the dominating DMUs from the analysis in the VWBM- and the BM-model. A comparison of columns 2-4 in Table 1 reveals that i) the R-model always provides the best (lowest) score of the three models, ii) if scores from the VWBM- and the BM-model differ then for almost all countries the VWBM-model provides the best (lowest) score compared to the BM-model, and iii) in one case (Thailand) the BM-model provides a slightly better score than the VWBM-model. This general pattern is what we would expect. If we consider a dominating DMU consisting partly of a DMU with low output production operating in a harsh environment and a DMU with a large output production operating in a favorable environment, then we would expect that the solution provided from the BM-model is no longer feasible when the weights λ_j are replaced with $\lambda_j^{VWBM} = \lambda_j u^T Y_j \left[\sum_{j=1}^{n} \lambda_j \left(u^T Y_j\right)\right]^{-1}$. The introduction of the volume of output in the environmental constraint implies that the weight becomes much larger on the environmental index from the large DMU operating in a favorable environment, which tends to increase the resulting convex combination of socio-index above the level of the index of the DMU being evaluated. On the other hand, if we consider a dominating DMU consisting partly of a DMU with low output production operating in a favorable environment and a DMU with a large output production operating in a harsh environment, then we would expect that the solution provided by the BM-model remains feasible when the weights λ_j are replaced with λ_j^{VWBM}.

Consider the scores from the VWBM- and the BM-model for *Italy* being dominated by primarily Japan (large output) and Turkey (smaller output). Japan is operating in a much more favorable environment compared to Turkey. Imposing the weights (0.345, 0.631) from the BM-optimal solution on the λ-values in the VWBM-model we get (0.609,0.296) which implies an output weighted environmental score of 0.70, which is way above the socio-index for Italy (0.49). Hence, the size of Japan is prohibitive when a dominating DMU is to be formed in the VWBM-model and Korea is substituted for Japan in the solution in the VWBM-model. Korea is smaller than Japan as measured by level of output, and the results is a lower VWBM-efficiency score compared to the

BM-score.

Consider the scores from the VWBM- and the BM-model for *Thailand* being dominated primarily by Mexico and Portugal in the BM-model. Mexico is very large and operates in a rather harsh environment Compared to Portugal. Hence, we are in the opposite situation compared to the Italy-case and would expect the BM-solution to be feasible in the VWBM-model. Indeed, imposing the dominance solution from the BM-model into the VWBM-model, we get that the output weighted socio-index of the dominating combination is 21.7, which is nearly feasible (Thailand index is 19.0).

Insert Table 2 here

Now let us illustrate the performance of the TE-DEA-model with a common handicap function and its relation to the 2SR-DEA-model. Table 2 presents the results from the various models. Column (1) contains the BCC-scores, column (2) the scores from the VWBM-model (7) and column (3) the corresponding endogenous handicap value g, which here varies over the DMUs. Column (5) contains the scores from the TE-DEA-model with a common handicap function, [(10),(11)], and the common g is estimated to 4.24×10^{-3}. The scores in column (5) reflect managerial efficiency when we control for environmental differences using a common handicap function. Defining the environmental impact in column (4) as the ratio of the BCC index and the TE-DEA-index, we get a decomposition of the BCC-score into an environmental and a managerial effect.

Notice that the VWBM-indices in column (2) by construction are lesser than or equal to the BCC-indices in (1), but the indices in column (5) do not posses this characteristic. For Czech, New Zealand, Norway and Slovak we get a higher index in the TE-DEA-model compared to the BCC-index, a phenomenon that is reflected by an index below one in column (4). Two of these countries are efficient in the BCC-model, but the estimated common handicap function indicates that they should be able to expand output. Inspecting column (3) one observes the obvious fact that these four countries may maintain their BCC score by setting $g = 0$. But an estimated common $g = 4.24 \times 10^{-3}$ based on a joint minimization of the sum of residuals implies that New Zealand and Norway become slightly inefficient and Czech and Slovak - with BCC-scores of 1.150 and 1.219 - get slightly higher scores with the common handicap function imposed. These four countries share a characteristic, namely rather high values of the environmental index. Hence these countries operate in very friendly environments, but obtain their efficiency score by insisting on $g = 0$. Although these DMUs are on the boundary of the convex hull estimator of the BCC-production possibility set, the estimated environmental effect implies that the relatively friendly environment of these DMUs should allow for a better performance than observed.

The following estimation allows for a comparison with the 2SR-DEA-model. The BCC scores have been regressed on the environmental indices using a truncated regression model with a lower bound of one on the independent variable. Following Ray (1998,2004), we use the residuals as "measures" of managerial inefficiency. To facilitate a comparison of TE-DEA and 2SR-DEA an additive shift of 0.19 of the residual is used. This shift factor is determined such that the average environmental index from the two models (in columns 4 and 8) are of the same magnitude. Hence the 2SR-DEA environmental index in column (8) is the sum of the shifted residuals and the BCC indices[19].

A regression on the full sample implies a very small (and insignificant) beta, and as a consequence there is very little variation in the environment indices in column (8). Hence, if we believe in these results, almost all variation in the BCC-indices is due to managerial inefficiency. However, this does not seem to be a reasonable conclusion, since it is caused by another problem related to the use of 2SR-DEA. Some of the included countries, most prominently Mexico and Uruguay, get a BCC score equal to one because their input-output vector in some sense specialize, and not necessarily because they perform "well". Mexico is characterized by a very low teacher to student ratio compared to the rest of the countries. Uruguay is the smallest DMU and is therefore deemed

[19] This ad hoc procedure provides the environmental index as the conditional mean of the BCC score added to the shift factor. The 2SR_DEA does not provide any estimation of the level of the environmental impact in a decomposition of the BCC-score into an environmental and a managerial part. This is really an econometric identification problem since no information is provided on how to disentangle the estimated intercept term into a part related to the environmental and a part related to the managerial impact. Notice that the estimation using the TE-DEA-model [(10),(11)] with a common handicap function provides all the information needed for the decomposition of the overall index into these two aspects. Ray suggests a shift of the residuals by adding the smallest residual.

efficient in a BCC-model. To get comparable results from the 2SR-DEA-model we provide an additional experiment by deleting these observations from the regression (together with four other rather special countries). On this reduced sample we get more meaningful results with a numerically larger beta (-0.00312). The same procedure as above is used to get the BCC index decomposed based on shifted residuals from this regression. The results are included in columns (6) and (7). Notice the rather high degree of agreement between the scores in columns (4) and (6) and the agreement between $-\beta = 3.13 \times 10^{-3}$ and $g = 4.24 \times 10^{-3}$ from [(10),(11)].

6 Summary and conclusion

In this paper we propose *a target efficiency DEA-model* that allows for the inclusion of environmental variables in a one stage model without the decrease in discrimination power that characterizes the Banker Morey model. The model does not suffer from the problems of the two stage model (Ray 1988) and a decomposition of the overall technical efficiency into target and environmental efficiency follows from a comparison with the BCC-model.

Within the DEA literature designing efficiency models that control for environmental characteristics is typically done using one of two approaches: i) The BM-model, which includes the environmental variables directly into the DEA-model as non-discretionary variables, or ii) a two stage regression model proposed by Ray in 1988, where the first stage consists of a DEA estimation ignoring environmental differences and the second stage of a regression of the obtained scores on environmental variables. The target efficiency model is an attractive alternative to the two stage regression model. Contrary to the Banker-Morey model both models allow for the inclusion of environmental variables without significant loss of discriminatory power. Estimation of the target efficiency scores requires a solution of one large non-linear optimization problem providing both a joint estimation of target efficiency scores for all DMUs and an estimation of **a single** common scalar expressing the environmental impact on efficiency for each environmental effect.

The general idea behind the TE-DEA-model is the formulation of the VWBM-model combined with a simultaneous estimation of all scores and restrictions on the choice of "impact from environment", such that only a common effect for each environmental characteristic is allowed. Hence, a maintained hypothesis of a common impact from each of the included environmental characteristics over the set of DMUs is imposed. The impact from environment works only through restriction of feasible radial expansions or contractions of output and is independent of each DMU's choice of input or output mix. The same approach is used in (Simar and Wilson 2007) to derive a coherent data generating process for the two stage model.[20] In our formulation of the TE-DEA-model the environmental impact works through endogenous handicap functions, which has been shown to be accomplished by the underlying VWBM-model. By weighting the environment indices in the convex combinations used in the BM-model with the value of output we get a model with three desirable characteristics:

- The VWBM-model controls for the volume of output when determining whether a potential dominating convex combination of inputs and outputs from a virtual DMU is allowed as dominating unit in the evaluation of any specific DMU.
- The VWBM-model makes the environmental impact work through an endogenous increase/decrease of produced output by a linear handicap function. The evaluated DMU is, as in standard DEA, allowed to pick the virtual multipliers along with the endogenous part of the handicap function, such that the estimated score is as close to one as possible. Evaluating each DMU separately in this flexible setup implies that a dependence of the handicap function on the chosen input and output mix is reflected in the estimated scores.
- Since environmental impact works through radial modifications of the output vectors, the framework allows for a simultaneous estimation of all scores, where we insist upon a common endogenous radial increase/decrease of the produced output. This approach is the basic

[20] It is not obvious that the impact from the environment in general works independently of the choice of input and output mix. Input or output mix may in some cases have significant impact on how environment contracts or expands output. If this is the case, then using the TE-DEA model will provide biased results. This problem corresponds to maintaining an assumption of a Hick-neutral technical progress in a situation where the movement of the frontier over time is different for different output mixes.

foundation for the proposed TE-DEA-model. Observe that the handicap function in the TE-DEA-model works independently of the choice of input and output mix.

An application of the general models on a data set comparing the education production in a number of OECD countries has been presented. Two outputs, two inputs and one environmental variable have been included in the analysis. It has been illustrated that the inclusion of an environmental variable in TE-DEA does not necessarily imply that the discrimination power decreases. Two BCC-efficient countries operating in very friendly environments turn out to be TE-DEA inefficient. They obtain a score equal to one in the VWBM-model by insisting upon no effect from environment. Hence, comparing the BCC and the TE-DEA-model does not by structure provide any ordering on the obtained scores. BCC-efficiency does *not* imply target efficiency.

References

Banker, R. D., Charnes, A. and Cooper, W. W.: 1984, Some models for estimating technical and scale inefficiencies in data envelopment analysis, *Management Science* **30**(9), 1078–1092.

Banker, R. D. and Morey, R. C.: 1986, Efficiency analysis for exogenously fixed inputs and outputs, *Operations Research* **34**(4), 513–521.

Charnes, A., Cooper, W. W. and Rhodes, E.: 1978, Measuring the efficiency of decision-making units, *European Journal of Operations Research* **2**, 429–444.

Coelli, T., Rao, D. S. P. and Battese, G. E.: 1998, An introduction to efficiency and productivity analysis, *Kluwer Academic Publisher* .

Farrell, M. J.: 1957, The measurement of productive efficiency, *The Journal of the Royal Statistical Society A* **120**, 253–281.

Hall, M. and Winsten, C.: 1959, The ambiguous notion of efficiency., *The Economic Journal* **69**(273), 71–86.

OECD: 2002-2004a, Education at a glance - OECD indicators 2002-2004, *OECD, Paris* .

OECD: 2004b, Learning for tomorrow's world - first results from PISA 2003, *OECD, Paris* .

OECD: 2004c, Problem solving for tomorrows world - first measures of cross-curricular competencies from PISA 2003, *OECD, Paris* .

Paradi, J. C., Vela, S. and Yang, Z.: 2004, Assessing bank and bank branch performance, *in Handbook on Data Envelopment Analysis, eds. Cooper, Seiford and Zhu* pp. 349–399.

Ray, S. C.: 1988, Data envelopment analysis nondiscretionary inputs and efficiency: An alternative interpretation., *Socio-economic Planning Science* **22**, 167–176.

Ray, S. C.: 1991, Resource use efficiency in public schools: A study of conneticut data, *Management Science* **37**(12), 1620–1628.

Ray, S. C.: 2004, Data envelopment analysis: Theory and techniques for economics and operations research, *Cambridge University Press, Cambridge* .

Ruggiero, J.: 1996, On the measurement of technical efficiency in the public sector., *European Journal of Operations Research* pp. 553–565.

Ruggiero, J.: 2004, Performance evaluation in education., *in Handbook on Data Envelopment Analysis, eds Cooper, Seiford and Zhu* pp. 323–348.

Simar, L. and Wilson, P. W.: 2007, Estimating and inference in two-stage semi-parametric models of production processes, *Journal of Econometrics* **136**, 31–64.

Winsten, C. B.: 1957, Discussion on Mr. Farrel's paper., *Journal of the Royal Statistical Society* **120**, 282–284.

Yang, Z. and Paradi, J. C.: 2003, Evaluating competitive banking units using 'handicapped' data envelopment analysis to adjust for system differences, *Workingpaper* .

Country	R-model	VW-BM-model	BM-model	BCC-model
Australia	1.0099	1.1367	1.1939	1.2189
Austria	1.3086	1.4438	1.4450	1.4646
Belgium	1.0522	1.1935	1.2247	1.2835
Czech Republic	1.1502	1.1502	1.1502	1.1502
Denmark	1.0024	1.1280	1.1337	1.1349
Finland	1	1	1	1
France	1.0293	1.1206	1.172	1.2304
Germany	1	1.0208	1.0264	1.0396
Greece	1.1837	1.3852	1.4203	1.5511
Hungary	1.1473	1.2308	1.2348	1.2589
Ireland	1	1	1	1
Italy	1	1.1031	1.1945	1.3053
Japan	1	1	1	1
Korea	1	1	1	1
Mexico	1	1	1	1
Netherlands	1	1.1739	1.1872	1.1948
New Zealand	1	1	1	1
Norway	1	1	1	1
Portugal	1	1	1	1.2803
Slovak Republic	1.2189	1.2189	1.2189	1.2189
Spain	1	1.0121	1.1021	1.1728
Sweden	1	1	1	1
Turkey	1	1	1	1.2252
Brazil	1	1	1	1
Indonesia	1	1	1	1.1936
Thailand	1	1.0344	1.0065	1.3072
Uruguay	1	1	1	1

Table 1. The efficiency scores from this application.

				Common g: g = 0.00425		Reduceed sample regression beta = -0.00312		Full sample regression beta = -0.00084	
	BCC-model (1)	VW-BM-model (2)	DMU specific g (3)	Environment index TE-DEA- model (4)=(1)/(5)	Managerial index TE-DEA-model (5)	Environment index 2SR-DEA (6)	Managerial index 2SR-DEA (7)=(1)/(6)	Environment index 2SR-DEA (8)	Managerial index 2SR-DEA (9)=(1)/(8)
Australia	1.219	1.137	0.00395	1.068	1.142	1.061	1.149	1.057	1.153
Austria	1.465	1.444	0.00430	1.014	1.444	0.993	1.475	1.042	1.406
Belgium	1.284	1.194	0.00387	1.064	1.206	1.049	1.223	1.055	1.217
Czech Republic	1.150	1.150	0	0.974	1.181	0.965	1.192	1.036	1.110
Denmark	1.135	1.128	0.00546	1.005	1.129	0.998	1.137	1.043	1.088
Finland	1	1	0	1	1	0.984	1.016	1.040	0.962
France	1.230	1.121	0.00394	1.098	1.121	1.039	1.185	1.052	1.169
Germany	1.040	1.021	0.00400	1.018	1.022	0.981	1.060	1.039	1.000
Greece	1.551	1.385	0.00434	1.118	1.387	1.066	1.455	1.059	1.465
Hungary	1.259	1.231	0.00409	1.022	1.232	1.004	1.254	1.044	1.205
Ireland	1	1	0	1	1	1.052	0.950	1.055	0.948
Italy	1.305	1.103	0.00480	1.167	1.119	1.099	1.188	1.066	1.224
Japan	1	1	0	1	1	0.954	1.048	1.033	0.968
Korea	1	1	0	1	1	1.006	0.994	1.045	0.957
Mexico	1	1	0	1	1	1.208	0.828	1.092	0.915
Netherlands	1.195	1.174	0.00545	1.016	1.176	1.032	1.158	1.051	1.137
New Zealand	1	1	0	0.974	1.027	1.001	0.999	1.044	0.958
Norway	1	1	0	0.932	1.073	0.964	1.037	1.036	0.966
Portugal	1.280	1	0.00330	1.280	1	1.194	1.072	1.089	1.176
Slovak Republic	1.219	1.219	0	0.972	1.255	0.966	1.262	1.036	1.177
Spain	1.173	1.012	0.00470	1.149	1.021	1.112	1.055	1.069	1.097
Sweden	1	1	0	1	1	0.977	1.023	1.039	0.963
Turkey	1.225	1	0.00340	1.225	1	1.179	1.040	1.085	1.129
Brazil	1	1	0	1	1	1.073	0.932	1.060	0.943
Indonesia	1.194	1	0.00794	1.095	1.090	1.185	1.007	1.087	1.098
Thailand	1.307	1.034	0.00825	1.183	1.105	1.197	1.092	1.090	1.200
Uruguay	1	1	0	1	1	1.145	0.873	1.077	0.928
Average	**1.157**	**1.087**		**1.051**		**1.055**		**1.056**	

Table 2. The efficiency scores from the TE-DEA-model with a common handicap function, from the 2SR-DEA based on the full and a reduced sample

Target and Technical Efficiency in DEA - Controlling for Environmental Characteristics, Extra Online Material.

O. B. Olesen and N. C. Petersen
The University of Southern Denmark

September 29, 2008

Appendix 1: The data set:

	PISA score in reading (in 10^9) (1)	Pisa score others (in 10^9) (2)	School hours per year (in 10^9) (3)	Teachers (in 10^7) (4)	Parental education (5)	Enrollment (6)	Hours per year per student (7)	Teachers per 100 students (8)	Public to total expenditure (9)
Australia	1.3316	1.3340	2.5943	2.0274	61.1	2534273	1023.7	8	84.6
Austria	0.3686	0.3762	0.8056	0.7511	81.9	751125	1072.5	10	96
Belgium	0.5631	0.5788	1.1163	1.1662	64.6	1110707	1005	10.5	94.4
Czech Republic	0.4821	0.5119	0.8556	0.7401	90.5	986834	867	7.5	91.9
Denmark	0.2137	0.2179	0.3733	0.3386	80.5	434063	860	7.8	97.9
Finland	0.2675	0.2691	0.3972	0.3593	84.7	492133	807	7.3	99.3
France	2.9203	3.0235	6.1032	4.7672	67.9	5885441	1037	8.1	93
Germany	4.1208	4.2457	7.4306	5.5352	85.6	8386651	886	6.6	80.8
Greece	0.3525	0.3419	0.7941	0.7538	59.4	746348	1064	10.1	91.6
Hungary	0.4855	0.5019	0.9320	0.8766	78.6	1007601	925	8.7	92.9
Ireland	0.1701	0.1657	0.2957	0.2309	63.7	329905	896.3	7	95.7
Italy	2.1236	2.1156	4.5538	4.3752	49.4	4464498	1020	9.8	97.9
Japan	4.2807	4.6667	7.5197	5.7580	94	8593992	875	6.7	91.6
Korea	2.0275	2.0640	3.2913	1.9360	77.8	3796167	867	5.1	78.5
Mexico	3.7500	3.6728	10.9472	3.0959	15.6	9381408	1166.9	3.3	86.7
Netherlands	0.7150	0.7349	1.4866	0.8500	69.9	1393373	1066.9	6.1	94.8
New Zealand	0.2393	0.2412	0.4371	0.2799	79.6	458811	952.6	6.1	0
Norway	0.1857	0.1820	0.3072	0.3567	90.8	371539	826.8	9.6	99.2
Portugal	0.3887	0.3807	0.7175	0.9359	20	813810	881.7	11.5	99.9
Slovak Republic	0.3130	0.3302	0.5913	0.4937	90.3	667154	886.3	7.4	98.1
Spain	1.5275	1.5411	2.8837	2.7337	45.3	3178670	907.2	8.6	93.1
Sweden	0.4794	0.4735	0.6907	0.6805	86.8	932234	740.9	7.3	99.9
Turkey	1.2170	1.1639	2.3218	1.5731	24.7	2759831	841.3	5.7	0
Brazil	10.4084	9.6181	20.6723	14.2122	57.3	25840339	800	5.5	0
Indonesia	5.3964	5.2618	18.0169	7.7781	22.7	14141985	1274	5.5	76.4
Thailand	2.3450	2.3660	6.5171	3.1273	19	5584504	1167	5.6	97.8
Uruguay	0.1308	0.1277	0.2750	0.2078	35.1	301230	913	6.9	93.5

Table A1. The data set. The two inputs (3) and (4) are instruction hours per year and number of teachers. The two outputs (1) and (2) are enrolled students weighted by the PISA Reading Score and by the average PISA score related to the natural science questionnaires.

	Reading score	Average of 3 scores	Total intended instruction time per year	Size of teaching staff	Parental education attainment	Enrollment
	Units: 10^9	Units: 10^9	Units: 10^9	Units: 10^7		
Mean	1.733467	1.722475	3.81211	2.442237	62.84	3901653
Minimum	0.130779	0.127678	0.2775023	0.207848	15.6	301229.7
Maximum	10.40839	9.618079	20.67227	14.21219	94.0	25840339
Range	10.27761	9.490401	20.39725	14.00434	78.4	25539109

Table A2 – Summary statistics of our data sample. The two inputs are instruction hours per year and number of teachers. The two outputs are enrolled students weighted by the PISA Reading Score and by the average PISA score related to the natural science questionnaires.

Country	Dominating units, VW-BM model	Dominating units, BM model
Australia	(Korea Portugal Turkey)	(Japan Korea Portugal)
Austria	(Korea Portugal Sweden Turkey)	(Korea Portugal Sweden Uruguay)
Belgium	(Korea Portugal Sweden)	(Korea Portugal Sweden)
Czech Republic	(Finland Korea Sweden)	(Finland Korea Sweden)
Denmark	(Finland Ireland Portugal)	(Finland Ireland Norway)
Finland		
France	(Japan Korea Brazil)	(Japan Turkey Brazil)
Germany	(Japan Korea Brazil)	(Japan Mexico Turkey Brazil)
Greece	(Finland Portugal Sweden Turkey)	(Korea Portugal Sweden Uruguay)
Hungary	(Korea Portugal Sweden)	(Korea Portugal Sweden)
Ireland		
Italy	(Korea (.198) Turkey (.691) Brazil (.111))	(Japan (.345) Turkey (.631) Brazil (.024))
Japan		
Korea		
Mexico		
Netherlands	(Ireland Korea Mexico Turkey)	(Korea Mexico New Zealand Uruguay
New Zealand		
Norway		
Portugal		
Slovak Republic	(Finland Korea Sweden)	(Finland Korea Sweden)
Spain	(Korea Turkey Brazil)	(Japan Korea Portugal)
Sweden		
Turkey		
Brazil		
Indonesia		
Thailand	(Mexico (.473) Turkey (.521) Brazil (.006))	(Mexico (.503) Portugal (.464) Brazil (.033))
Uruguay		

Table A3. The dominating DMUs from the analysis in the volume weighted BM- and the BM-model.

Appendix 2: Implementation of the model (a Mixed Integer Linear Programming model).

With more than one output, as illustrated in the application, we have to use the more general model (7) which clearly is nonlinear and not even a convex optimization model. Hence, we can not be sure that solving the model by a non-linear solver will provide a global maximum. To overcome this problem we reformulate (7) as a MILP[1]. To illustrate the general idea we will focus on the situation with only one environmental non-discretionary variable ($p = 1, g \in \mathbb{R}_+$) :

$$
\begin{array}{llcll}
\min & v^t X_{j_0} + v_0 & & & \\
s.t. & u^t Y_j - v^t X_j - v_0 - w^t Y_j \left(Z_j - Z_{j_o} \right) & \leq & 0 \quad \forall j & (1.1) \\
 & u^t Y_{j_0} & = & 10^3 & (1.2) \\
 & w_{kj} - \sum_{j=1}^{50} \widetilde{w}_{kj} & = & 0 \quad k = 1, \ldots, s & (1.3) \\
 & \widetilde{w}_{kj} - b_j M_{kj} & \leq & 0 & (1.4) \\
 & \widetilde{w}_{kj} - (1 - b_j)(-M_{kj}) - u_k \times 2^{(j-40)} & \geq & 0 \quad j = 1, \ldots, 50, \forall k & (1.5) \\
 & \widetilde{w}_{kj} - u_k * 2^{(j-40)} & \leq & 0 & (1.6) \\
 & & & & \\
 & u \in \mathbb{R}_+^s, v \in \mathbb{R}_+^m, g \in \mathbb{R}_+, v_0 \in \mathbb{R} & & \widetilde{w}_{kj} \geq 0, b_j \in \{0, 1\} \, \forall j &
\end{array} \tag{1}
$$

where M_{kj} is a large (but not too large) number, $\forall k, j$. The product of the decision variables $gu^t Y_j$ in (7) is here substituted for $w^t Y_j$. The binary structure implies that

$$w_{kj} = \sum_{j=1}^{50} \widetilde{w}_{kj} \text{ and } \widetilde{w}_{kj} = \begin{cases} 2^{(j-40)} u_k & \text{if } b_j = 1 \\ 0 & \text{if } b_j = 0 \end{cases}$$

Hence, we get an optimal value of g from the MILP as $g^* = \sum_{j=1}^{50} b_j^* 2^{(j-40)}$. The implications of the two possible values of binaries b_j follows from:

		Constraints derived from (1.4-6)	Combined effect
$b_j = 0$	$\Longrightarrow$	$\widetilde{w}_{kj} \leq 0 \wedge \widetilde{w}_{kj} \geq (-M_{kj}) + u_k \times 2^{(j-40)} \wedge \widetilde{w}_{kj} \leq u_k * 2^{(j-40)}$	$\widetilde{w}_{kj} = 0$
$b_j = 1$	$\Longrightarrow$	$\widetilde{w}_{kj} \leq M_{kj} \wedge \widetilde{w}_{kj} \geq u_k \times 2^{(j-40)} \wedge \widetilde{w}_{kj} \leq u_k * 2^{(j-40)}$	$\widetilde{w}_{kj} = u_k * 2^{(j-40)}$

Solving (1) we know for sure that we have an optimal (or near optimal solution). However, computational experience shows that it in certain cases is difficult to get the correct optimal values (and dual values) to (7) by solving (1). Hence, we recommend that (1) is solved to get a near optimal solution, which

[1] This MILP is formulated using a reformulation from Williams, H. P.: "Model Building in Mathematical Programming". 2. edition. Wiley. New York 1989, page 197)

is given as initial solution to a nonlinear solver, which then quickly determines a precise solution satisfying the Kuhn Tucker conditions to (7).