Measuring Eco-inefficiency: A New Frontier Approach

Authors
Chen, Chien-Ming
Delmas, Magali A
Publication Date
2011-01-19

UNIVERSITY OF CALIFORNIA, LOS ANGELES

UCLA Institute of the Environment and Sustainability

Working Paper Series

Measuring Eco-inefficiency: A New Frontier Approach

Authors:

Chien-Ming Chen,

Institute of the Environment and Sustainability, University of California, Los Angeles, CA 90095-1496.

cmchen@ioes.ucla.edu

Magali Delmas

Anderson School of Management and Institute of the Environment and Sustainability, University of California, Los Angeles, CA 90095-1496

delmas@ucla.edu

January 19, 2011

WP# 4

Measuring Eco-inefficiency: A New Frontier Approach

Chien-Ming Chen,
Institute of the Environment and Sustainability, University of California, Los Angeles, CA 90095-1496.
cmchen@ioe.ucla.edu

Magali Delmas
Anderson School of Management and Institute of the Environment and Sustainability, University of California, Los Angeles, CA 90095-1496
delmas@ucla.edu

Abstract

Increasing social concerns over the environmental externalities associated with business activities are pushing firms to identify activities that create economic value with less environmental impact and to become more eco-efficient. Over the past two decades, researchers have increasingly used frontier efficiency models to evaluate productive efficiency in the presence of undesirable outputs, such as greenhouse gas emissions. In this paper, we identify critical flaws of existing frontier models and show that under these models eco-inefficient firms can be identified as eco-efficient. We develop a new eco-inefficiency frontier model that rectifies these problems. Our model allows us to calculate, for each firm, an eco-inefficiency score and improvements in outputs necessary to attain eco-efficiency. We demonstrate, through a Monte-Carlo experiment that our eco-inefficiency model provides a more reliable measurement of corporate eco-inefficiency than the existing frontier models. In the simulation experiment we develop a production function of multiple desirable and undesirable outputs that extends the classical Cob-Douglas function of a single output. The multi-output production function allows for greater flexibility in the simulation analysis of frontier models.

Keywords: Environmental Performance; Eco-Efficiency; Nonparametric Frontier Methodology.

1 INTRODUCTION

Increasing social concerns over the environmental externalities of business activities are pushing managers to devise strategies to mitigate environmental impact (Porter and Reinhardt 2007). Common examples of these strategies include pollution prevention, waste reduction, recycling, closed-loop supply chain management, and environmental management systems (Klassen and McLaughlin 1996; Corbett and Kleindorfer 2001; King and Lenox 2002; Corbett and Klassen 2006; Delmas and Toffel 2008). In all cases, managers are faced with the fundamental question of the impact of these strategies on their corporate performance (King and Lenox 2002; Klassen and Vachon 2003). As a firm typically utilizes multiple input resources to produce outputs, managers need to consider a variety of input and output variables when it comes to corporate performance assessment. Depending on the type of firms, input variables can include labor, capital assets, investments in new product development, and raw materials. Output variables from a production process usually include products, services, or revenue, as well as undesirable by-products such as greenhouse gas emissions and wastes. The potential trade-off relationships among input and output variables make it very challenging for managers to aggregate these variables and present the information as a simple index, which helps facilitate decision-making and identifying room for improvements. In this paper, we develop an eco-inefficiency model capable of aggregating multiple inputs and outputs into an eco-inefficiency score.

Despite the direct relationship between undesirable outputs and other variables in the production process, most prior studies take a single dimensional view on environmental performance assessment. Specifically, researchers often use a ratio between pollution quantities and economic activities as the indicator for environmental performance; for example, a ratio between CO2 quantities and gross domestic product (GDP), electricity consumption, or sales (Cleveland and Ruth 1999; Verbruggen 2009). Although these ratios are easy to understand and interpret, they fail to consider multiple inputs and outputs. In addition, these simple indexes cannot provide reliable information for performance improvement. A firm with a lower index score than its competitors does not necessarily mean the firm has attained better corporate performance, because a lower index score may come at the cost of other input and output variables.

Frontier methodologies can provide a composite inefficiency score that represents the observed firm's distance to the best practice eco-efficiency frontier (Charnes et al. 1978; Kuosmanen and Kortelainen 2005; Färe et al. 1989, 2005). The efficiency frontier represents the collection of firms that produce more desirable outputs with fewer inputs and undesirable outputs than the other firms in the sample. The efficiency frontier also represents the boundary condition that a firm can achieve under the current production technology. Frontier methodologies use a mathematical programming model to extrapolate the efficiency frontier based on the input and output quantities of the sampled firms. A firm's inefficiency score is measured by the improvements in outputs necessary for this firm to reach the extrapolated frontier (i.e., increase desirable output quantities and reduce undesirable output quantities), given the firm's current input level. Therefore we can also identify a benchmark target for according to the value of its inefficiency score.

While several studies have developed frontier models to evaluate eco-inefficiency (e.g., Berg et al. 1992; Färe et al. 1989; Chung et al. 1997; Seiford and Zhu 2002), our analysis shows that the current frontier models may have significant flaws. Specifically, the current frontier models may identify inefficient firms as eco-efficient, and in other cases firms' inefficiency scores may improve with an increased amount of undesirable outputs. These models, however, have been widely used in many subsequent studies. In fact, a bibliographical search in the ISI Journal Citation Report shows that these existing frontier models have received more than 400 citations.[1]

In this paper we identify the cause of the problem of existing models, and build on the nonparametric frontier approach to develop an eco-inefficiency model that overcomes the validity problem of current frontier models. We use a Monte-Carlo simulation experiment to compare the performance of our eco-inefficiency model and four current representative frontier models as recently identified by Hua and Bian (2007), namely the directional distance function (DDF model; Chung et al. 1997), the hyperbolic model (Färe et al.1985), the Seiford and Zhu model (SZ model; Seiford and Zhu 2002), and the "undesirable output as input" (UINP) model (Berg et al. 1992). The simulation results show that our eco-inefficiency model outperforms under different sample sizes and numbers of output variables. From the results we also find that our eco-inefficiency model on average produces a more precise assessment of the inefficiency effect than four other models. This new model has important implications to analyze undesirable outputs beyond the environmental context. Undesirable outputs are present in many operations context. Undesirable outputs include debts or loans, accidents, delays, corporate social irresponsibility, defective products, and waste (e.g., Chen and Delmas 2011; Callens and Tyteca 1999; Park and Weber 2006; Pathomsiri et al. 2008). Our model can therefore provide a useful tool to evaluate operational efficiencies in these contexts.

Our paper makes an additional contribution to the operations research literature by extending the current frontier models simulation methodology, which only allows for simulation of a single desirable output, to a general framework capable of simulating multiple desirable and undesirable outputs. Our simulation framework allows for greater flexibility in the analysis of frontier models for measuring eco-inefficiency.

In the next section we introduce the frontier methodology and the four current frontier models. We illustrate the shortcomings of these four models by using the production data of 30 paper mills in the U.S. In Section 3 we present our eco-efficiency model and demonstrate its advantages. In Section 4, we use a Monte-Carlo experiment to compare the performance of our and the other frontier models. In the Section 5 we summarize our findings and contributions.

2 FRONTIER METHODOLOGY AND EXISTING MODELS

We begin this section by introducing the formulation of the efficiency frontier. Next we review four frontier models that have been developed to deal with undesirable outputs.

2.1 Fundamental concepts of frontier methodologies

The nonparametric frontier methodology, also known as data envelopment analysis (DEA), has been extensively used in the operation literature to evaluate firms according to their multiple inputs and outputs (Charnes et al. 1978; Banker et al. 1984). The frontier methodology uses linear programming to convert multiple inputs and outputs of firms into a relative efficiency score. In the linear programming model, a piecewise linear industry best practice frontier is constructed using the observations in the sample. The set of feasible production plans, or technology set, are the input-output combinations enveloped by the frontier. If a firm is on this frontier, it is considered efficient. If a firm is not on the frontier, the distance to the best practice frontier represents the firm's inefficiency.

We now describe the efficient frontier model in a linear programming form. In the model, we consider three vectors. The inputs $X = (x_1,..., x_M)$ are the resources used to produce the desirable outputs $Y = (y_1,..., y_N)$ and undesirable outputs $U = (u_1,..., u_P)$. Given that we observe k firms in our sample, the production technology set can be formulated as follows (Charnes et al. 1978; Färe and Grosskopf 2004):

$$\Omega := \{(X, Y, U): X \text{ can produce } Y \text{ and } U\} \tag{1-1}$$

$$= \{(X, Y, U): \sum_{k=1}^{K} z_k x_{km} \le x_m, \text{ for } m = 1,..., M\} \tag{1-2}$$

$$\sum_{k=1}^{K} z_k y_{kn} \ge y_n, \text{ for } n = 1,..., N \tag{1-3}$$

$$\sum_{k=1}^{K} z_k u_{kp} = u_p, \text{ for } p = 1,..., P \tag{1-4}$$

$$z_k \ge 0, \text{ for k} = 1,..., \text{K}\} \tag{1-5}$$

where $(x_{k1},..., x_{kM})$, $(d_{k1},..., d_{kN})$ and $(u_{k1},..., u_{kN})$ are the input and output vectors of the *k*th firm in the sample, and z_k is the intensity variable associated with the *k*th firm. The z_k variable indicates the importance of the *k*th firm in constructing the efficient frontier for a specific point (X, D, U) in the production set.

The constraints (1-2) to (1-5) form a polyhedron also referred as the production set. The production set is the collection of feasible inputs and outputs (X, D, U) under the current production technology. The production set is similar to the feasible region in linear programming. Points in the production set are those achievable (and therefore feasible) under the current technology constraint. The production set as defined in (1) has the following important properties (Färe and Grosskopf 2004):

Property 1: $(X, Y, U) \in \Omega$ and $X_1 \ge X$ implies $(X_1, Y, U) \in \Omega$

Property 2: $(X, Y, U) \in \Omega$, then $Y_1 \le Y$ implies $(X, Y_1, U) \in \Omega$

Property 3: $(X, Y, U) \in \Omega$ implies $(X, \theta Y, \theta U) \in \Omega$ for $0 \le \theta \le 1$

A key presumption underlying these three properties is that, if (X, Y, U) is observed, and then it is by definition a member of the production set. In other words, all firms in the sample are deemed achievable production combinations under the current production technology. By these three properties, the frontier methodologies extrapolate the entire production set based on the input-output observations in the sample. The first two properties mean that, if (X, Y, U) is observed, then any input-output vector with higher inputs (i.e., (X_1, Y, U)) or fewer desirable outputs (i.e., (X, Y_1, U)) is a member of the production set. These two properties are called the strong disposable assumption, because inputs or outputs can change unilaterally without compromising each other. The third property, also called the weak disposable assumption, indicates that if we reduce the undesirable outputs of (X, Y, U), its desirable outputs decrease in tandem. This property is associated with the equality constraint (1-5), and can be contrasted with the strong disposability in the first two properties. The weak disposability property only applies to undesirable outputs, because we assume that producers cannot dispose freely of the undesirable outputs, so producers need to divert either their inputs or desirable outputs to cover the costs. For example, electric utility plants may install carbon capture devices to reduce their greenhouse gas emissions (Gibbins and Chalmers 2008).

If undesirable outputs can be generated without subsequent costs or damage, undesirable outputs are said to be strongly disposable and the production set is the same as (1) except we replace (1-5) with (2):

$$\sum_{k=1}^{K} z_k u_{kp} \leq u_p, \text{ for } p = 1,\ldots,P \qquad (2)$$

In Figure 1 we show a production set with a desirable output *y* and an undesirable output *u* in order to illustrate how to compute an inefficiency score. The horizontal axis represents the undesirable output u and the vertical axis represents the desirable output y. We divide the output quantity of each firm by its input quantity in order to evaluate firms' eco-inefficiency based on *y* and *u*. Firms with a high eco-efficiency are those situated in the upper-left corner of the graph, because firms located in that area produce more desirable outputs with low undesirable outputs. We use piecewise linear segments to extrapolate the eco-efficient frontier by linking firms in the upper-left corner. Firms on the frontier are considered eco-efficient because no other firms in the production set can produce more desirable outputs and fewer undesirable outputs. In Figure 1, the frontier is the line segment *oabcd* if we assume that undesirable output is weakly disposable, and is *oabce* if assume that undesirable output is strongly disposable. It is also important to note that the *cd* portion of the efficient frontier is dominated by the point *c*; i.e., *c* has higher output *y* and fewer undesirable output *u*. We call the *cd* portion of the frontier the *misspecified efficient frontier*. As we explained earlier, in the strong disposability assumption, undesirable outputs are free, and therefore firms do not need to allocate resources to compensate for the emissions of undesirable outputs. The difference in disposability assumption is characterized by the inequality signs for the undesirable output constraints (1-4). As a result, the production set associated with the strong disposability assumption is larger than that under the weak disposability assumption (see Figure 1).

***

[Insert Figure 1 about here]

***

In a frontier model, the inefficiency score of a firm represents the firm's distance to the efficient frontier. We can calculate the inefficiency score as an optimization problem, in which the efficient frontier is the boundary of the feasible region and the inefficiency index is the objective function. Therefore, the inefficiency index determines both the direction of the evaluated firm toward the frontier, as well as how the distance between the firm and the frontier is calculated. As shown in Figure 1, different frontier models may adopt different assumptions on the production set (e.g., weak disposability or strong disposability assumption) and different inefficiency indexes. The inefficiency score and the benchmark target therefore depend on these two settings. For example, firm *f* in Figure 1 is eco-inefficient because it is not on the efficient frontier, whereas firms *a*, *b* and *c* are eco-efficient. However, firm *f* can move in different directions to reach the frontier. Next we introduce four frontier models for undesirable outputs.

2.2 Current frontier models for undesirable outputs

In this section we give a brief overview of four representative frontier models: the directional distance function (DDF model; Chung et al. 1997), the hyperbolic model (Färe et al.1989), the Seiford and Zhu model (SZ model; Seiford and Zhu 2002), and the “undesirable output as input” (UINP) model (Berg et al. 1992). We first introduce the UINP and SZ models, which both assume strong disposability on undesirable outputs and treat undesirable outputs as variables to be minimized in the formulation. These two models use the traditional DEA model mathematical formation. Second we introduce the DDF and hyperbolic model, which utilize a weak disposability assumption on undesirable outputs. This second set of models has been used widely in various industry contexts, including banks, electricity industries, industry efficiency, provincial governments, agriculture, and airports (e.g., Lee et al. 2002; Picazo-Tadeo et al. 2005; Park and Webber 2006; Watanabe and Tanaka 2007; Pathomsiri et al. 2008; Cuesta and Zofio 2007; Zofio and Prieto 2001).

The UINP approach simply treats undesirable outputs as inputs, as firms are expected to minimize their input consumption (Berg et al. 1992). Therefore in implementation the UINP model is identical to the traditional DEA model; namely,

$$\max\{\theta_{UINP} \mid (X, \theta_{UINP} Y, U) \in \Omega_{UINP}\}, \tag{3}$$

where Ω_{UINP} is constructed by replacing the (1-4) of Ω with $\sum_{k=1}^{K} z_k u_{kp} \le u_p$, for $p = 1, \dots, P$. The inefficiency score represents the extent a firm can scale up its desirable outputs, given its current inputs and undesirable outputs. For this reason, the UINP model has been criticized for not being representative of the production process, because undesirable outputs are modeled as inputs (Seiford and Zhu 2002). The inefficiency score θ is associated only with the desirable outputs Y, and therefore from this score, we cannot calculate the efficient level of undesirable output (we can only calculate the efficient level of desirable output; see Figure 2 for an illustration).

Seiford and Zhu (2002) take a more heuristic approach to undesirable outputs. The SZ model substitutes undesirable output variables by auxiliary output variables. These new variables are computed by adding a positive scalar to the original undesirable outputs after multiplying them by minus one. The SZ model deals with undesirable outputs by transforming undesirable output variables as in (4).

$$\tilde{U} = -U + W, \qquad (4)$$

where W is a predetermined vector making the new undesirable vector $\tilde{U}$ positive for all firms. Next the new undesirable vector $(X, D, \tilde{U})$ is used to construct the production set $\tilde{\Omega}$ under the strong disposability assumption (see (2)). Thus maximizing these new output variables is equivalent to reducing the underlying undesirable outputs. The inefficiency score of the SZ model is obtained from (5)

$$\max\{\theta_{SZ} \mid (X, \theta_{SZ} Y, \theta\tilde{U}) \in \tilde{\Omega}\}, \qquad (5)$$

Therefore by maximizing the objective function θ_{SZ} in (5), we are actually scaling up Y and scaling down U at the same time (see Figure 2). We should also note that, the inefficiency score θ of (5) depends on the choice of translation vector W.

The DDF and hyperbolic models use a weak disposability assumption. They have identical production sets but differ in their inefficiency indexes, which are illustrated in Figure 2. In the DDF model, firms follow a predetermined direction (g^Y, g^U) towards the efficient frontier; the inefficiency score θ_{DDF} is the optimal value of problem (6):

$$\max\{\theta_{DDF} \mid (X, Y + \theta_{DDF} g^{Y}, U - \theta_{DDF} g^{U}) \in \Omega\}, \tag{6}$$

In the DDF model, we can designate the directional vector with price information or using preferences for outputs. While there are a number of commonly used directional vectors (see Färe et al. (2008), p.553), the literature does not have a clear guideline or rule about choosing a directional vector. More importantly, the DDF inefficiency scores can vary with different directional vectors. As noted in Färe and Grosskopf (2004), "…clearly [the directional] efficiency depends on the choice of the directional vector (p.9)…However, we do not have a general rule for determining those vectors (p.10)."

In the hyperbolic model, the inefficiency of a firm is measured by expanding the firm's desirable outputs and contracting undesirable outputs. The inefficiency score θ_{hyper} is obtained from (7):

$$\max\{\theta_{hyper} \mid (X, \theta_{hyper} Y, U / \theta_{hyper}) \in \Omega\}, \tag{7}$$

Therefore the locus of projecting a firm to the efficient frontier will be hyperbolical (see Figure 2). Note that the hyperbolic model is a nonlinear and non-convex optimization problem, and therefore the model is difficult to solve, especially for a large sample.

***

[Insert Figure 2. about here]

***

The modeling assumptions and ranges of efficiency scores of these four models are summarized in Table 1. Note that in all four models, the efficiency status is achieved when a firm obtains the lower-bound value (i.e., one or zero), which means that further expansion of desirable outputs and reduction of undesirable outputs is impossible.

***

[Insert Table 1. about here]

***

In the next section, we use data from paper mill production to test these four models. We illustrate that these models not only fail to capture actual fluctuations in undesirable outputs, but also tend to produce misleading efficiency measurement results.

2.3 Illustrative examples: Assessing the eco-inefficiency of paper mills

The data used in this section consist of the empirical inputs and outputs of 30 paper mills operating in the U.S. in 1976. This data set also appears in Färe et al. (1989) and Seiford and Zhu (2002). We generate inefficiency scores using four inputs (fiber, energy, capital and labor), one desirable output (paper) and four undesirable outputs (biochemical oxygen demand, total suspended solids, particulates and sulfur oxides) (Färe et al., 1989; Seiford and Zhu, 2002). Based on the paper mill data, we construct two scenarios to test the four models covered in the previous section. Our purpose is to verify how sensitive the models are in detecting increases in undesirable outputs. In the first scenario, we use the original input output data. In the second scenario, we double the undesirable outputs of *the evaluated firm*, while all other data remain unchanged. For example, when firm *a* is evaluated in the second scenario, we will double firm *a*'s all undesirable outputs (BOD, total suspended solids, particulates and Sox) to be $2u_{ap}$ for $p = 1,...,P$. Intuitively we are expecting that firms' efficiency does not increase when emissions increase. That is, firms should become more inefficient when their emissions are doubled, which represents a massive surge in emissions. So if in the course of our experiment, we find problems in these models, then further experimentation with greater increases in undesirable can be safely omitted.

***

[Insert Table 2 and 3 about here]

***

Table 2 provides descriptive statistics and the experimental results are shown in Table 3. Now we examine the results from the DDF model (columns 2 to 3) and the hyperbolic model (columns 4 to 5). Compared with other models, the DDF and hyperbolic models have a higher proportion of efficient mills in both scenarios. More importantly, some mills become more efficient with increases in undesirable outputs (*after doubling their undesirable outputs)*. For example mill 14 becomes more efficient in scenario 2. In the case of DDF, 10% of the mills become more efficient and with the hyperbolic model 23% become more efficient. Therefore these two models are not very responsive to increases in undesirable outputs.

While the UINP and SZ scores don't show decreases in inefficiency scores, SZ scores do not vary with changes in undesirable outputs. This means that the score is insensitive to changes in undesirable outputs. Around 70% of the mills receive the same inefficiency scores in scenario 2. For 30% of the mills with changes, the average inefficiency increase in scenarios 2 is less than 0.2%. This result warrants further investigation about the sensitivity of the SZ model regarding changes in undesirable outputs. Our simulation will confirm these irregularities.

In the paper mill example, we find that the DDF and hyperbolic inefficiency scores can become lower, so firms appear to be more "efficient" in the presence of increased undesirable outputs, while the UINP and SZ models do not face the same issue.

3 MATHEMATICAL FORMULATIONS

3.1 Eco-inefficiency model

We argue that the weak disposability assumption in the DDF and hyperbolic models might create a problem of decreased inefficiency with increased undesirable outputs. First, we will show graphically the intuition behind our model. Second we will present the mathematical formulations of our model.

We can illustrate the problem with previous models in Figure 1. For example, firm f obtains an inefficiency score of θ under the weak disposability assumption. Then we increase firm *f*'s undesirable outputs to *f**. As shown in Figure 1, the inefficiency score becomes θ^*, which is closer to the efficiency frontier under the weak disposability assumption. Under strong disposability assumption the efficient frontier is 'oabce' while under the weak disposability assumption it is 'oabcd'. Clearly θ is larger than θ^*, and hence, under the weak disposability assumption, firm *f** appears to be more efficient than *f*. If firm f increases its undesirable output further, it can overtake firm *d* and becomes efficient.

The reverse situation is similarly problematic: if a firm manages to cut its undesirable output from the position of f* to f, it will be considered less efficient in the model. We can attribute this problem to the characteristics of the pre-determined directional vector or hyperbolic curve of the conventional efficiency measure.

Our model overcomes this problem by allowing firms to select their own directions for improvement to reach the efficiency frontier. Our eco-inefficiency model is presented below:

$$E(x,d,u) = \max \frac{1}{N+P}\left(\sum_{n=1}^{N} \tilde{g}_n^y \Big/ y_{1n} + \sum_{p=1}^{P} \tilde{g}_p^u \Big/ u_{1p} \right) \tag{8-1}$$

$$s.t. \sum_{k=1}^{\mathrm{K}} \mathrm{z}_{\mathrm{k}} x_{km} \leq x_{1m}, m = 1,\ldots, M \tag{8-2}$$

$$\sum_{k=1}^{\mathrm{K}} \mathrm{z}_{\mathrm{k}} y_{kn} \geq y_{1n} + \tilde{g}_n^y, n = 1,\ldots, N \tag{8-3}$$

$$\sum_{k=1}^{\mathrm{K}} \mathrm{z}_{\mathrm{k}} u_{kp} = u_{1p} - \tilde{g}_p^u, p = 1,\ldots, P \tag{8-4}$$

$$z_k \geq 0, \tilde{g}_n^y \geq 0, \tilde{g}_p^u \geq 0, \text{for all } k, n \text{ and } p \tag{8-5}$$

The eco-inefficiency model uses the additive inefficiency index similar to the DDF model ((8-3) and (8-4)). The additive inefficiency index can be contrasted with the radial inefficiency index in the UINP and SZ models, which assume evaluated firm should reach the efficiency frontier by proportionally changing its undesirable and desirable outputs. We should note that in practice there is no guarantee that firms would always improve their efficiency by decreasing undesirable outputs and increasing desirable outputs proportionally. Thus it would be unrealistic to make this assumption. Another benefit of formulation (8) is that the benchmark target for each firm must be efficient, while the radial inefficiency measure could identify dominated points as benchmark targets (Tone 2001). We also choose to maximize the objective function in order to assure that the evaluated firm is benchmarked with an efficient firm on the frontier. The variables $\tilde{g}_n^y$ and $\tilde{g}_p^u$ in model (8-1) represent the amount of output improvements that the evaluated firm can make to reach its benchmark target on the efficiency frontier. Correspondingly, the objective function is the average magnitude of these improvements. For example, a score of 0.5 would mean that the firm can increase its desirable outputs by 50% and reduce undesirable outputs by 50%.

The objective value of equation (8-1) represents the overall degree of output efficiency. It is calculated as the average amount of potential output improvement divided by the observed output value, y_{1n} and u_{1p} in equation (8-1). The index value ranges from zero to infinity. A zero value means that the evaluated firm is on the efficiency frontier and has no slack values (hence the firm is *efficient*). If a firm's score is positive, the larger the value, the more inefficient the firm is. The constraints of this problem are similar to those of the DDF model. Therefore we also assume that undesirable outputs are weakly disposable, namely the reduction of undesirable outputs is not free, and will entail some loss of desirable outputs.

***

[Insert Figure 3 about here]

***

We illustrate our model in Figure 3, where we also consider one desirable and one undesirable output. In our model, instead of using a fixed direction to reach the frontier, the evaluated firm (u, y) is free to choose an improvement direction that maximizes its potential for improvement and therefore its potential efficiency (8-1). We will show that this flexibility in choosing an improvement direction helps avoid problems associated with the weak disposability assumption on undesirable outputs.

The eco-inefficiency score provides an aggregate measure of a firm's relative efficiency compared to other firms in the sample. After solving the eco-efficiency model, however, we can also identify the efficiency target that the evaluated firm can emulate. Specifically, the benchmark target for firm *k* can be obtained as:

$$(x_{km}, y_{kn} + \tilde{g}_n^{y*}, u_{kp} - \tilde{g}_p^{u*}) \text{ for all } m, n \text{ and } p \tag{9}$$

where $(\tilde{g}_n^{y*}, \tilde{g}_p^{u*})$ is the optimal solution to model (7).

3.2 Properties of the eco-inefficiency model

In this section we will show some important properties of the model. Proofs of these results are provided in Appendix-A. Theorem 1 shows that our eco-efficiency model is unit-invariant in inputs and all outputs:

Theorem 1. $E(x_{km}, y_{kn}, u_{kp})$ *is homogeneous of degree zero in* x_{km}, y_{kn} **, and** u_{kp} *; i.e., if we replace the original data* (x_{km}, y_{kn}, u_{kp}) *by* $(\alpha x_{km}, \beta y_{kn}, \gamma u_{kp})$ *for all k, where* α , β **, and** γ *are arbitrary positive numbers, we still have* $E(\alpha x_{km}, \beta y_{kn}, \gamma u_{kp}) = E(x_{km}, y_{kn}, u_{kp})$ *for all k.*

The homogeneity (or units invariance) property is useful because it facilitates comparisons of efficiency across different systems without worrying about the measurement units of inputs and outputs. The "unit-less" property of efficiency scores has also long been recognized as important in engineering and science (see discussions and examples in Chapter 1 of Charnes et al. 2007). Without the homogeneous property, the inefficiency scores would depend on the unit of measurement (e.g., in pounds, kg, or tons; or in Euros or dollars). This would make the interpretation and comparison of the scores more difficult. Traditional DEA models, where all outputs are desirable outputs, are endowed with the homogeneous property (Charnes et al. 2007). We can easily verify that the hyperbolic, SZ, and the UINP models also possess the homogeneous property. This means that proportional changes on both sides of the constraints in (1) can cancel each other. Note that these three models all have a multiplicative type of inefficiency indexes. The DDF model, however, does not have the homogeneous property, even when all outputs are desirable. This limitation relates to our earlier discussion that the DDF score can be influenced by the choice of directional vectors.

Another important property that needs to be carefully verified is the quality of the eco-efficiency measure. Ideally, we would expect that eco-efficient firms, as identified by the model, should be "at least as good as" any members in the technology set. Conversely, firms will be regarded as inefficient only when they have an eco-performance inferior to any feasible units in the technology set. To answer this question, we need to first define the dominance relationship in the technology set.

Definition 1 (Domination relationship). The production plan $(x_{km}, y_{kn}, u_{kp}) \in \Omega$ *is non-dominated if there does not exist any* $(x_{km}, y'_{kn}, u'_{kp}) \in \Omega$ *such that* $(x_{km}, y'_{kn}, u'_{kp}) \neq (x_{km}, y_{kn}, u_{kp})$ *while* $y'_{kn} \geq y_{kn}$ *and* $u'_{kp} \leq u_{kp}$ *Otherwise* (x_{km}, y_{kn}, u_{kp}) *is dominated.*

The next theorem shows that the eco-efficiency status is equivalent to the non-dominance status in the technology set.

Theorem 2. $E(x_{km}, y_{kn}, u_{kp}) = 0$ *if and only if* (x_{km}, y_{kn}, u_{kp}) *is non-dominated in* Ω.

Theorem 2 also implies that our eco-inefficiency model will also identify non-dominated benchmark target points. Graphically, it means that the eco-inefficiency model will always locate points on the efficiency frontier as benchmark target points (see Figure 3). Algebraically, the theorem implies that the constraints on undesirable outputs (8-4) are always binding, and therefore the type of disposability assumptions on undesirable outputs (i.e., (8-4)) is not going to influence our eco-inefficiency score.

Theorem 2 provides a convenient way to check whether a firm has been misclassified as an efficient firm in the DDF and hyperbolic models:

Corollary 1. For a firm is efficient in the DDF or hyperbolic model ($\theta_{DDF} = 0$ *or* $\theta_{hyper} = 1$*) but inefficiency in the eco-inefficiency model (i.e.,* $E(x_{km}, y_{kn}, u_{kp}) > 0$*), then the firm is dominated in* Ω *.*

Corollary 1 applies to firms located on the misspecified efficient frontier due to the weak disposability assumption (see Figure 1 the ‘cd’ line). These firms are dominated points in the production set, but in the DDF and hyperbolic models these firms may be identified as efficient (see Figures 1 and 2). If a firm appears to be efficient in these two models but inefficient in the eco-inefficiency model, this firm must be dominated (therefore inefficient) in the production set. From our earlier application to the paper mill production data, firms whose efficient targets are on the misspecified efficient frontier in the DDF and hyperbolic models can obtain misleading inefficiency scores (see Figure 1). We can similarly verify whether a firm has the above problem by calculating their efficient targets under these two models. Then we can apply Corollary 1 and see if the firm's eco-inefficiency score is equal to zero.

4 MONTE-CARLO EXPERIMENT

The paper mill example presented earlier offers some initial evidence about the drawbacks of the current frontier approaches for eco-efficiency. To further confirm these limitations, we employ a Monte-Carlo experiment and generate random samples to compare our model with the other four frontier models. We begin by describing the production function used in the simulation.

4.1 Production function

In the production economics literature, researchers have typically utilized the Cobb–Douglas production function to generate the input and output samples because of its flexibility and simplicity (e.g., Grosskopf 1996; Zhang and Bartels 1998; Coelli 2005; Banker and Natarajan 2008; Banker et al. 2010; Kuosmanen and Johnson 2010). We follow Banker and Natarajan (2008) and Banker et al. (2010) and use the following production function with one input:

$$\log y = f(x) + \upsilon - \mu = \log(a_0 + a_1 x + a_2 x^2 + a_3 x^3) + \upsilon - \mu \qquad (10)$$

In equation (10), the output quantity (y) is the sum of the Cobb-Douglas function f(x), a random noise term (υ), less the inefficiencies (μ) in the production process. The Cobb-Douglas function $f(x)$ comprises a linear cubic function of the input variable x. This function corresponds to the maximal output quantity that is technically achievable by using x. Then the function $f(x)$ forms the efficient frontier that we use to benchmark firm performance. The term υ stands for sampling errors as commonly seen in most econometric models, and μ represents the productive inefficiency effect. The random variable υ is typically assumed to follow a standard normal distribution, while μ is assumed to follow a one-sided distribution such as a half normal distribution and is non-negative (Coelli et al. 2005). We illustrate the production function in Figure 4. In Figure 4 we plot a hypothetical Cobb–Douglas production function that has one input x and output y. Observed input-output quantities are represented by asterisks that are on both the upper and lower side of the production frontier. The deviation from the production function (e.g., y*-y0) results from the mixed influence from the noise and inefficiency terms (i.e., exp($\upsilon - \mu$)).

***

[Insert Figure 4 about here]

***

The production function (10) leads to a single output. However, the evaluation of eco-efficiency requires the consideration of multiple outputs and therefore we need a model for both desirable and undesirable outputs. One approach used in prior studies is to model undesirable outputs as inputs in the production function (Koop 1998). This approach is akin to the UINP model and therefore is endowed with similar limitations (see Table 1). To avoid these potential limitations, Fernández et al. (2002) use two production functions to estimate the technical and environmental efficiencies separately (i.e., the production of desirable and undesirable outputs, respectively). The production function of desirable outputs depends on inputs only and the production function of undesirable outputs depends on desirable outputs. This assumption, however, can be overly strong in many situations, because we can expect that technical and environmental efficiencies of a firm should be correlated.

In this paper we develop a framework of multiple desirable and undesirable outputs based on the concept of Fernández et al. (2002). However, we take a different approach and model the technical and environmental efficiencies as two correlated random variables. Specifically, we generalize the single output function (10) to a multiple output production function $F(x)$ of N desirable outputs $(y_1,...,y_N)$ and P undesirable outputs $(u_1,...,u_p)$ as:

$$\begin{pmatrix} \log y_1 \\ \log y_2 \\ \vdots \\ \log y_N \\ \log u_1 \\ \log u_2 \\ \vdots \\ \log u_P \end{pmatrix} = F(x) + \vec{\upsilon} - \vec{\mu} = \log Ax + \begin{pmatrix} \upsilon \\ \upsilon \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \upsilon \end{pmatrix} - \begin{pmatrix} \mu_y \\ \mu_y \\ \vdots \\ \vdots \\ -\mu_u \\ \vdots \\ \vdots \\ -\mu_u \end{pmatrix} \tag{11}$$

where A denotes the coefficient matrix with each row having a similar structure to $f(x)$ in (10).

As in the univariate production function, the random noise υ (11) has a standard normal distribution across different outputs. For the inefficiency effect, we distinguish between the *productive inefficiency* μ_y and the *environmental inefficiency* μ_u. The negative sign of the environmental inefficiency terms indicates that environmental inefficiency will cause firms to produce more undesirable outputs. Figure 5 illustrates the relationship between inputs, outputs, and the two inefficiency terms.

***

[Insert Figure 5 about here]

***

Specifically, μ_y and μ_u are the inefficiency effects associated with the production of desirable and undesirable outputs, respectively. The variable μ_y, the *productive inefficiency* term, is nonnegative and can reduce the desirable output quantities in $F(X)$. On the other hand, μ_u, the *environmental inefficiency* term, has the effect of increasing undesirable outputs quantities from the efficient level in $F(X)$. In this paper we assume that μ_y and μ_u are correlated, but are not likely to be perfectly correlated. Prior studies have found differences in firms corporate profitability and environmental performance and there is a debate in the literature on whether improved environmental performance could lead to financial gains (e.g., Klassen and McLaughlin 1996; King and Lenox 2002).

Based on the assumption made for the conventional production function (10), we assume the production function (13) has a bi-variate half normal distribution: $(\mu_y, \mu_u) \sim | N_2(0, \Sigma) |$, Σ is a semi-positive definite variance covariance matrix. The joint distribution function of variance covariance matrix (μ_y, μ_u) is (Johnson et al. 2002, pp.326-327):

$$p(\mu_y, \mu_u) = \frac{1}{\pi \mathrm{S}_{\mu_y} \mathrm{S}_{\mu_u} \sqrt{1-\rho^2}} \exp\left(\frac{-(\mu_y / S_{\mu_y})^2 - (\mu_u / S_{\mu_u})^2}{2(1-\rho^2)} \right) \cosh\left(\frac{\rho \mu_y \mu_u}{(1-\rho^2)\mu_y \mu_u} \right) \quad (12)$$

Note that the marginal distribution (μ_y or μ_u) is a half-normal distribution, which matches the distributional assumption made in the univariate production function (10).

The variance covariance matrix Σ can be written as a function of the standard deviations of μ_y and μ_u and the correlation coefficient between of μ_y and μ_u (ρ) as (Gut 2009, p.126):

$$\Sigma = \begin{bmatrix} \mathrm{S}_{\mu_y}^2 & \rho * \mathrm{S}_{\mu_y} \mathrm{S}_{\mu_u} \\ \rho * \mathrm{S}_{\mu_y} \mathrm{S}_{\mu_u} & \mathrm{S}_{\mu_u}^2 \end{bmatrix} \quad (13)$$

This covariance structure allows us to vary the correlation between the productive and environmental efficiency terms by assigning different values to ρ in a simulation experiment. Based on Equation (11), the production function used in our experiment is given by:

$$
\begin{pmatrix} \log y_1 \\ \log y_2 \\ \vdots \\ \log y_N \\ \log u_1 \\ \log u_2 \\ \vdots \\ \log u_P \end{pmatrix} = \begin{bmatrix} \log(-37+48x-12x^2+x^3) \\ \vdots \\ \vdots \\ \log(-37+48x-12x^2+x^3) \end{bmatrix} + \begin{pmatrix} \upsilon \\ \upsilon \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \upsilon \end{pmatrix} - \begin{pmatrix} \mu_y \\ \mu_y \\ \vdots \\ \vdots \\ -\mu_u \\ \vdots \\ \vdots \\ -\mu_u \end{pmatrix} \tag{14}
$$

In equation (14), we assign the coefficients of x in $F(x)$ according to Banker and Natarajan (2008), such that the production function has the desirable properties of being "continuous, monotone increasing, and concave" in the range of [1, 4]. To simplify the experimental setup, we let all outputs share the same coefficient values in the production function (14), but the output values (i.e., Y_N and u_P) are contingent on the noise terms associated with each output (υ), as well as the productive or environmental inefficiency effects (i.e., μ_Y and μ_u). Once we specify the input and the two stochastic terms in (14), we can conveniently calculate the output vector on the left-hand-side of (14).

4.2 Evaluation criteria

With the simulated input and output data, we can apply the five frontier models discussed in Section 2 and obtain inefficiency scores corresponding to the models. Comparing the inefficiency scores with the inefficiency variables in the simulation can indicate the performance of these frontier models. In this section we introduce two criteria, namely validity and consistency that we will use to evaluate the performance of the five frontier models.

4.2.1 Correlation criterion

The validity of the frontier models hinges on how well the inefficiency scores correspond to the true inefficiency status of firms. To measure the validity of frontier models, we calculate the rank correlation between the inefficiency scores and the simulated inefficiency terms, which we operationalize as the inefficiency effect that frontier models are supposed to detect. We calculate rank correlation because inefficiency scores obtained from different frontier models may have their specific inefficiency indexes (see Table 1), and therefore rank correlation provides a more consistent assessment.

In other words, we expect that the rankings that we derive from the inefficiency scores correlate highly with the “real” rankings, which we generate through simulation. Regarding the choice of correlation measures, we use the Kendal’s tau (τ) rank correlation coefficient. Kendall's tau (τ) measures the degree of agreement between the generated and measured efficiency rankings. The tau (τ) coefficient ranges from -1 and 1, where “1” means a perfect match between two ranking distributions, and “-1” conversely suggests that one ranking distribution is the opposite of the other. See Kendall and Gibbons (1990) for an in-depth exposition of the Kendall's tau (τ) statistic.

In the production function (13), eco-inefficiency consists of productive inefficiency μ_y and environmental inefficiency μ_u (see Figure 4). We use the average of these two inefficiency terms as the proxy of simulated eco-inefficiency ($\mu_{avg} = (\mu_y + \mu_u)/2$). Then we calculate the rank correlation coefficient between μ_{avg} and the inefficiency score θ obtained from a frontier model:

$$\tau = corr_{Kendall}(\mu_{avg}, \theta) \qquad (15)$$

4.2.2 Error rate criterion

In the paper mill example, we observed that some mills’ inefficiency scores decreased after we doubled their undesirable outputs. This is a clear indication of the problems raised by current frontier models, since the inefficiency score should be non-decreasing as the firm produces more pollution.

To measure the degree of inconsistency of frontier models, we record the number of times that the inefficiency score decreases (therefore the firm appears to be *less* eco-inefficient) after we experimentally double all undesirable outputs of the evaluated firm. When a firm's inefficiency score decreases in this situation, we call the above situation an instance of *an error*. More specifically, we define the error rate for a frontier model as:

$$\delta = \sum_{k=1}^{K} \tilde{\theta}_k \Big/ K \text{, where } \tilde{\theta}_k = 1 \text{ if } \theta_k - \theta_k^* > 0 \text{ and } \tilde{\theta}_k = 0 \text{ otherwise.} \qquad (16)$$

In (16), θ_k stands for the inefficiency score of firm *k* obtained using the original data, while θ_k^* is the inefficiency score that we obtain from the same frontier model, but computed with the firm *k*'s undesirable outputs doubled. Therefore $\tilde{\theta}_k$ is equal to one when the firm *k* provides an instance of an error, as we define earlier. In other words, δ calculates the likelihood that an instance of an error will occur for each firm in the sample.

4.3 Parameters

Table 4 lists the simulation parameters in the experiment. We generate an input variable from a continuous uniform distribution between 1 and 4 to obtain the desirable properties for the production function. We implement three different sample sizes in the experiment (25, 50, and 100). These correspond to small, medium, and large sample sizes in the applications of frontier models (Banker et al. 1993; Zhang and Bartels 1998). We also vary the number of outputs to verify whether the output dimensionality impacts the performance of frontier models.

***

[Insert Table 4 about here]

***

We select three different values for the correlation parameter ρ between the productive and environmental inefficiency terms, as shown in Equations (12) and (13). We consider three different situations: the productive inefficiency term has a low correlation ($\rho = 0.2$), moderate correlation ($\rho = 0.4$), or high correlation ($\rho = 0.8$) with the environmental inefficiency term. In the simulation, we also test the performance of the model with more output variables. We do this by multiplying the number of output variables by two (i.e., two desirable and two undesirable outputs).

In addition, we follow prior simulation studies on frontier models and assume that the noise term has a standard normal distribution and the inefficiency term a half-normal distribution (Pastor et al. 2002; Coelli 2005; Green 2005; Banker et al. 2010; Kuosmanen and Johnson 2010). As noted in Section 4.1, we assume that the productive and environmental inefficiency terms have a bivariate half-normal distribution, and these two terms have equal variances for their marginal distributions. We designate the variance parameters for the inefficiency distribution as $\sigma_{\upsilon}^{2} = 0.36$ and $\sigma_{u}^{2} = 5.06$.[2] The above variance parameter values are chosen for two main reasons. First, to represent as much as possible a real life situation, we want our experiment to include a moderate measurement error. In our experiment, the ratio between the variances of simulated inefficiency and the noise distributions is equal to 5.13, which, according to Banker and Natarajan (2009), corresponds to a situation with moderate measurement errors. Second, we want the inefficiency score distributions that we obtain from the simulated input-output data to be comparable to that in a real situation. By using the chosen variance values for the noise and inefficiency terms, the average eco-inefficiency score in our experiment is equal to 0.388, which is close to the average eco-inefficiency that we obtained in our prior evaluation of U.S. electric utility firms (Chen et al. 2010).[3]

4.4 Results

We replicate the simulation experiment 1000 times under the parameter values shown in Table 4. We summarize the simulation outputs for the two- and four-output model in Tables 5 and 6, which display the average performance statistics of the five models under different sample sizes and inefficiency correlation coefficients ρ. The results of the six- to ten-output model are illustrated in Figures 6 and 7. In the experiment, we evaluate the frontier models by the average rank correlation coefficients τ and the consistency measure δ.

***

[Insert Table 5, Table 6, and Figure 6 about here]

***

4.4.1 Result 1: correlation criterion

First we look at the correlation criterion, which corresponds to the rank correlation between inefficiency scores and simulated inefficiency (τ). Overall, the eco-inefficiency model shows a higher validity criterion value $(\tau = 0.474)$ than the other four models ($\tau =$ 0.032, 0.364, 0.450, and 0.352 ; see Table 5).

To compare the τ values of these five models under different sample sizes and ρ values, we apply the two-sample t test under the 1% significance level to examine whether the eco-inefficiency model has a higher average τ values than the other four models. For a specific sample size and the ρ value, the frontier model with a significantly higher average τ value than the other models is marked with double asterisks. For example, when the ρ is equal to 0.2 and sample size equal to 25, the average τ value of the eco-inefficiency model is equal to 0.401, and, according to the t-test, the eco-inefficiency model has a higher τ value than the SZ (average $\tau = 0.039$), hyperbolic (average $\tau = 0.367$), UINP (average $\tau = 0.376$), and DDF models (average $\tau = 0.332$) . From Table 5, it is clear that the eco-inefficiency model has the highest mean τ value than all the other four models in all test scenarios—the only exception occurs when we set ρ to 0.8 and the sample size equal to 100, and there the average τ value of the eco-inefficiency model is not significantly different from that of the DDF model (column 5 of Table 5). However, the tie in the τ value of the two models disappears when we increase the number of outputs from two to four and higher (see the last column of Table 5 and Figure 6).

In Table 5, we show the simulation results of the two-output and four-output models. We also implemented models that have six, eight, and ten outputs and the results are shown in Figure 6. In all cases, we find that the eco-inefficiency model has the highest τ value among the five models (again under the 1 % significance level). To illustrate the effect of increasing the output dimension on the validity criterion, we plot the validity criterion value in Figure 6 for the experiment with 2, 4, 6, 8 and 10 outputs. In the figure the eco-inefficiency model forms a clear contrast with the SZ, hyperbolic, and DDF models. The validity criterion values of these three model decrease markedly as the number of outputs increases from 2 to 10. The UINP model has its validity criterion values close (and lower than) to those of the eco-inefficiency model.

We also find that the correlation coefficient τ tends to be higher with a larger sample. Specifically, the average τ values of the frontier models (excluding the SZ model; see the explanation below) grow by 14.3 % when the sample size increases from 25 to 50, and by 8.6% when the sample size is enlarged from 50 to 100. The SZ model, however, shows the opposite results: its τ value drops by 22.4% and 20.5% when the sample size changes from 25 to 50, then to 100. Several studies have pointed out the advantage of utilizing a larger sample in the frontier analysis (see, e.g., Grosskopf (1996) and the references therein). With a larger sample, the frontier model has a higher likelihood to get a finer estimation of the frontier, and therefore we can reduce errors when calculating the inefficiency score (e.g., Grosskopf 1996; Zhang and Bartels 1998). In this regard, our experimental results of all five except for the SZ model are consistent with prior findings in the literature. The ρ parameter has a positive effect on the validity criterion for all five frontier models. That is, when the productive and environmental inefficiency terms are more highly correlated, frontier models have a higher likelihood of getting a precise inefficiency ranking of the sample.

4.4.2 Result 2: error rate criterion

Now we turn to Table 6 that shows the δ values of the five models. The δ represents the likelihood that the inefficiency score of an observation in the sample decreases after we double the undesirable output quantities of this observation. The eco-inefficiency, SZ and UINP models exhibits $\delta = 0.00$ in the experiment, which means that we do not find any instances of an error for these three models. However, the hyperbolic and DDF models show positive δ values: the DDF model has average δ values of 0.144% and 0.140% for the two-output and four-output models. When the sample size is set to 25, the average of δ is higher (0.26 ~0.30). The hyperbolic model has a lower average δ values than the DDF model (0.07% for the two-output model and 0.08% for the four-output model). In addition, the δ value is higher with a smaller sample and higher number of outputs (see Figure 6 and Table 6). The error rate appears to be lower when the number of outputs increases (Figure 6). Our conjecture is that the probability that higher output dimensionality makes an observation less probable to have its inefficiency score calculated based on the misspecified efficient frontier (see, e.g., Figure 1 line ‘cd’). A final note is that, although the average error rates from the experiment seem low, we illustrate next that errors are much more likely to occur for firms that have relatively high amounts of desirable and undesirable outputs under the DDF and hyperbolic model.

***

[Insert Figure 7 about here]

***

In Figure 7, we use the simulated data to illustrate the cause of errors in the DDF and hyperbolic models. The figure contains 100 points of simulated desirable and undesirable outputs. Points that have led to errors under the DDF model are circled in the dashed line and points that have yield errors under the hyperbolic model are marked with asterisks. We have two observations from the figure. First, points that have errors, errors as defined in Section 4.2.2, are those that identify benchmark targets on the misspecified efficient frontier when doubling undesirable outputs. Therefore, the more undesirable outputs a firm produces (compared with an average firm in the sample), the more likely the firm will experience errors under the DDF and the hyperbolic model. Second, the hyperbolic model is less prone to errors than the DDF model. The result echoes what we see in our simulation experiment (see Figure 6). Note that this is due to the directional vector we choose, and may not always hold true as a general result. But we should also mention that the literature has not yet provided a guideline for choosing a direction vector to avoid the problem identified in this paper.

5 DISCUSSION AND CONCLUSION

As environmental awareness and pressure increases, there are pressing needs for managers and policymakers to use effective tools to assess corporate performance according to firms' input consumption, products, and undesirable outputs that could create negative externalities to the natural environment and society. However, undesirable outputs, such as greenhouse gas emissions or other hazardous substances, usually do not have a fully functioning market that provides objective information about the relative importance of different factors. Consequently, aggregating multiple productive and ecological factors into a comprehensive and representative index becomes a real challenge to both academics and practitioners.

In this paper, we develop a new model to evaluate firms' eco-inefficiency by using the nonparametric frontier methodology. Our model allows us to construct the best-practice efficient frontier based on observed input-output quantities without the need to make explicit prioritization assumptions. Our model can calculate an eco-inefficiency score in the presence of multiple inputs and outputs. Weights for inputs and outputs are generated automatically through an optimization procedure, such that the evaluated firm will be assigned a set of "optimal weights" that maximizes the firm's efficiency relative to the other firms in the sample. Using this score, firms can derive the corresponding efficient target. The eco-inefficiency score can help firms understand their competitive standing in their own industry and provide a concrete benchmark target for subsequent efficiency improvement activities.

Our paper also makes major contributions to the frontier literature. We identify a major issue associated with the classical assumption on undesirable outputs in production economics. Specifically, we show that, under this assumption, frontier models may generate unreasonable estimations of eco-inefficiency scores and identify targets that are actually dominated in the production set. We compare our model with four current frontier models. The results from the Monte-Carlo experiment show that our approach provides a more robust measurement than these four frontier models. In the experiment, the eco-inefficiency model has attained higher rank correlations with the simulated inefficiency effect than current models across all experimental conditions. The result indicates that our eco-inefficiency model can provide a more precise ranking of inefficiencies than the current models. We have proved that eco-inefficiency model is guaranteed to identify non-dominant points on the frontier and therefore our model rectifies the inconsistency problem due to the classical assumption on undesirable outputs. The simulation results confirm our analytical findings and show that the eco-inefficiency score is non-decreasing with increased undesirable outputs. The simulation model we employed extends the traditional single-output production in the literature, which can only generate a single desirable output variable. We propose a new simulation framework amenable to the production process of multiple desirable and undesirable outputs. Our multi-output production function therefore allows for greater flexibility and opens up a new path for the analysis of frontier models.

Our model has important implications for operations research and is not limited to the measurement of productive efficiency for operations involving environmental negative externalities. Indeed many operations produce undesirable outcomes. This includes accidents, delays, defective products and waste. Our model can also be used for the measurement of efficiency frontiers in these situations.

Per definition, if outputs are undesirable then the firm should seek to minimize them. Therefore we need an accurate frontier model that accommodates this. Carbon dioxide, along with other greenhouse gases, is still unregulated in most cases. This is the situation where companies need to resort to the quantity-based efficiency measure. In this case the eco-inefficiency score will be a quantity-based measure that indicates the evaluated firm's distance to the frontier.

We assume that firms should minimize their inefficiency derived from undesirable outputs. For example, the goal of a firm can be to reduce its costly toxic waste. Regarding environmental emissions, the question is clear for regulated emissions, which are required to be minimized by law. However, there might be some other reasons for firms to reduce these emissions, such as reputation effects for example. A good example is carbon dioxide, which many of the largest firms are managing to curb, with the strong belief that mismanagement of the environmental practice can endanger corporate long-term sustainability. More generally, we are seeing increasing evidence about the impact of firms' environmental records and stance toward sustainability on corporate performance. An important body of empirical literature shows that improved environmental performance leads to better corporate performance (see the following meta-analyses Ambec and Lanoie, 2008; Margolis and Walsh, 2003; Orlitszky et al., 2003). For example, greener suppliers are more likely to secure their market share because green suppliers reduce buyer's environmental risk (Delmas and Montiel, 2008, Delmas and Nairn-Birch, 2010). Some firms reduce their greenhouse gas emissions to reduce their risks, drive innovation and save costs (CDP report 2009).

We also want to point out some promising research directions. As firms' environmental performance is receiving growing attention from market and governments, more firms are interested in the potential interactions between corporate eco-efficiency and different aspects of firm operations and management. The eco-inefficiency score provides an ideal proxy for eco-inefficiency in econometric models as a dependent or independent variable. See Banker and Natarajan (2008) for an updated procedure about how to regress inefficiency scores on independent variables of interests. One of the limitations of deterministic frontier models considered in this paper (as opposed to stochastic frontier models; see Chapter 9 of Coelli et al. (2005) for an introduction) is that they do not consider the influence from statistical noise. As a result, the eco-inefficiency score may be sensitive to outliers in the sample or sampling errors. Therefore a useful direction is to incorporate a stochastic term into the frontier formulation (e.g., Olesen and Petersen 1995; Post 2001; Post et al. 2002). Another promising direction is to carry out sensitivity analysis using bootstrapping (e.g., Simar and Wilson 1998) for the eco-inefficiency model. In this paper we focus on the situation where output price information is unavailable, and we calculate the eco-inefficiency score based on the "quantities" of inputs and outputs. It is also possible to calculate the revenue efficiency of a firm when price information for output variables exists (see Appendix-B).

References

Ambec, S., P. Lanoie. 2008. Does it pay to be green? A systematic overview. *The Academy of Management Perspectives* **22** (4) 45-62.

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

Banker, R.D., V.M. Gadh, W.L. Gott. 1993. A Monte Carlo comparison of two production frontier estimation methods: corrected ordinary least squares and data envelopment analysis. *European Journal of Operational Research* **67** (3) 332-343.

Banker D., R. Natarajan. 2008. Evaluating contextual variables affecting productivity using data envelopment analysis. *Operations Research* **56** (1) 48-58.

Banker, R.D., Z.E. Zheng, R. Natarajan. 2010. DEA-based hypothesis tests for comparing two groups of decision making units. *European Journal of Operational Research* **206** 231-138.

Berg, S.A., F.R. Forsund, E.S. Jansen.1992. Malmquist indices of productivity growth during the deregulation of Norwegian banking 1980-89. *The Scandinavian Journal of Economics* **94** 211-228.

Callens, I., D. Tyteca. 1999. Towards indicators of sustainable development for firms: A productive efficiency perspective. *Ecological Economics* **28**(1) 41-53.

Carbon Disclosure Project (CDP), Supply Chain Report 2009. http://www.cdproject.net/.

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

Cooper, W.W., L.M. Seiford, K. Tone. 2007. Data envelopment analysis: a comprehensive text with models, applications, references and DEA-solver software. Springer Verlag.

Chen, C.-M, M.A. Delmas. 2011. Measuring corporate social performance: an efficiency perspective. *Production and Operations Management*, forthcoming.

Chen, C.-M, M.A. Delmas, M. Montes-Sancho. 2010. Eco- vs. productive efficiency: A new approach to effective and comparative performance analysis. Center for Corporate Environmental Performance, UCLA Institute of the Environment & Sustainability.

Cherchye, L., B. De Rock, F. Vermeulen. 2008. Analyzing cost efficient production behavior under economies of scope: a nonparametric methodology. *Operations Research* **56**(1) 204-221.

Chung, Y.H., R. Färe, S. Grosskopf. Productivity and undesirable outputs: a directional distance function approach. *Journal of Environmental Management* **51**(3) 229-240.

Cleveland, C., M. Ruth. 1999. Indicators of dematerialization and materials intensity of use. *Journal of Industrial Ecology* **2** (3) 15-50.

Coelli, T. 2005. An Introduction to Efficiency and Productivity Analysis. Springer.

Corbett, C.J., P.R. Kleindorfer. 2001. Environmental management and operations management: introduction to part 1 (manufacturing and eco-logistics). *Production and Operations Management* **10**(2), 107-111.

Corbett, C.J., R.D. Klassen. 2006. Extending the horizons: Environmental excellence as key to improving operations. *Manufacturing & Service Operations Management* **8**(1) 5-22.

Cuesta, R.A., J.L Zofio. 2007. Hyperbolic efficiency and parametric distance functions: With application to Spanish savings banks. *Journal of Productivity Analysis* **24** (1) 31-48.

Delmas, M, M. Toffel. 2008. Organizational responses to environmental demands: Opening the black box. *Strategic Management* Journal **29**(10) 1027-1055.

Delmas, M., I. Montiel. 2008. The Diffusion of voluntary international management standards: responsible care, ISO 9000 and ISO 14001 in the Chemical Industry. *Policy Studies Journal* **36**(1): 65-93.

Delmas, M., N. Nairn-Birch. 2010. Is the Tail Wagging the Dog: An Empirical Analysis of Corporate Carbon Footprint and Financial Performance. Working paper. Center for Corporate Environmental Performance, UCLA Institute of the Environment & Sustainability.

Färe, R., S. Grosskopf, D. Margaritis. 2008. Efficiency and productivity: Malmquist and more. In" The Measurement of Productive Efficiency and Productivity Growth," edited by H. Fried, C.A.K. Lovell, and S.S. Schmidt. Oxford University Press.

Färe, R., S. Grosskopf, C.A.K. Lovell, C. Pasurka. 1989. Multilateral productivity comparisons when some outputs are undesirable: a nonparametric approach. *The Review of Economics and Statistics* **71** 90-98.

Färe, R., S. Grosskopf. 2004. *New directions: efficiency and productivity*. Kluwer Academic Pub.

Färe, R., S. Grosskopf, D.W. Noh, W. Weber. 2005. Characteristics of a polluting technology: theory and practice. *Journal of Econometrics* **126**(2) 469-492.

Fernández, C., G. Koop, M. Steel. 2002. Multiple-output production with undesirable outputs: an application to nitrogen surplus in agriculture. *Journal of the American Statistical Association* **97** 432-442.

Gibbins, J., H. Chalmers. 2008. Carbon capture and storage. *Energy Policy* **36**(12) 4317-4322.

Greene, W. 2005. Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics* **126** (2) 269-303

Grosskopf, S. 1996. Statistical inference and nonparametric efficiency: A selective survey. *Journal of Productivity Analysis* **7** 161-176.

Gut, A. 2009. An Intermediate Course in Probability. Springer.

Huan, Z., Y. Bian. 2007. DEA with undesirable factors. In "Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis." Edited by J. Zhu and W. Cook. Springer.

Johnson, N.L., S. Kotz, N. Balakrishnan. 2002. Continuous Multivariate Distributions: Models and Applications. John Wiley & Sons.

Kendall, M., J.D. Gibbons. 1990. Rank Correlation Methods. Oxford University Press.

King, A.A., M.J. Lenox. 2002. Exploring the locus of profitable pollution reduction. *Management Science* **48**(2) 289-299.

Klassen, R.D., C.P. McLaughlin. 1996. The impact of environmental management on firm performance. *Management Science* **42**(8) 1199-1214.

Klassen, R., S. Vachon. 2003. Collaboration and evaluation in the supply chain: the impact on plant-level environmental investment. *Production and Operations Management* **12**(3) 336-352.

Koop, G. 1998. Carbon dioxide emissions and economic growth: a structural approach. *Journal of Applied Statistics* **25**(4) 489-515.

Kuosmanen, T., M. Kortelainen. 2005. Measuring eco-efficiency of production with data envelopment analysis. *Journal of Industrial Ecology* **9**(4) 59-72.

Kuosmanen, T., A. Johnson. Data envelopment analysis as nonparametric least-squares regression. *Operations Research* **58** (1) 149-160.

Lee, J., J. Park, T. Kim. 2002. Estimation of the shadow prices of pollutants with production/environment inefficiency taken into account: a nonparametric directional distance function approach. *Journal of Environmental Management* **64**(4) 365-375.

Margolis, J., J. Walsh. 2003. Misery loves companies: Rethinking social initiatives by business.

Administrative Science Quarterly **48** (2) 268-305.

Olesen, O.B., N.C. Petersen. 1995. Chance constrained efficiency evaluation. *Management Science* **41**(3), 442-457.

Orlitzky, M., F.L. Schmidt, S.L. Rynes. 2003. Corporate social and financial performance: A meta-analysis. *Organization Studies* **24**(3) 403-441.

Park, K., W. Weber. 2006. A note on efficiency and productivity growth in the Korean banking industry, 1992-2002. *Journal of Banking and Finance* **30**(8) 2371-2386.

Pastor, J.T., J.E.L. Ruiz, I. Sirvent. 2002. A statistical test for nested radial DEA models. *Operations Research* **50**(4) 728-735.

Pathomsiri, S., A. Haghani, M. Dresner, R. Windle. 2008. Impact of undesirable outputs on the productivity of US airports. *Transportation Research Part E* **44**(2) 235-259.

Picazo-Tadeo, A., E. Reig-Martinez, F. Hernandez-Sancho. 2005. Directional distance functions and environmental regulation. *Resource and Energy Economics* **27**(2) 131-142.

Porter, M.E., F. Reinhardt. 2007. Grist: A Strategic Approach to Climate. Forethought. *Harvard Business Review* **85**(10) 22-26.

Post, T. 2001. Performance evaluation in stochastic environments using mean-variance data envelopment analysis. *Operations Research* **49**(2) 281-292.

Post, T., L. Cherchye, T. Kuosmanen. 2002. Nonparametric efficiency estimation in stochastic environments. *Operations Research* **50**(4) 645-655.

Seiford, L. M., J. Zhu. 2002. Modeling undesirable factors in efficiency evaluation. *European Journal of Operational Research* **142**(1) 16-20.

Simar, L., P. Wilson. 1998 .Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science* **44**(1) 49-61.

Tone, K. 2001. A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research* **130**(3) 498–509.

Verbruggen, A. 2009. Beyond Kyoto, plan B: a climate policy master plan based on transparent metric. *Ecological Economics* **68** 2930-2937.

Watanabe, M., K. Tanaka. 2007. Efficiency analysis of Chinese industry: A directional distance function approach. *Energy Policy* **35**(12) 6323-6331.

Zhang, Y., R. Bartels. 1998. The effect of sample size on the mean efficiency in DEA with an application to electricity distribution in Australia, Sweden and New Zealand. *Journal of Productivity Analysis* **9** (3) 187-204.

Zofio, J.L., A.M. Prieto. 2001. Environmental efficiency and regulatory standards: the case of CO2 emissions from OECD industries. *Resource and Energy Economics* 23(1) 63-83.

APPENDIX-A: PROOFS OF THE THEOREMS

Proof of Theorem 1: For all firms $k = 1,...,K$, we first substitute input i and output n , respectively by $\tilde{x}_{ki} = \alpha x_{ki}$ and $\tilde{d}_{kn} = \beta d_{kn}$, where α and β are arbitrary positive numbers. For input i , it is straightforward to prove the homogeneity, since form Equation (8-2) we can derive

$$\sum_{k=1}^{K} z_k \tilde{x}_{ik} \le \tilde{x}_{ik} \Rightarrow \alpha \sum_{k=1}^{K} z_k x_{ik} \le \alpha x_{ik} \Rightarrow \sum_{k=1}^{K} z_k x_{ik} \le x_{ik} \tag{17}$$

For output n , observe first that the term associated with this output in the objective function (8-1) is now rewritten as $\tilde{g}_i^d \big/ \tilde{d}_{ki} = (\tilde{g}_i^d / \beta) \big/ d_{ki}$. In addition, in Equation (8-3), we obtain

$$\sum_{k=1}^{K} z_k \tilde{d}_{kn} \le \tilde{d}_{kn} + \tilde{g}_n^d \Rightarrow \beta \sum_{k=1}^{K} z_k d_{kn} \le \beta d_{kn} + \tilde{g}_n^d \Rightarrow \sum_{k=1}^{K} z_k d_{kn} \le d_{kn} + \tilde{g}_n^d \big/ \beta. \tag{18}$$

Since in model (8) we restrict that $\tilde{g}_n^d \ge 0$, it follows that $\tilde{g}_n^d \big/ \beta \ge 0$. By observing $\tilde{g}_k^d \big/ \beta$ in (8-1) and (8-3), we can obtain problem (8) in its original form. The homogeneous property of undesirable outputs can be proved analogously. □

Proof of Theorem 2: Consider an arbitrary input output vector $(x, d, u) \in \Omega \subset \Re^{M \times N \times P}$. Without loss of generality, suppose that (x, d, u) is dominated by some $(x, d', u') \in \Omega$. Then there must exist a N-by-P non-negative vector $(g^d, g^u) \ne 0$, for which $(x, d', u') = (x, d + g^d, u - g^u)$. It follows that there must also exist nonnegative $z = (z_1, ..., z_k)$ satisfying

$$\sum_{m=1}^{M} z_k x_{km} \le x', \sum_{n=1}^{N} z_k d_{kn} \ge d', \text{ and } \sum_{p=1}^{P} z_k u_{kp} = u' \tag{19}$$

, which shows that (z, g^{d}, g^{u}) is a feasible solution to the eco-efficiency model (8). Given that $(g^{d}, g^{u}) \neq 0$, we obtain $E(x, d, u) > 0$, where $E(.)$ is defined in (8).

Conversely, we can show that when $E(\alpha, \beta, \gamma) > 0$, the optimal solution (z^{*}, g^{d*}, g^{u*}) can be used to construct a vector $(x, d'', u'') = (x, d + g^{d*}, u - g^{u*}) \in \Omega$, which dominates (x, d, u).

APPENDIX-B: REVENUE INEFFICIENCY INDEX

To calculate the revenue efficiency of a firm (X^*, Y^*, U^*), the revenue efficiency is defined as (20):

$$(P_D'Y^* - P_U'U^*) \Big/ \max_{(D,U)}\{P_D'Y - P_U'U \mid (X^*, Y, U) \in \Omega\} \qquad (20)$$

where P_D and P_U are prices vectors for desirable outputs and undesirable outputs, respectively. The revenue efficiency index is equal to a firm's current revenue divided by the maximum revenue achievable by using inputs X^*. We should note that the revenue inefficiency index is well-defined only when the maximal revenue is non-zero. See Cherchye et al. (2008) for a related discussion on calculating cost inefficiency through the frontier model. Compared with the eco-efficiency concept where firms pursue the "efficiency" status in multiple inputs and outputs, the computation of the revenue (or cost or profit) inefficiency is relatively straightforward, because we can expect rational firms to follow price signals and optimize their production accordingly. In this situation the frontier model becomes a single-objective problem of revenue maximization and cost minimization. In this sense, estimating the economic inefficiency is less of a modeling issue but more of an empirical issue (i.e., can we get a reasonably accurate price estimate?) As we also primarily focus on the situation where price information is unavailable, we did not pursue this direction further.

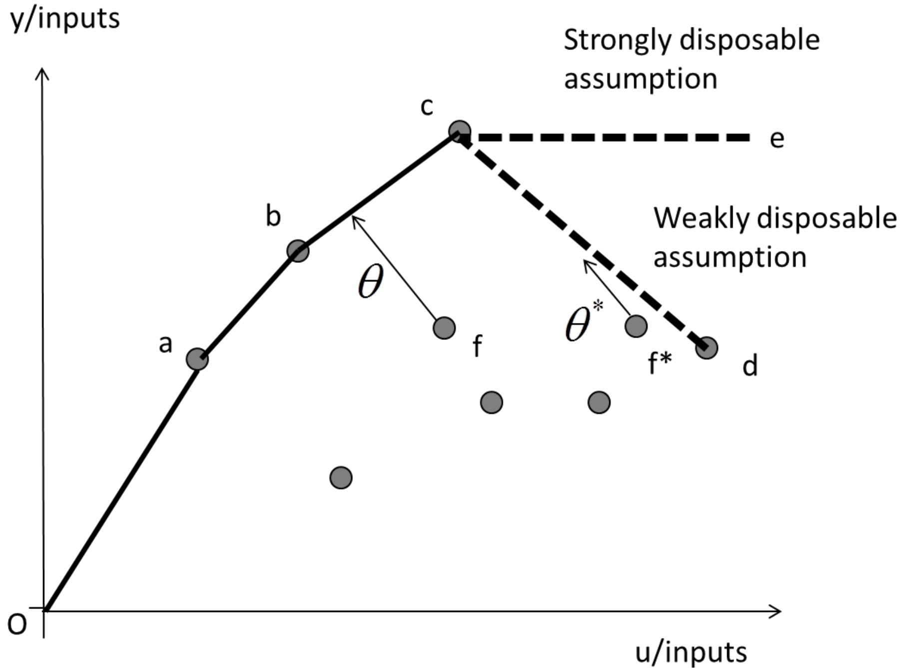
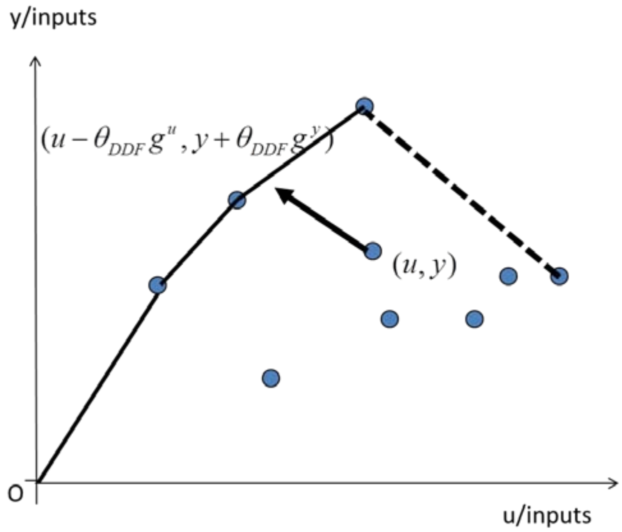
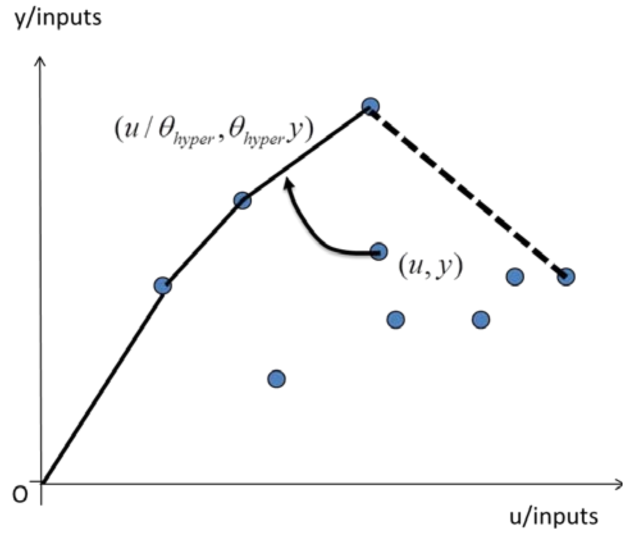


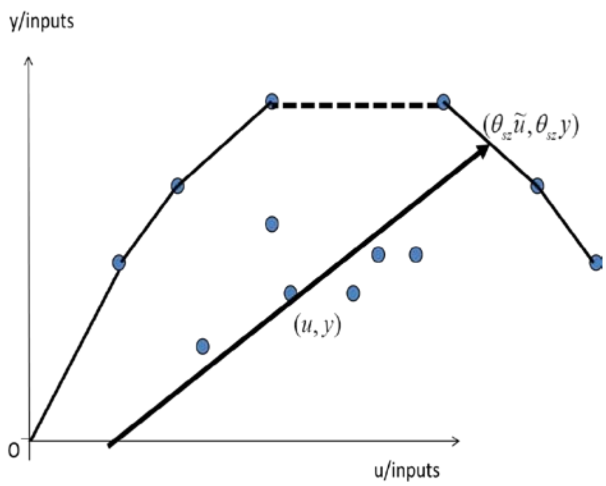
Figure 1 Illustration of the efficiency measure



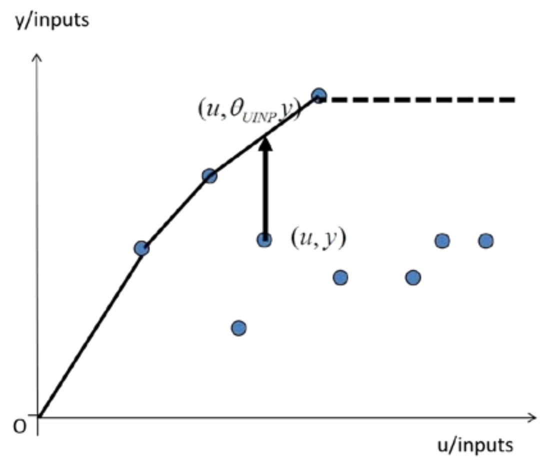
(a) DDF model



(b) Hyperbolic model



(c) SZ model



(d) UINP model

Figure 2 Illustration of the four frontier models for undesirable outputs

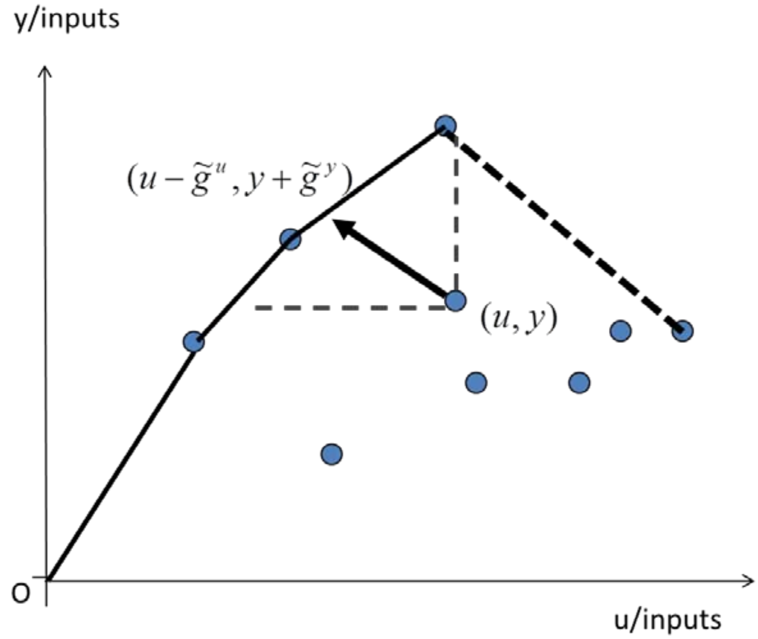


Figure 3 Illustration of the proposed eco-inefficiency model

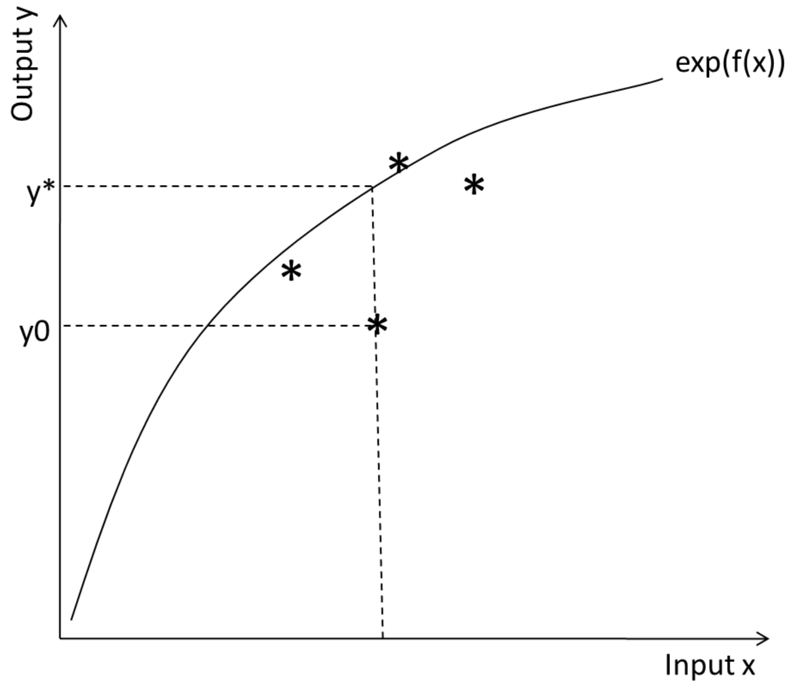


Figure 4 Illustration of the Cobb-Douglas function

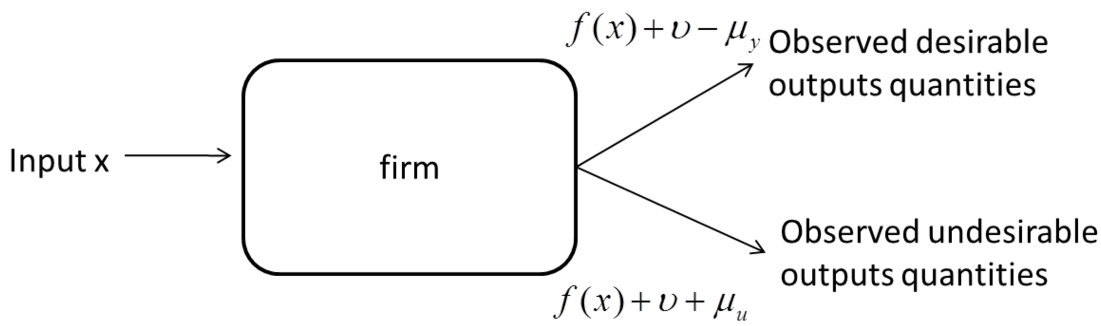


Figure 5 Productive and environmental inefficiency

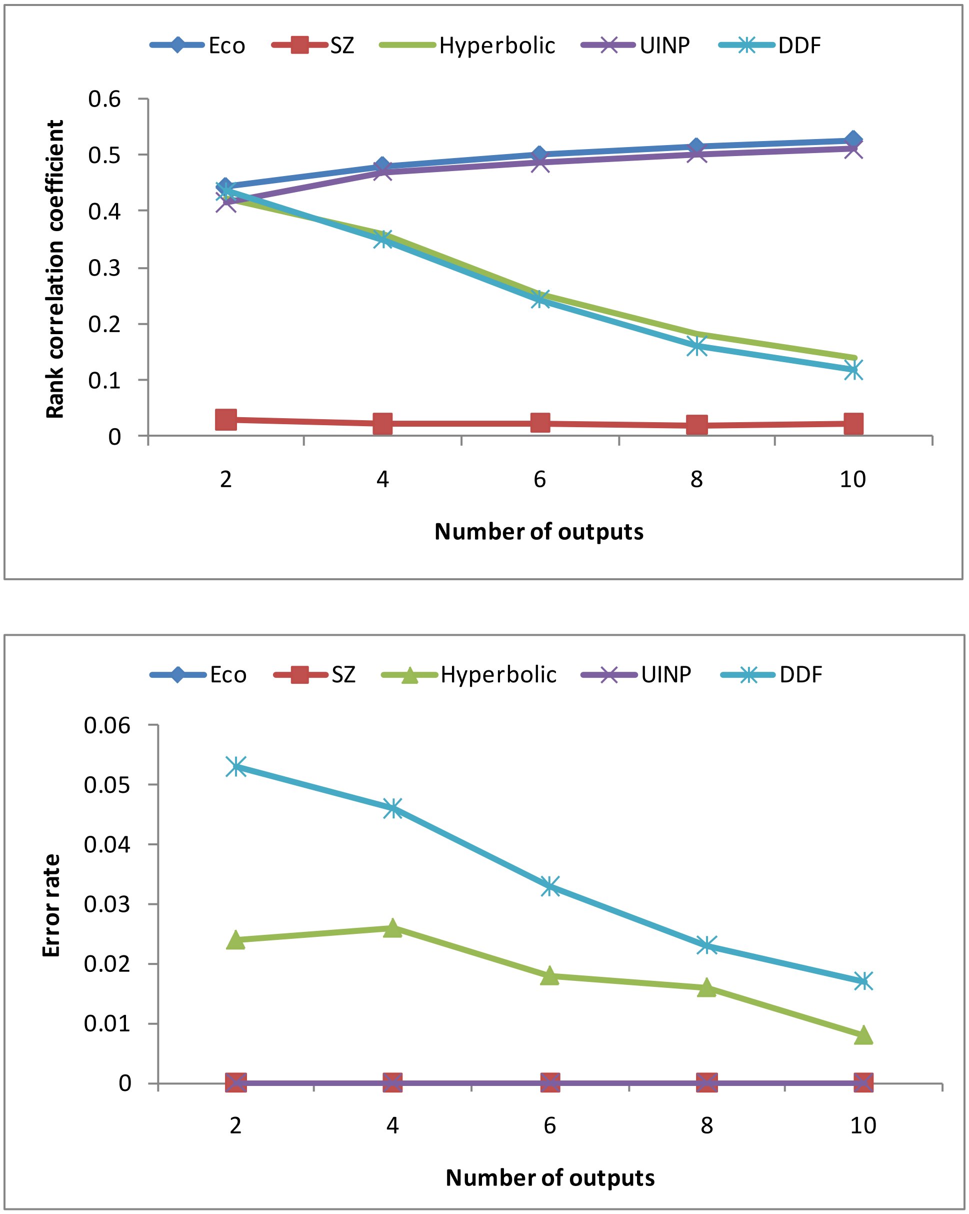


Figure 6 Kendall's tau values with different numbers of outputs (sample size 100, $\rho = 0.4$)

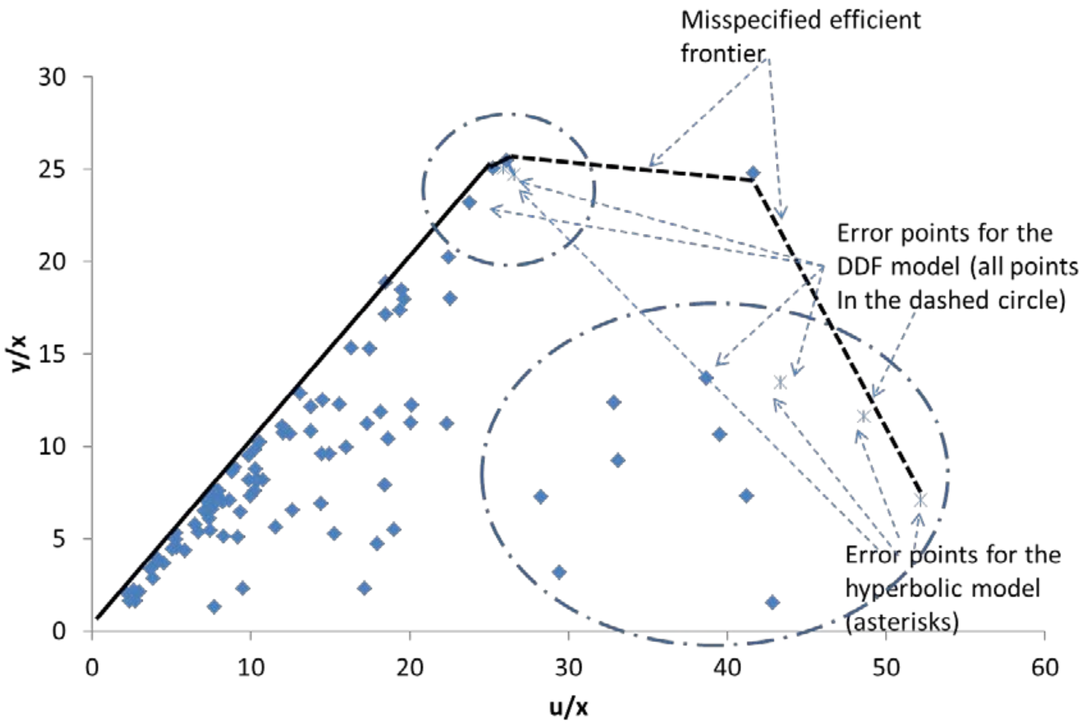


Figure 7 Illustration of 100 simulated data points (n=100, $\rho = 0.4$)

Table 1 Modeling assumptions and ranges of scores

Efficiency measures	Assumptions on undesirable outputs	Score range	Potential limitations
Hyperbolic efficiency model	Weakly disposable	$[1,\infty)$	The model is a non-linear optimization problem and therefore could be difficult to solve
Directional distance function	Weakly disposable	$[0,\infty)$	The model requires us to specify a directional vector beforehand, and the inefficiency score will vary for different choices of directional vectors
UINP model	Treated as inputs	$[1,\infty)$	Not representative of the production process; the model cannot provide the benchmark value for undesirable outputs.
Seiford and Zhu's model	Strongly disposable	$[1,\infty)$	The model requires us to specify a translation vector beforehand, and the inefficiency score will vary for different choices of translation vectors

Table 2 Descriptive statistics of the paper mill data (n=30)

Variable	Mean	Std. Dev.	Min	Max
Fiber	103997.20	65671.23	14743.00	312910.00
Energy	2285863.00	1415598.00	304031.00	5771544.00
Capital	78500000.00	49700000.00	18100000.00	262000000.00
Labor	1107302.00	767867.10	163993.00	3144336.00
Paper	106615.60	65494.73	1800.00	293000.00
Biochemical oxygen demand (BOD)	3014.00	3376.71	86.79	13318.19
Total suspended solids	1807.54	1896.37	17.38	9015.50
Particulates	327.23	596.22	2.84	2284.27
Sox	2730.19	3136.69	1.26	12129.65

Table 3 Efficiency scores from different models

	DDF		Hyperbolic model		SZ model		UINP	
mills	Original data	Undesirable outputs x2	Original data	Undesirable outputs x2	Original data	Undesirable outputs x2	Original data	Undesirable outputs x2
1	0	0	1	1	1.25	1.25	1.12	1.22
2	0	0	1	1	1.39	1.39	1	1
3	0	0	1	1	1	1	1	1
4	0	0	1	1	1.22	1.22	1.04	1.18
5	15.31	33.92	1.36	1.32	1.52	1.52	1.39	1.41
6	0	0	1	1	1	1	1	1
7	0	0	1	1	1	1	1	1
8	0	0	1	1	1.33	1.33	1.21	1.37
9	0	0	1	1	1.01	1.01	1	1
10	0	0	1	1	1	1	1	1
11	0	0	1	1	1.13	1.13	1.12	1.13
12	0	0	1	1	1.14	1.14	1.15	1.15
13	0	0	1	1	1.53	1.53	1.41	1.43
14	187.19	0	1.16	1	1.27	1.27	1.21	1.21
15	0	0	1	1	1	1	1	1
16	0	0	1	1	1	1	1	1
17	61.79	0	6.56	1	1	1	1	1
18	0	0	6.95	1	1.15	1.15	1	1
19	0	0	1	1	1.45	1.45	1.45	1.45
20	0	0	1.77	1	1.76	1.76	1.54	1.54
21	0	0	1	1	1	1	1	1
22	0	0	1	1	1.35	1.35	1.28	1.32
23	0	0	1	1	1.30	1.30	1.29	1.31
24	0	0	1	1	1.28	1.28	1	1
25	0	0	1	1	1.44	1.44	1	1
26	0	0	1	1	1.18	1.18	1.04	1.12
27	0	0	1	1	1	1	1	1
28	0	0	1	1	1	1	1	1
29	0	0	1	1	1.00	1.00	1	1
30	29.69	0	1.07	1	1.12	1.12	1.08	1.12
Avg.	9.80	1.13	1.43	1.01	1.19	1.19	1.11	1.13
Std. Dev.	35.76	6.19	1.46	0.06	0.21	0.21	0.16	0.17

Table 4 Experiment parameters

Parameters	Value
Distribution of the input variable (x)	Uniform[1,4]
Sample size	[25,50,100]
The number of desirable and undesirable outputs	2, 4, 6, 8,and 10
Correlation between the productive and environmental inefficiency terms (λ)	[0.2,0.4,0.8]
Error term distribution (υ)	N(0,0.36)
Inefficiency term distribution (μ)	|N(0, 5.06)|
Covariance matrix of the two inefficiency terms (Σ)	$\begin{bmatrix} 5.06 & \rho * 5.06 \\ \rho * 5.06 & 5.06 \end{bmatrix}$

Table 5 Average Kendall's tau rank correlation coefficients (τ)

		Two-output model			Four-outputs model		
Frontier model	Sample size	$\rho = 0.2$	$\rho = 0.4$	$\rho = 0.8$	$\rho = 0.2$	$\rho = 0.4$	$\rho = 0.8$
Eco-inefficiency model	25	0.401**	0.414**	0.501**	0.435**	0.461**	0.551**
(average: 0.474)	50	0.425**	0.436**	0.514**	0.453**	0.470**	0.563**
	100	0.427**	0.443**	0.524**	0.459**	0.480**	0.573**
SZ model	25	0.039	0.040	0.052	0.033	0.037	0.042
(average: 0.032)	50	0.031	0.034	0.036	0.025	0.026	0.036
	100	0.026	0.029	0.028	0.018	0.023	0.025
Hyperbolic model	25	0.367	0.374	0.462	0.210	0.217	0.253
(average: 0.364)	50	0.400	0.409	0.486	0.278	0.290	0.346
	100	0.412	0.427	0.506	0.339	0.358	0.420
UINP model	25	0.375	0.386	0.466	0.416	0.441	0.520
(average: 0.450)	50	0.398	0.407	0.479	0.441	0.456	0.540
	100	0.402	0.416	0.491	0.449	0.470	0.554
DDF model	25	0.332	0.345	0.432	0.188	0.195	0.232
(average: 0.352)	50	0.387	0.401	0.484	0.261	0.273	0.333
	100	0.418	0.436	0.527**	0.331	0.350	0.420

Note: Given a specific sample size and ρ value, the model"**" means that it has a significant higher τ value than all the other models at the 1% significance level.

Table 6 Error rates (δ in %) of the frontier models

		Two-output model			Four-outputs model		
Frontier model	Sample size	$\rho = 0.2$	$\rho = 0.4$	$\rho = 0.8$	$\rho = 0.2$	$\rho = 0.4$	$\rho = 0.8$
Eco-inefficiency model	25	0.0	0.0	0.0	0.0	0.0	0.0
	50	0.0	0.0	0.0	0.0	0.0	0.0
	100	0.0	0.0	0.0	0.0	0.0	0.0
SZ model	25	0.0	0.0	0.0	0.0	0.0	0.0
	50	0.0	0.0	0.0	0.0	0.0	0.0
	100	0.0	0.0	0.0	0.0	0.0	0.0
Hyperbolic model	25	0.15	0.16	0.12	0.18	0.17	0.13
	50	0.084	0.064	0.068	0.092	0.056	0.056
	100	0.026	0.024	0.020	0.018	0.026	0.016
UINP model	25	0.0	0.0	0.0	0.0	0.0	0.0
	50	0.0	0.0	0.0	0.0	0.0	0.0
	100	0.0	0.0	0.0	0.0	0.0	0.0
DDF model	25	0.3	0.236	0.252	0.268	0.26	0.28
	50	0.116	0.122	0.094	0.098	0.11	0.106
	100	0.067	0.053	0.058	0.045	0.046	0.043

[1] From the ISI Web of Knowledge website, Chung et al. (1997) receives 111 citations, Färe et al. (1989) receives 180 citations, and Seiford and Zhu (2002) have 47 citations (data retrieved Nov. 30 2010). The citation record of Berg et al. (1992) is not found in ISI Web of Knowledge. We find from Google Scholar that Berg et al. (1992) has 145 citations on Nov. 30 2010.

[2] The values of the variances of the two stochastic terms are taken from the Banker and Natarajan (2008). In Banker and Natarajan (2008), they specify the variances as $\sigma_{\upsilon} = 0.04$ and $\sigma_b = 0.15$. We multiple these two values by 15 times in order to obtain the desired average mean inefficiency, while maintain a noise-to-signal ratio similar to that in Banker and Natarajan (2008).

[3] In Chen et al. (2010), we evaluate the eco-inefficiency of 85 US electric utility firms based on total sales (in MWH), and three types of undesirable gas emissions (i.e., CO_2, NO_x, and SO_2), and four inputs including plant values, total operations and maintenance expenditure, labor costs, and purchased electricity. The eco-inefficiency score in Chen et al. (2010) has an average of 0.357 and standard deviation 0.697. Using the parameter values in Table 4 and with the same number of output variables as in Chen et al. (2010), we obtain a sample of eco-inefficiency scores that average at 0.388 and have a standard division of 0.770. The average efficiency and standard division are deliberately set higher than what we observe in Chen et al. (2010) because prior studies have generally confirmed that high dimensions in the frontier model (i.e., more input and output variables) are associated with lower average inefficiency estimates and variations of the scores (e.g., Dyson et al. 2001; Zhang and Bartels 1998).