PRODUCTION FRONTIER METHODOLOGIES AND EFFICIENCY AS A PERFORMANCE MEASURE IN STRATEGIC MANAGEMENT RESEARCH

Chien-Ming Chen
Nanyang Business School,
Nanyang Technological University, Singapore 639798
Email: cmchen@ntu.edu.sg

Magali A. Delmas
Institute of the Environment & Sustainability,
Anderson School of Management,
University of California, Los Angeles, CA 90095
Email: delmas@ucla.edu

Marvin B. Lieberman
Anderson School of Management,
University of California, Los Angeles, CA 90095
Email: marvin.lieberman@anderson.ucla.edu

Forthcoming *Strategic Management Journal*

Abstract

The measurement of corporate performance is central to strategic management research. A common objective of this research is to identify top performers in an industry and their sources of competitive advantage. Despite this focus on best firms and practices, most researchers utilize statistical methods that identify average effects in a sample, and they assess a single performance dimension while ignoring other relevant dimensions. Emphasis on purely financial measures can overlook the fact that a firm's efficiency in transforming resources has been shown as a major source of competitive advantage. In this article we demonstrate how frontier methodologies, such as Data Envelopment Analysis and the Stochastic Frontier approach, can address these challenges. We provide an illustration based on longitudinal data from U.S. and Japanese automobile producers.

Keywords: Performance Measurement, Data Envelopment Analysis, Stochastic Frontier Analysis, Research methodology, Efficiency.

Electronic copy available at: http://ssrn.com/abstract=2337697

INTRODUCTION

Research in strategic management is focused on finding strategies and attributes that enable an organization to outperform its competitors. One or more firms in an industry are often argued to have attained 'competitive advantage' relative to the majority of rivals (Porter, 1985). A common challenge facing strategic management scholars is to objectively identify the leading competitors and assess the reasons for their superiority.

The purpose of this article is to present an introduction to a useful methodology, the frontier methodology, which measures the *relative efficiency* of firms in transforming resources to achieve business goals. Although the idea of relative efficiency is closely tied to the concept of competitive advantage (which occurs when an organization acquires or develops some combination of attributes that allows it to outperform its competitors) surprisingly few studies have used frontier methods in strategic management research. A search for studies using these approaches in three prestigious management journals (*Strategic Management Journal*, *Academy of Management Journal*, and *Management Science*) yields only 16 articles as compared to hundreds of articles using profit measures such as Return on Assets or Sales.[1] This note demonstrates the potential of these methods for studies in strategic management and provides practical guidance on the relative benefits of the different approaches.

Identifying firms that possess advantage over competitors is a straightforward exercise if performance can be succinctly captured by a single measure. However, the strategic management literature points to a diverse array of objectives and actions regarding the creation of competitive

[1] A search for profit, return on sales, return on assets, and Tobin's q on the publisher's website of these three journals yields 4,174 articles (access date: July 22 2013).

Electronic copy available at: http://ssrn.com/abstract=2337697

advantage (e.g., Dyer, 1996; Douglas and Judge Jr., 2001; Hillman and Keim, 2001). As a consequence, researchers are faced with a range of performance measures that relate to various aspects of corporate activities, including accounting, finance, operations, marketing and corporate social responsibility, and often there is no clear guideline to select valid measures for corporate performance (Dess and Robinson Jr., 1984; Waddock and Graves, 1997; Ray, Barney and Muhanna, 2004; Hull and Rothenberg, 2008; Godfrey, Merrill, and Hansen, 2009). For example, in a survey of 374 studies published in the Strategic Management Journal from 1980 to 2004, Combs, Crook and Shook (2005) found that 56 different measures were used to operationalize the corporate performance construct. It is seldom the case that a single firm will top the list all the available performance measures. Identifying the firms that define the best performance frontier across the relevant measures is an important task that is seldom performed. Some of the methods described in this article can be applied to perform this task.

Although performance includes multiple dimensions, studies in strategic management focus most commonly on firms' financial returns. Superior profitability is achieved through some combination of cost efficiency and the ability of the firm to charge a price premium (Porter, 1985). Strategies designed to enhance a firm's efficiency are often quite different from those oriented toward charging a higher price. Yet few empirical studies consider these two sources of advantage separately. By providing a technique to assess efficiency, the frontier methods described in this article enable an understanding of firm performance that is deeper than a mere comparison of profitability or financial returns. Indeed, as we demonstrate below, the major American automotive companies demonstrated relatively strong financial performance in the 1980s and 1990s, even though their Japanese rivals maintained higher efficiency. But when the market segments that sustained the US producers—trucks and SUVs—contracted in the wake of increasing oil prices, the financial

performance of the US companies collapsed, leading to the bankruptcies of General Motors (GM) and Chrysler in 2009.

Frontier methods have been designed to assess an individual firm's performance relative to the best performers in an industry and are therefore well suited to address the challenges of measuring competitive advantage (Majumdar, 1998; Durand and Vargas, 2003; Delmas and Tokat, 2005; Lieberman and Dhawan, 2005; Knott and Posen, 2005). Frontier methods represent performance by an efficiency score, calculated as the firm's distance to the best practice industry frontier. The efficient frontier is estimated directly through the observed inputs and output(s) of each firm.

Frontier models can be used by strategy researchers to test theories of various factors that lead to competitive advantage. They are also particularly suited to conceptualize and measure firm-specific capabilities. As Dutta, Narasimhan and Rajiv (2005, p. 278) point out, 'one can think of capabilities as the efficiency with which a firm uses the inputs available to it (i.e., its resources, such as R&D expenditure), and converts them into whatever output(s) it desires….'.

This research note covers the two main frontier approaches: Data Envelopment Analysis (DEA) (Charnes, Cooper, and Rhodes, 1978; Banker, Charnes, and Cooper, 1984), and the Stochastic Frontier Approach (SFA) (Aigner, Lovell, and Schmidt, 1977; Meeusen and van den Broeck, 1977; Kumbhakar and Lovell, 2003).[2] A fundamental distinction between them is that DEA is nonparametric whereas SFA is a parametric approach. This distinction makes these two approaches have their own specific areas of strength, as we elaborate in this note.

[2] Several papers introduce both SFA and DEA (e.g., Hjalmarsson, Kumbhakar and Heshmati, 1996; Coelli et al., 2005; Murillo-Zamorano, 2004). These sources provide information from a technical or economic standpoint without specific focus on strategic management.

We begin by introducing the fundamentals of the DEA and SFA methods. We then illustrate the methodologies with data on U.S. and Japanese automobile producers. This illustration demonstrates the advantages of the frontier methods as well as caveats that researchers should consider when applying these methods or interpreting work that make use of them.

FRONTIER METHODOLOGIES

In frontier methodologies, a firm's performance is measured in terms of distance from the industry's efficient frontier. The efficient frontier is a function that indicates the maximum attainable level of output corresponding to a given quantity of input. The frontier is estimated based on the observed inputs and outputs of all firms in the industry (or a representative sample). For example, consider firms that employ two inputs, labor and capital, to produce one or more outputs, such as cars, trucks and consumer loans. The efficient frontier represents the maximum amount of output(s) that can be produced from a specific amount of labor and capital. Each firm's relative efficiency can be defined, based on the distance between the firm's actual output and the estimated 'best practice' frontier (e.g., expressed as the ratio of the firm's observed output relative to the fully-efficient output). DEA and SFA are two alternative approaches through which the industry efficient frontier and the firm-specific efficiencies can be estimated.

Basic DEA model

DEA generates the efficient frontier through a mathematical optimization model (Charnes *et al.*, 1978; Banker *et al.*, 1984). The DEA frontier is a linear surface or 'piecewise hyperplane' extrapolated from all efficient firms in the sample such that the inefficient firms are 'enveloped' by the frontier. We illustrate the DEA frontier in Figure 1(a), where we consider a simple case with one input and one output variable for seven firms (firm *a* to firm *g)*. The figure depicts this frontier, along with the input and output of the seven firms. The DEA frontier is the line that goes from the origin *O*

through firm *a* that corresponds to the highest ratio of output to input. The area below the frontier consists of feasible yet inefficient input-output combinations.[3] Firm *a* is therefore efficient, while firms *b* to *g* that are below the frontier are inefficient. Figure 1(a) also contains an OLS regression line with the intercept set to zero. It illustrates that the OLS model would underestimate the frontier because it does not permit inefficiency and assumes that deviation from the average input-output correspondence is purely random.[4]

Inefficiency is measured by a firm's distance to the frontier. For example, the DEA efficiency score of firm *f* is calculated as $\overline{Oi}$ (observed output level) divided by $\overline{Oi^*}$ (efficient output level given firm *f*'s input). Therefore firm *f* is inefficient with an efficiency score less than one.[5] Firm *a* is on the DEA frontier and therefore it is efficient with an efficiency score of one. The DEA model works similarly when there are multiple outputs; in that case the efficiency score is calculated as the possible percentage increase of 'all outputs,' given the current input. We provide detailed mathematical formulations of DEA in Appendix A.

As shown in the above example, the DEA efficiency score is calculated based on input and output quantities. In the presence of multiple inputs and outputs, prior studies show that inadvertent aggregation of different performance measures can sometimes result in overlooked competitive strength in some performance dimensions (Ray *et al.*, 2004; Chen and Delmas, 2011). DEA does not

[3] Here we assume the production technology has constant returns-to-scale (CRS), which, in the presence of multiple inputs and outputs, means that a proportional change in a firm's inputs (e.g., all inputs are increased by 50%) should lead to the same proportional change in a firm's outputs (all outputs should increase 50%).

[4] See Coelli *et al.* (2005) pp.258 – 259 for statistical tests for the existence of inefficiency effects. [5] The DEA model used to measure output inefficiency is presented in (5) in Appendix A. We should note that the efficiency score obtained from formulation (5) (i.e., the θ^{output}) is the reciprocal of our definition provided here, and the score is greater than or equal to one. For example, firm b's DEA score from formulation (5) is equal to $\overline{Oi^*}/\overline{Oi}$, which is greater than 1. We provide a more in-depth discussion in Appendix A.

[5] The DEA model used to measure output inefficiency is presented in (5) in Appendix A. We should note that the efficiency score obtained from formulation (5) (i.e., the θ^{output}) is the reciprocal of our definition provided here, and the score is greater than or equal to one. For example, firm b's DEA score from formulation (5) is equal to $\overline{Oi^*}/\overline{Oi}$, which is greater than 1. We provide a more in-depth discussion in Appendix A.

require explicit weight specifications or assumptions about the production function and probability distributions for technical inefficiency. The DEA model calculates weights for each firm through an optimization procedure, which is detailed in Appendix A. The weights are calculated based on which input(s) a specific firm excels at utilizing, or which output(s) a firm excels at generating in comparison to the other firms in the sample. By assigning higher weights to the input and output variables a specific firm excels in utilizing or generating, and low weights to the others, the algorithm maximizes the performance of each firm in light of its particular competence.[6] In addition, DEA does not specify a specific production function, and the efficiency scores are obtained from solving linear programming problems. This feature enhances the computational convenience of DEA and reduces the risk of model misspecifications, but the trade-off is that DEA is a deterministic approach and can be sensitive to outliers.

***

[Insert Figure 1(a) and Figure 1(b) about here]

***

Basic SFA model

In contrast to the DEA frontier, the SFA specification involves a production or cost function with an error term that includes two components: a random noise effect and an inefficiency effect (Aigner and Chu, 1968; Aigner *et al.*, 1977; Meeusen and van den Broeck, 1977). The stochastic efficient

[6] This flexibility in the assignment of weights can lead to a situation where a large proportion of firms are found to be efficient, or a specific firm appears efficient because it specializes in a few rather than most outputs or inputs, and thus DEA loses its discriminatory power. Lack of discriminatory power can also occur when the number of input and output variables is high relative to the sample size: having more variables increases the likelihood that firms are found efficient because DEA optimizes the weights of their 'niche' inputs or outputs. If we have prior knowledge about the relative importance of inputs and outputs, it is possible to constrain the weight of one input or output relative to that of another in the DEA models. For example, we can impose that the weight of input A must be three times to five times higher than that of input B. Note that when weight restriction constraints are imposed, the efficient frontier and therefore the efficiency scores would both change. See Cooper, Ruiz, and Sirvent (2011) for a detailed discussion about weight restriction methods in DEA.

frontier assumes that production outputs are subject to random shocks that are not under the direct control of firms. Specifically, let y_{it} denote the output of firm *i* produced at time *t*, and let $X_{it} = (x_{it1}, x_{it2}, \dots, x_{itm})$ be the collection of m inputs or resources consumed for the purpose of producing y_{it}. For example, a firm's inputs may include its capital assets and employees, and the output can be revenue, value-added (revenue minus materials costs), or physical quantity of output. The stochastic production function with panel data is given by:

$$y_{it} = f(X_{it})e^{v_{it}}e^{-\mu_{it}} \tag{1}$$

On the right hand side of (1), there are three components. The first involves a production function *f* that transforms the various input factors X_{it} into output in the case of a fully-efficient firm. The second and the third components represent the random and stochastic inefficiency factors that capture the difference between the observed output, y_{it}, and $f(X_{it})$, the output that a fully-efficient firm *i* would secure in the absence of uncertainty. Specifically, v_{it} is the random error assumed to be standard normal $N(0, \sigma_v^2)$, and μ_{it} represents the inefficiency effect, which is non-negative and often assumed to follow a half-normal distribution (i.e., $|N(0, \sigma_\mu^2)|$) (Coelli *et al.*, 2005; Kumbhakar and Lovell, 2003).[7] The production function $f(X_{it})$ is commonly assumed to be a Cobb-Douglas production function,[8] which enables us to rewrite (1) as:

[7] Other distributional assumptions for the inefficiency terms have also been used in the literature, including gamma and exponential distributions. The choice of distributions may influence the predicted firm efficiency score, but not efficiency ranking (Coelli *et al.*, 2005, p.252). The z-test or the likelihood ratio statistics can be used to test whether the half-normal SFA model is adequate; see Coelli *et al.* (2005), p.259.

[8] The Cobb-Douglas function is a widely used functional form to specify the relationship between multiple inputs to an output. However, there are other functional forms in common use. For example, Knott and Posen (2005) adopt the Translog function in their SFA model. See Coelli *et al.* (2005) p. 211 for a list of commonly used functional forms.

$$y_{it} = e^{\beta_0}\left(\prod_{p=1}^{m} x_{itp}^{\beta_p}\right)e^{\nu_{it}}e^{-\mu_{it}} \tag{2}$$

We can express Equation (2) in a linear form after taking logarithms on both sides:

$$\log y_{it} = \beta_0 + \sum_{p=1}^{m} \beta_p \log x_{itp} + \nu_{it} - \mu_{it} \tag{3}$$

We illustrate the SFA frontier model (3) in Figure 1(b), which considers the same firms *a* to *g* that appeared in Figure 1(a). The frontier displayed in Figure 1(b) is determined based on maximum likelihood estimation (MLE) of the parameters β, and observations therefore deviate from the frontier as a result of joint effect of the random noise and inefficiency. The constant term β_0 is negative, indicating there is a minimal input requirement before any output can be produced. For firm *a*, the net effect from inefficiency and noise is positive, and therefore firm *a* is above the frontier $f(X_{it})$. For other firms, the net effect is negative and they are below the frontier and inefficient. The inefficiency of firm *i* at time period *t* can then be calculated as:

$$TE_{it} = e^{-\mu_{it}} \tag{4}$$

In implementation, the separation of the inefficiency effect from the random noise effect is made possible by the distributional assumptions, which allow the efficiency index in (4) to be estimated.

The SFA frontier can be contrasted with the DEA frontier in Figure 1(a), where all firms are on or below the frontier, and firms are driven below the frontier only by the inefficiency effect. The positions of firms in the SFA model are determined by both the inefficiency effect and the random error. Firms can temporarily lie above the frontier given the random error ν_{it}. Another important observation is that, since the inefficiency term μ_{it} is a continuous random variable, we will observe

$\mu_{it} = 0$ (so $TE_{it} = 1$) with a zero probability, which means we will not observe any firms on the SFA efficient frontier. In contrast, we can identify at least one firm on the DEA efficient frontier.

The SFA approach requires an assumption on the functional forms of the production function and the inefficiency term, whereas DEA only requires much weaker assumptions on the production possibility set, such as convexity and minimal extrapolation (Banker *et al.*, 1984). SFA recognizes that there may be errors in data or measurement of the underlying efficiency. DEA assumes that there are no errors; therefore, any error will be reflected in the efficiency score. Another weakness of DEA is that it defines the frontier of the most efficient firms within the sample; if the sample is too small, the frontier may not be representative of the potentially most efficient frontier of the industry because of missing observations.

DEA and SFA in the strategy literature

We compare various features of DEA and SFA in Table 1 and describe how these methods have been used in the strategy literature based on an analysis of studies published in three prestigious journals: *Strategic Management Journal*, *Academy of Management Journal*, and *Management Science* listed in Appendix C. In these studies, the most represented industries are the financial and electric utility sectors. This mix of industries is similar to that of the more general DEA literature, which also includes many studies of airlines and healthcare (see, e.g., Chilingerian and Sherman, 2004). Because these industries are highly regulated, publicly available data on outputs and resource inputs are more readily available to researchers than in other industries.

***

[Insert Table 1 about here]

***

In the strategic management field, it is often of crucial interest to seek an explanation for a firm's performance deviation below the leading firms in the industry. One technique applicable to

both the DEA and SFA models is to proceed with the analysis in two stages. In the first stage, the DEA or SFA models calculate the efficiency scores of the firms in the sample. In the second stage, the efficiency scores obtained in the first stage are regressed on a collection of explanatory variables.[9] Eight of the studies presented in Appendix C use efficiency scores as a dependent variable (Delmas and Tokat, 2005; Schefczyck, 1993; Cummins, Weiss and Zi, 1999; Majumdar and Marcus, 2001; Majumdar and Venkataraman, 1998; Knott and Posen, 2005; Wu and Knott, 2006; Knott, Posen and Wu, 2009).

Although commonly used, this two-stage approach is often criticized because assumptions are required for the inferences made in the second stage to be statistically valid. For example, Wang and Schmidt (2002) show that the SFA two-stage approach can generate substantially biased estimates in both stages. To avoid these problems, a recommended option within SFA is to express the inefficiency term μ_{it} as a function of explanatory variables (Battese and Coelli, 1995; Lieberman and Dhawan, 2005). Here the inefficiency term μ_{it} is assumed to follow a truncated-normal distribution with a mean equal to $Z_{it}\delta$ (i.e., $|N(Z_{it}\delta, \sigma_{\mu}^2)|$), where Z_{it} is a vector of explanatory variables and δ is a vector of parameters to be estimated.[10] Thus, the econometric structure of SFA allows for simultaneously estimating the impacts of inputs and exogenous factors on outputs. The two-stage approach for DEA has also been criticized for lacking a sound statistical foundation (Simar and Wilson, 2007), although Banker and Natarajan (2008) show that the two-stage approach for DEA can yield statistically consistent coefficient estimators under certain general distributional assumptions. Johnson and Kuosmanen (2012) further show that the estimators remain statistically

[9] The regression models commonly used in the second stage include the ordinary least square (OLS), censored regression (e.g., probit and Tobit models), truncated regression, and panel data models. See the discussion in Simar and Wilson (2007) and Knott, Posen, and Wu (2011).

[10] This option can now be implemented within standard statistical packages, such as STATA and R.

consistent even when the first-stage input and output variables in DEA are correlated with the second-stage variables in the regression model. However, the precision of the second-stage estimates will decrease when this correlation is high and when the data have high statistical noise. Johnson and Kuosmanen (2012) develop a robust one-stage estimation approach to address some of the limitations of the traditional two-stage approach for DEA.

One advantage of DEA is that it readily allows for multiple outputs, while SFA in the production frontier form described above is most applicable to a single output or aggregate measure, such as sales or value added.[11] In Appendix C, we see that almost all studies that use DEA have multiple outputs. For example, Delmas, Russo, and Montes-Sancho (2008) in their study of the electricity sector, use three outputs that correspond to three different cost and market structures: low voltage sales, high voltage sales and sales for resale. Having multiple outputs, however, also implies that researchers need to be careful in selecting a meaningful set of output variables for their DEA model. Different outputs may represent different conceptualizations and orientations of business performance. Majumdar and Venkataraman (1998) present an excellent example of comparing DEA efficiency scores that are calculated based on different selections of outputs.

Another advantage of DEA is that it is capable of handling inputs and outputs expressed in different measurement units. For example, Dutta *et al.*, (2005) measured firm capabilities by using patent counts weighted by citations as an output, and R&D and marketing expenses as inputs. This allows researchers to measure capabilities and intangible resources, which are often difficult to

[11] The SFA cost efficiency model can accommodate multiple outputs; see Knott and Posen (2005) for an illustration. Other SFA models for multiple inputs and outputs do exist (e.g., Coelli and Perelman, 1999; Banker, Conrad and Strauss, 1986). However, studies have identified several statistical issues regarding these models; see Chap. 10 in Coelli *et al.* (2005) for a further discussion.

approximate in monetary terms. It also allows researchers to conceptualize capabilities as the ability to combine efficiently a number of resources to engage in productive activity (Dutta *et al.*, 2005).

Both SFA and DEA can incorporate panel data to estimate efficiency scores and other parameters. SFA incorporates the panel information via the traditional econometric framework (Aigner et al.1977; Battese and Coelli, 1995).[12] In Appendix C, all the SFA studies use panel data, where cost efficiency is estimated based on a pooled sample across the observation years (Miller and Parkhe, 2001; Dutta *et al.*, 2005; Lieberman and Dhawan, 2005; Knott and Posen, 2005; Wu and Knott, 2006; Knott, Posen, and Wu, 2009).

In DEA there are two different approaches to deal with panel data. The first approach is to calculate the DEA scores year by year. This means that the DEA score of year *t* is determined based on the data of year *t*.[13] This approach is recommended over calculating DEA scores based on multi-year data, because incorporating multi-year data to estimate the frontier would raise the concern that firms may be compared with the best performers operating under different technological conditions (Majumdar and Venkataraman, 1998; Majumdar and Marcus, 2001; Delmas and Tokat, 2005; Delmas *et al.*, 2008; Delmas and Montes-Sancho 2010). The other approach is to calculate productivity change between two consecutive periods in DEA by using the Malmquist productivity index (Färe, Grosskopf, and Norris, 1994; Cummins *et al.*, 1999; Banker, Chang and Natarajan, 2005). The Malmquist index indicates how much of each firm's total factor productivity change from one period to the next is due to frontier shift and how much is due to its efficiency change.

[12] We should note that SFA can work with a cross-sectional sample, but maximum likelihood estimation of the model requires strong statistical assumptions on error components, and efficiency scores cannot be estimated consistently with a cross-sectional sample (Schmidt and Sickles, 1984).

[13] Recently Chen and van Dalen (2009) used a panel vector autoregressive (PVAR) regression model to estimate the correlation between frontiers in different years. They incorporated the PVAR estimations in the first approach of calculating the efficiency scores for panel data.

Finally, for both DEA and SFA, there are a number of software options available. We provide a list of commonly used of software packages in Appendix B.

ILLUSTRATION: JAPAN AND U.S. AUTOMOTIVE SECTORS

Data

To illustrate the DEA and SFA methodologies, we use data from the U.S. and Japanese automotive sectors from 1977 to 1997 (Lieberman and Dhawan, 2005).[14] Performance of the automotive sector has been of significant interest to strategy researchers (Dyer, 1996; Jiang, Tao, and Santoro, 2010), and the U.S. and Japan represented the top motor vehicle producing countries in the world during the sample period (OICA, 2005-2006). The data contain a balanced panel of 11 automobile producers and their input and output variables. We specify one SFA and two DEA models. In the SFA and the first DEA model, we consider two inputs, *capital* and *number of employees*, and one output, *value-added*.[15] Descriptive statistics of the inputs and outputs are provided in Table 2. We also compare the results of our efficiency frontier models with a more traditional profit measure.

***

[Insert Table 2 about here]

***

DEA model

In the DEA model, we consider two inputs, *capital* and *number of employees*, and one output, *value-added*. We apply the DEA model to cross-sectional yearly samples in the period of analysis. For

[14] Data for Japanese companies are from annual issues of the Daiwa Analyst's Guide. The Japanese data are limited to motor vehicle production within Japan; all transplant operations outside of Japan are excluded. The U.S. data are from the companies' annual financial reports and Standard & Poor's Compustat. Details of the calculation of the variables can be found in Lieberman and Dhawan (2005).

[15] *Value-added* represents the monetary value created and retained by the firm, and is calculated as the firm's sales during the fiscal year minus the costs of purchased materials and services.

illustrative purposes, we assume that the production technology exhibits constant returns-to-scale (CRS) in the DEA model. The CRS assumption implies that, regardless of firm size, expansion or reduction of a firm's inputs by a factor will lead to the same proportion of change in the firm's outputs. We assume CRS because the CRS technology is more intuitive when represented in graphics and also because our cross-sectional sample size is relatively small. Methods to impose a variable returns-to-scale assumption in DEA are described in Appendix A. Summary statistics of the efficiency scores are provided in Table 3.

***

[Insert Figure 2 and Table 3 about here]

***

Figure 2 contains a three-dimensional scatter plot of the inputs and output of the 1997 sample, as well as the corresponding efficient frontier (i.e., the shaded plane). The arrows in the figure represent each firm's trajectory to its efficient benchmark on the frontier. As in Figure 1, firms' efficiency scores are calculated as the 'observed value-added' divided by 'efficient value-added.' In this case efficiency scores range from 0.43 to 1, which represents firms on the efficiency frontier. The results show that Honda and Toyota have scores of 1 and are on the 1997 efficient frontier. The frontier under the CRS assumption is a plane containing these two firms and the origin. The efficient output levels for inefficient firms are points on the frontier that have the same input levels but a higher output level. For example, GM's input and output quantities (Capital, Labor, Value-added) in 1997 are: $16.8 billion of capital stock; 627,500 employees; and $6.6 billion of value-added. GM's efficiency score in that year is 0.43. Based on the efficiency score, we can calculate GM's efficient level of output (value-added) as: $6.6 billion / 0 .43 = $15.3 billion.

DEA allows the inclusion of multiple output variables. Similar to the single-output DEA, the multi-output DEA calculates efficiency scores by measuring the distance between a firm and the efficient frontier in the multi-dimensional space. In addition to value-added, we could for example introduce the number of trucks and the number of cars produced. Thus, outputs can be measured in terms of physical quantities or monetary values. When longitudinal panel data are available, we can also graphically depict how the industry efficient frontier changes over different time periods. To measure the productive changes of individual firms, however, we need to account for both the efficiency change (i.e., relative position to the efficient frontiers in two periods) and the frontier movements. Under the DEA framework, a commonly used approach to calculating productivity change is the Malmquist productivity index that measures productivity change over time; see also our earlier discussion about incorporating panel data in DEA.

SFA model

In the SFA model we use *capital stock* and *number of employees* as the inputs and *value-added* as the output. In the inefficiency component of this model, we follow Lieberman and Dhawan (2005) in using *Volume per plant* (Q/N) , *Work-in-progress/sales (*$WIP/sales$) and *Value-added/sales* ($V/sales$) as the explanatory variables for productive inefficiency. Our choice of variables in the SFA model follows Lieberman and Dhawan (2005). All independent variables used in SFA are in natural logarithms. The first variable (Q/N) provides a measure of plant scale, through which we can test whether scale economies at the level of individual manufacturing plants affect firms' efficiency. The second variable $(WIP/sales)$ serves as a proxy measure of factory management skills, as a large WIP/sales ratio implies that firms need to maintain a high inventory level to counter disruptions in production (Lieberman and Demeester, 1999; Lieberman and Dhawan, 2005). The third variable $(V/sales)$ indicates the level of vertical integration. Specifically, the SFA model is formulated as:

$$\log V_{it} = \beta_1 t + \beta_2 \log K_{it} + \beta_3 \log L_{it} + \nu_{it} - U_{it} \quad (3)$$

where the inefficiency effect is a function of the three explanatory variables and a random inefficiency term:

$$U_{it} = \delta_0 + \delta_1 \log\left(\frac{Q}{N}\right)_{it} + \delta_2 \log\left(\frac{WIP}{sales}\right)_{it} + \delta_3 \log\left(\frac{V}{sales}\right)_{it} + \mu_{it} \quad (4)$$

Table 3 tabulates the estimation results of the SFA model. The coefficient of the time variable (β_1) is positive and highly significant, which suggests that firms' efficiency tend to improve over time. The coefficients of the capital and labor variables are both significant; the sum of these two coefficients is larger than 1 (i.e., 1.036), which signifies a small degree of increasing returns to scale, based on the size of the firm as a whole. The coefficient for plant scale (δ_1) is negative and also highly significant, suggesting that the auto-makers with larger plants in general have gained higher efficiency (i.e., economies of scale exist at the level of individual plants, as well as for the firm as a whole). These findings are consistent with prior studies and reaffirms that exploiting economies of scale is an important factor in attaining efficiency in the automotive sector (Lieberman and Dhawan, 2005; Lieberman, Lau, and Williams, 1990).

In our results, the coefficients for the WIP/sales and Value-added/sales ratios are both insignificant. These results differ slightly from those in Lieberman and Dhawan (2005), which applies SFA to data from the automotive sector from 1965 to 1997. This might be due to a difference in the timeframe observed since our panel starts in 1977, after Japanese producers had made their major inventory reductions.

DEA, SFA and Profit

The DEA and SFA efficiency scores are summarized in Table 4. The significant correlation between these scores implies that they capture similar notions of firm performance in the case of our automotive sample. But to what degree are these measures of efficiency related to company profitability? To answer this question we used the available data to compute an annual financial rate of return (operating profit/net property plant and equipment) for each of the 11 automotive companies. As we can see in Table 4, this measure of profit is also significantly correlated with DEA and SFA but to a lesser level.

***

[Insert Table 4 about here]

***

In order to further compare these measures, we ranked the 11 firms based on the three alternative performance measures: DEA score, SFA score, and financial return. The results of this ranking are illustrated in Figure 3, where we have aggregated the data by country of origin. The contrast between the US and Japanese producers is striking. For the Japanese producers, the efficiency score rankings are uniformly higher than the profit rankings. The reverse is true of the U.S. producers (except during the U.S. recession of the early 1980s).

***

[Insert Figure 3 about here]

***

How did the Americans sustain these high financial returns despite comparatively low levels of efficiency? Although our analysis cannot answer this question directly, a better understanding of the U.S. political economy of the time can help us understand what happened. One major reason relates to tariff protection of a key segment – trucks and SUVs. The United States protects its

domestic market of trucks and vans with a 25 percent import tariff while regular cars only face a tariff of 2.5 percent (Ikenson, 2003). This protection allowed the Big Three US manufacturers to dominate the US market with more than 85 percent of pickup truck sales and benefit from high profit margins (Ikenson, 2003). These margins were particularly high from the mid-1980s to the early 2000s, when oil prices remained low favoring the development of the market for non-fuel efficient vehicles.

While the US car manufacturers were protected by high trade tariffs, Japanese automakers invested in efficiency improvements. In their comparison of a U.S. and a Japanese car plant, Abegglen and Stalk (1985: 105) show that U.S. plants needed 250 percent as many employees as the Japanese one to make a vehicle. The crucial difference was that the Japanese car manufacturers developed lean production systems whereas the U.S. car manufacturers were slow to do so (Womack, Jones, and Roos, 1990). In the 2000s, when high gas prices shifted the market towards more fuel efficient cars, the big three automobile manufacturers which had specialized in building trucks, were unable to adequately respond.[16]

In general, our application of DEA and SFA to a panel data set on 11 US and Japanese automobile producers illustrates several features of frontier methodologies, as well as issues to be considered in their application. Most fundamentally, we show how the methods characterize the distance of firms from the frontier defined by their most efficient competitors (typically, Toyota and Honda in most years of our sample). The automotive example demonstrates the ability of DEA to simultaneously accommodate multiple performance measures, and the ability of SFA to incorporate

[16] To further explain the patterns in figure 3, we regressed the company profit rates and efficiency scores on a series of variables, including the firm's degree of focus on trucks (measured as annual production of trucks and SUVs as a proportion of total vehicle output). These regressions showed that firms' profit rates were strongly and positively linked to truck production, whereas the DEA and SFA efficiency scores were not. Thus, truck and SUV production was unrelated to efficiency but highly related to profitability, particularly for the American companies.

tests of hypotheses about sources of advantage without resorting to a separate stage of statistical analysis. The example shows how corporate performance based on efficiency differs from profitability based on pricing power derived from barriers to entry.

CONCLUSION

Performance measurement is at the heart of strategic management research. This paper has provided an overview on the frontier methodology as a tool for performance measurement by strategic management scholars. Specifically, we have introduced the two most prevalent frontier methodologies, *Stochastic Frontier Analysis* and *Data Envelopment Analysis*, and offered a comparative discussion regarding the strengths and limitations of these two approaches.

The DEA and SFA methodologies provide a way of characterizing competitive advantage in an industry. Both methods focus on firms' efficiencies in converting inputs to outputs, and both allow researchers to identify top performing firms, which lie on or near the industry best practice frontier. DEA and SFA estimate the best-practice frontier and quantify the gap between the observed firm and the frontier (i.e., inefficiency). The efficiency score of a firm is defined as the ratio between the firm's present performance and the performance that the firm would have achieved if it were fully-efficient, based on the estimated frontier.

By providing a technique to assess efficiency, the frontier methods described in this paper enable an understanding of firm performance that is deeper than a mere comparison of company profits. Indeed, as we demonstrate, the major American automotive companies exhibited relatively strong financial performance in the 1980s and 1990s, even though their Japanese rivals maintained higher efficiency. But when the market segments that sustained the US producers— trucks and

SUVs—contracted in the wake of increasing oil prices, the financial performance of the US companies collapsed, leading ultimately to the bankruptcy of General Motors (GM) and Chrysler.

We see a great potential for the use of frontier methodologies in strategy research. Two different views on the sources of corporate profits have dominated strategy research: the industry view and the firm-efficiency view (McGahan and Porter, 1999). In the industry view, industry structure drives profit while in the efficiency view while in the efficiency view companies achieve profits in a line of business when they operate more efficiently than their competitors. Production frontier methodologies allow researchers to assess the efficient use of resources within the existing industry structure. Our auto example illustrates both types of effects: US producers earning high returns by dominating attractive (albeit protected) market segments, and Japanese producers, such as Toyota and Honda, earning their returns through greater efficiency.

The frontier methodologies have proved particularly useful when firm performance is characterized by multiple dimensions with different units of analysis. A key strength of DEA lies in its capability to simultaneously incorporate multiple inputs and outputs, a requirement for analysis of many industries and for studies that seek to incorporate non-financial measures of performance. It also allows the incorporation and comparison of variables with different units (such as, for example, number of employees, tons of input, and dollars of profit). By comparison, SFA is most applicable when multiple outputs can reasonably be aggregated into a single measure, or when price and quantity data are available for inputs and outputs so that a cost frontier model can be estimated. For problems with limited dimensions, the methods can provide an intuitive graphical interpretation of the efficient frontier in the industry and the sampled firms' relative distances.

DEA and SFA have wide potential applications in strategic management research. They provide vehicles for characterizing performance in ways that go beyond conventional analysis of

common financial measures. For example, the ability of DEA to deal with multiple outputs and different units may be particularly useful for resource based view (RBV) analyses of heterogeneous collections of resources, including physical capacities such as capital and machinery, as well as intangible properties such as technological know-how and managerial skills. In RBV, sustained competitive advantage is resulted from leveraging the organizational resources and capabilities that are valuable, rare, inimitable, and non-substitutable (Barney, 1991; Newbert, 2008). What constitutes 'valuable and inimitable resources' depend on both the distribution of the critical resources in the industry and the standing of a firm among its competitive peers. The ability of DEA to include such resources in a comparative way is well suited for RBV analyses when suitable measures are available. DEA and SFA can also facilitate the development of newer areas of strategic management research. For example, the field of frontier methodologies has seen emerging extension to topics such as the measurement of eco-efficiency and corporate social performance, where some of the outputs may be undesirable (such as pollution, labor issues, etc.; see Reinhard, Lovell, and Thijssen, 2000; Chen and Delmas, 2011, 2012). We encourage strategic management researchers to exploit frontier methodologies to explore agendas in these and other topic areas.

REFERENCES

Abegglen J, Stalk G. 1985. *Kaisha: The Japanese Corporation.* New York: Basic Books Inc.

Aigner DJ, Chu S. 1968. On estimating the industry production function. *American Economic Review* **58**(4): 826–839.

Aigner DJ, Lovell CAK, Schmidt P. 1977. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* **6**(1): 21–37.

Banker RD, Chang H, Natarajan R. 2005. Productivity change, technical progress, and relative efficiency change in the public accounting industry. *Management Science* **51**(2): 291–304.

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

Banker RD, Conrad, RF, Strauss RP. 1986. A comparative application of data envelopment analysis and translog methods: an illustrative study of hospital production. *Management Science* **32**(1), 30–44.

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

Barney J. 1991. Firm resources and sustained competitive advantage. *Journal of Management* **17**(1): 99–120.

Battese G, Coelli T. 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics* **20**(2): 325–332.

Bogetoft P, Otto L. 2001. *Benchmarking with DEA, SFA, and R.* Springer: New York, NY.

Charnes A, Cooper W, Rhodes E. 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research* **2**(6): 429–444.

Chen C-M, Delmas M. 2011. Measuring corporate social performance: an efficiency perspective. *Production and Operations Management* **20**(6): 789–804.

Chen C-M, Delmas M. 2012. Measuring eco-inefficiency: a new frontier approach. *Operations Research,* **60**(5): 1064–1079.

Chen C-M, van Dalen J. 2009. Measuring dynamic efficiency: Theories and an integrated methodology. *European Journal of Operational Research* **203**(3): 749–760

Cooper W, Ruiz JL, Sirvent I. 2011. Chapter 4: Choices and Uses of DEA Weights. In *Handbook on Data Envelopment Analysis,* Cooper WW, Seiford LM, Zhu J (eds). Springer: New York: 93–126.

Chilingerian JA, Sherman HD. 2004. Health Care Applications: from hospitals to physicians, from productive efficiency to quality frontiers. In *Handbook on Data Envelopment Analysis*, Cooper WW, Seiford LR, Zhu J (eds). Kluwer Academic Publisher: Boston; 299–321.

Coelli T, Rao DSP, O'Donnell CJ, Battese GE. 2005. *An Introduction to Efficiency and Productivity Analysis*. Springer Verlag: New York, NY.

Coelli T, Perelman S. 1999. A comparison of parametric and non-parametric distance functions: with application to European railway. *European Journal of Operational Research* **117**: 326–339.

Combs J, Crook T, Shook C. 2005. The dimensionality of organizational performance and its implications for strategic management research. In Ketchen D.B. (Ed.), *Research Methodology in Strategy and Management* **2**; 259–286.

Cummins J, Weiss M, Zi H. 1999. Organizational form and efficiency: The coexistence of stock and mutual property-liability insurers. *Management Science* **45**(9): 1254–1269.

Delmas M, Montes-Sancho M. 2010. Voluntary agreements to improve environmental quality: symbolic and substantive cooperation. *Strategic Management Journal* **31**: 575–601.

Delmas M, Tokat Y. 2005. Deregulation, governance structures, and efficiency: the U. S. electric utility sector. *Strategic Management Journal* **26**(5): 441–460.
Delmas M, Russo M, Montes-Sancho M. 2008. Deregulation and environmental differentiation in the electric utility industry. *Strategic Management Journal* **28**(2): 189–209.
Dess G, Robinson Jr R. 1984. Measuring organizational performance in the absence of objective measures: the case of the privately-held firm and conglomerate business unit. *Strategic Management Journal* **5**: 265–273.
Douglas W, Judge Jr. R. 2001. Total quality management implementation and competitive advantage: the role of structural control and exploration. *Academy of Management Journal* **44**(1): 158–169.
Durand R, Vargas V. 2003. Ownership, organization, and private firms' efficient use of resources. *Strategic Management Journal* **24**(7): 667–675.
Dutta S, Narasimhan O, Rajiv S. 2005. Conceptualizing and measuring capabilities: methodology and empirical application. *Strategic Management Journal* **26**: 277–285.
Dyer JH. 1996. Specialized supplier networks as a source of competitive advantage: Evidence from the auto industry. *Strategic Management Journal* **17**(4): 271–291.
Godfrey P, Merrill C, Hansen J. 2009. The relationship between corporate social responsibility and shareholder value: an empirical test of the risk management hypothesis. *Strategic Management Journal* **30**(4): 425-445.
Hillman AJ, Keim GD. 2001. Shareholder value, stakeholder management, and social issues: what's the bottom line? *Strategic Management Journal* **22**(2): 125–139.
Hjalmarsson L, Kumbhakar SC, Heshmati A. 1996. DEA, DFA and SFA: a comparison. *Journal of Productivity Analysis* **7**(2): 303–327.
Howard L, Miller J. 1993. Fair pay for fair play: Estimating pay equity in professional baseball with data envelopment analysis. *Academy of Management Journal* **36**(4): 882–894.
Hull C, Rothenberg S. 2008. Firm performance: The interactions of corporate social performance with innovation and industry differentiation. *Strategic Management Journal* **29(**7): 781–789.
Ikenson, D. J. 2003. The Big Three's shameful secret. Cato Institute Commentary. July 6. http://www.cato.org/publications/commentary/big-threes-shameful-secret [21 July 2013].
Jiang RJ, Tao QT, Santoro MD. 2010. Alliance portfolio diversity and firm performance. *Strategic Management* Journal **31**(10): 1136–1144.
Johnson AL, Kuosmanen T. 2012. One-stage and two-stage DEA estimation of the effects of contextual variables. *European Journal of Operational Research* **220**(2): 559–570.
Knott A, Posen H. 2005. Is failure good? *Strategic Management Journal* **26**(7): 617–641.
Knott M, Posen H, Wu B. 2009. Spillover asymmetry and why it matters. *Management Science* **55**(3): 373–388.
Kumbhakar SC, Lovell CAK. 2003. *Stochastic Frontier Analysis*. Cambridge University Press: New York, NY.
Lieberman MB, Lau LJ, Williams MD. 1990. Firm-level productivity and management influence: a comparison of US and Japanese automobile producers. *Management Science* **36**(1): 1193–1215.
Lieberman MB, Demeester L. 1999. Inventory reduction and productivity growth: linkages in the Japanese automotive sector. *Management Science* **45**: 466–485.
Lieberman MB, Dhawan R. 2005. Assessing the resource base of Japanese and U.S. auto producers: a stochastic frontier production function approach. *Management Science* **51**(7): 1060–1075.
McGahan, AM., Porter, ME. 1999. The persistence of shocks to profitability: comparing the market-structure and Chicago views. *Review of Economics and Statistics* **81**(1): 43–53.

Majumdar S. 1998. On the utilization of resources: perspectives from the US telecommunications industry. *Strategic Management Journal* **19**(9): 809–831.
Majumdar S, Marcus A. 2001. Rules versus discretion: The productivity consequences of flexible regulation. *Academy of Management Journal* **44**(1): 170–179.
Majumdar S, Venkataraman S. 1998. Network effects and the adoption of new technology: evidence from the US telecommunications industry. *Strategic Management Journal* **19**(11): 1045–1062.
Miller S, Parkhe A. 2001. Is there a liability of foreignness in global banking? An empirical test of banks' X-efficiency. *Strategic Management Journal* **23**(1): 55–75.
Meeusen W, van den Broeck J. 1977. Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review* **18**(2): 435–444.
Murillo-Zamorano LR. 2004. Economic efficiency and frontier techniques. *Journal of Economic Surveys* **18**(1): 33–77.
Newbert SL. 2008. Value, rareness, competitive advantage, and performance: a conceptual-level empirical investigation of the resource-based view of the firm. *Strategic Management Journal* **29**: 745–768.
OICA. 2005-2006. World Motor Vehicle Production. International Organization of Motor Vehicle Manufacturers. http://oica.net/category/production-statistics [20 July 2013].
Porter ME. 1985. *Competitive Advantage*. Free Press: New York.
Ray G, Barney J, Muhanna W. 2004. Capabilities, business processes, and competitive advantage: choosing the dependent variable in empirical tests of the resource-based view. *Strategic Management Journal* **25**: 23–37.
Reinhard S, Lovell CAK, Thijssen GJ. 2000. Environmental efficiency with multiple environmentally detrimental variables; estimated with SFA and DEA. *European Journal of Operational Research* **121**(2): 287–303.
Schefczyk M. 1993. Operational performance of airlines: an extension of traditional measurement paradigms. *Strategic Management Journal* **14**(4): 301–317.
Schmidt P, Sickles RC. 1984. Production frontiers and panel data. *Journal of Business & Economic Statistics* **2**(4): 367–374.
Simar L, Wilson PW. 2007. Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics* **136**(1): 31–64.
Thursby JG, Thursby MC. 2002. Who is selling the ivory tower? Sources of growth in university licensing. *Management Science* **48**(1): 90–104.
Waddock S, Graves S. 1997. The corporate social performance-financial performance link. *Strategic Management Journal* **18**(4): 303–319.
Wang HJ, Schmidt P. 2002. One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. *Journal of Productivity Analysis* **18**(2): 129–144.
Womack J, Jones D, Roos D.1990. *The Machine that Changed the World*. Harper-Perennial: New York.
Wu B, Knott A. 2006. Entrepreneurial risk and market entry. *Management Science* **52**(9): 1315–1330.

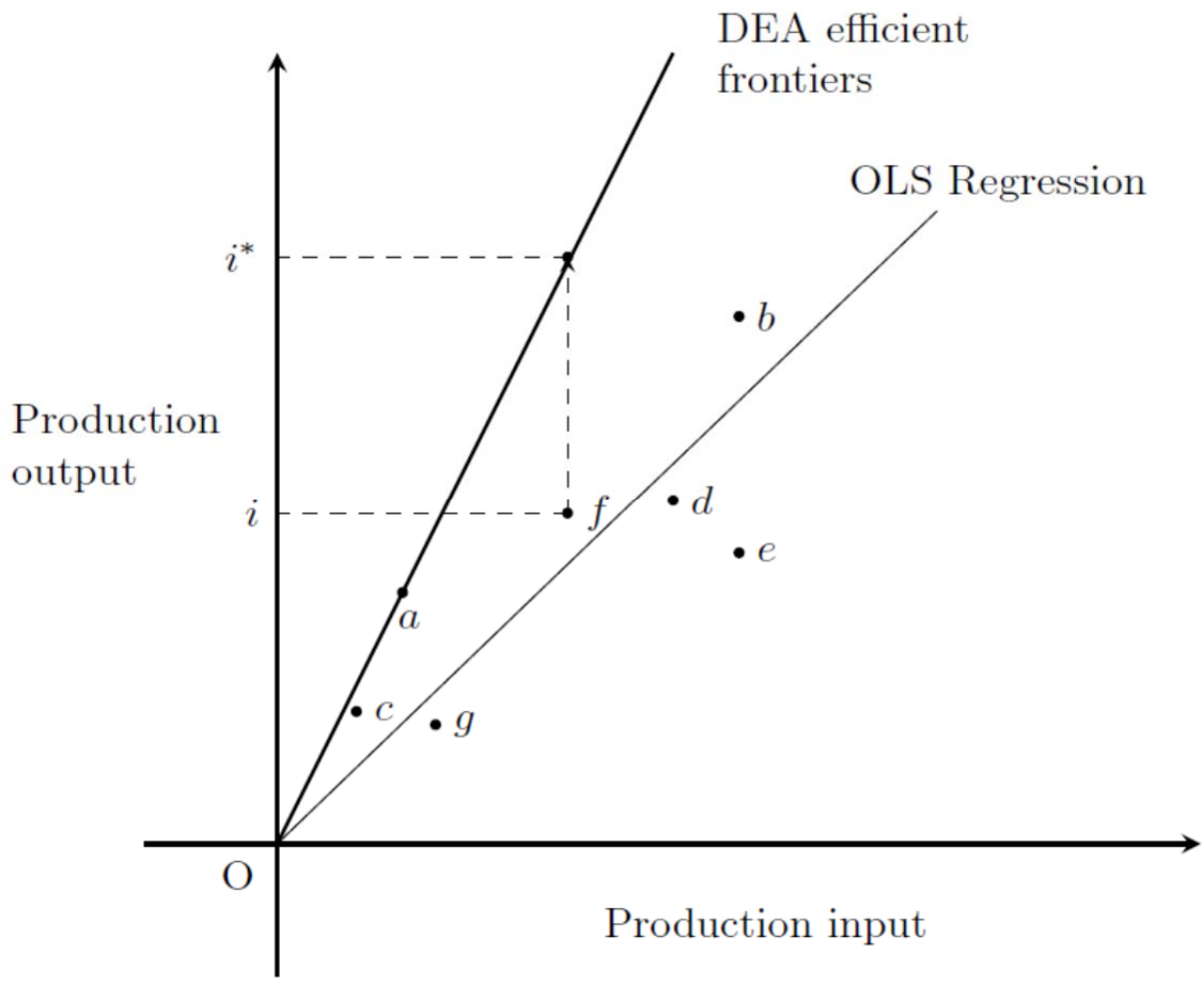

Figure 1(a) Illustration of the DEA frontier

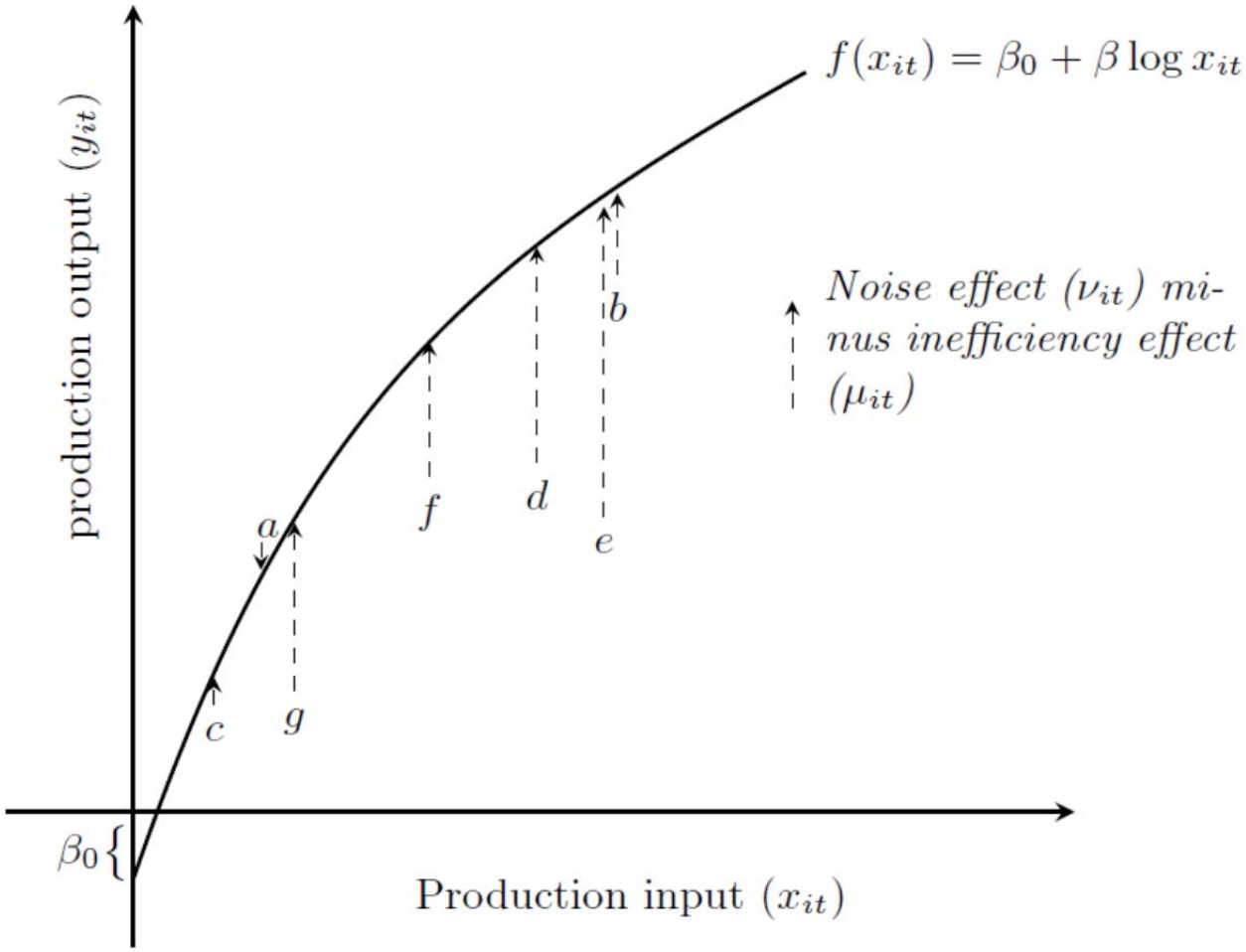

Figure 1 (b) Illustration of the SFA frontier

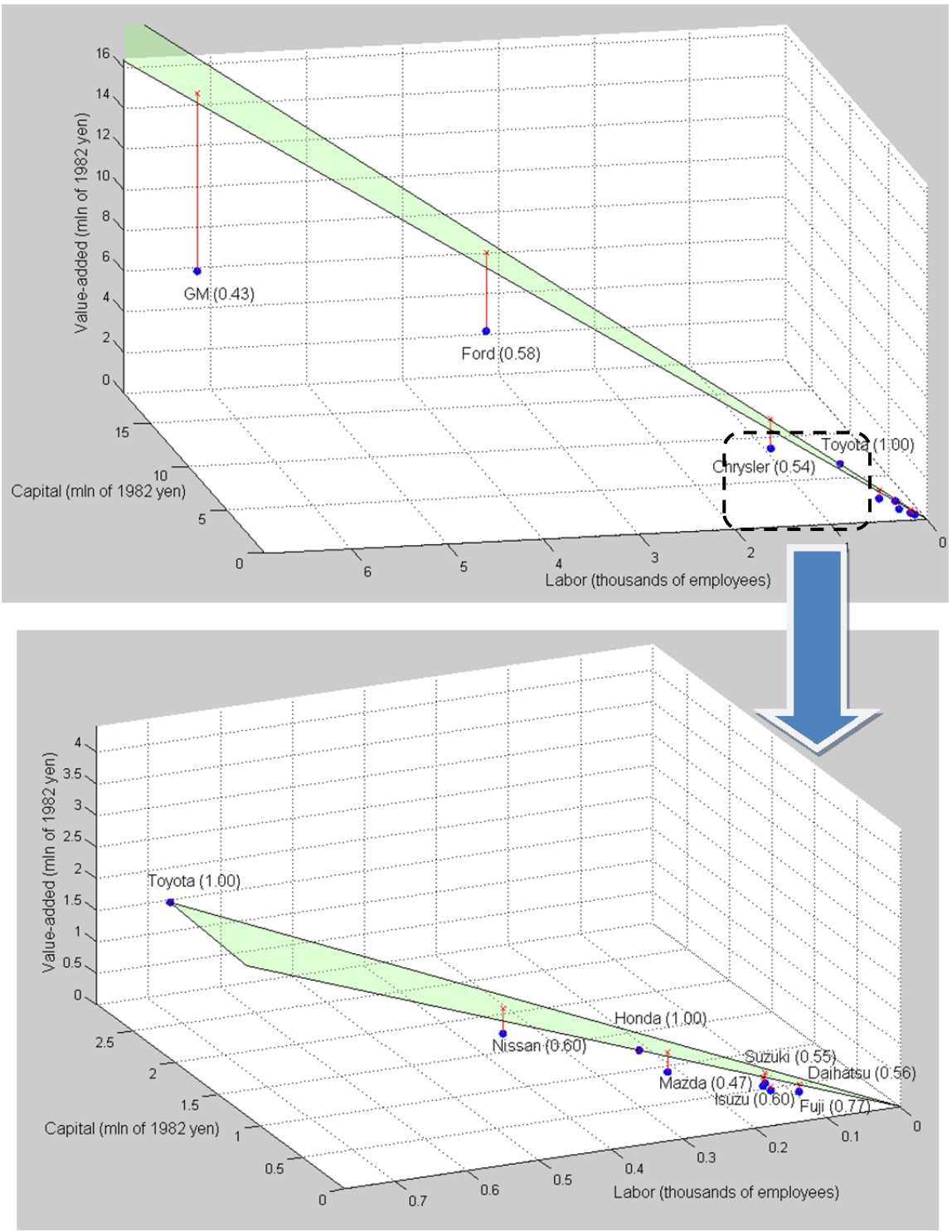

Figure 2 Three-dimensional efficient frontier (year 1997)

Figure 3 Efficiency and profit scores rankings

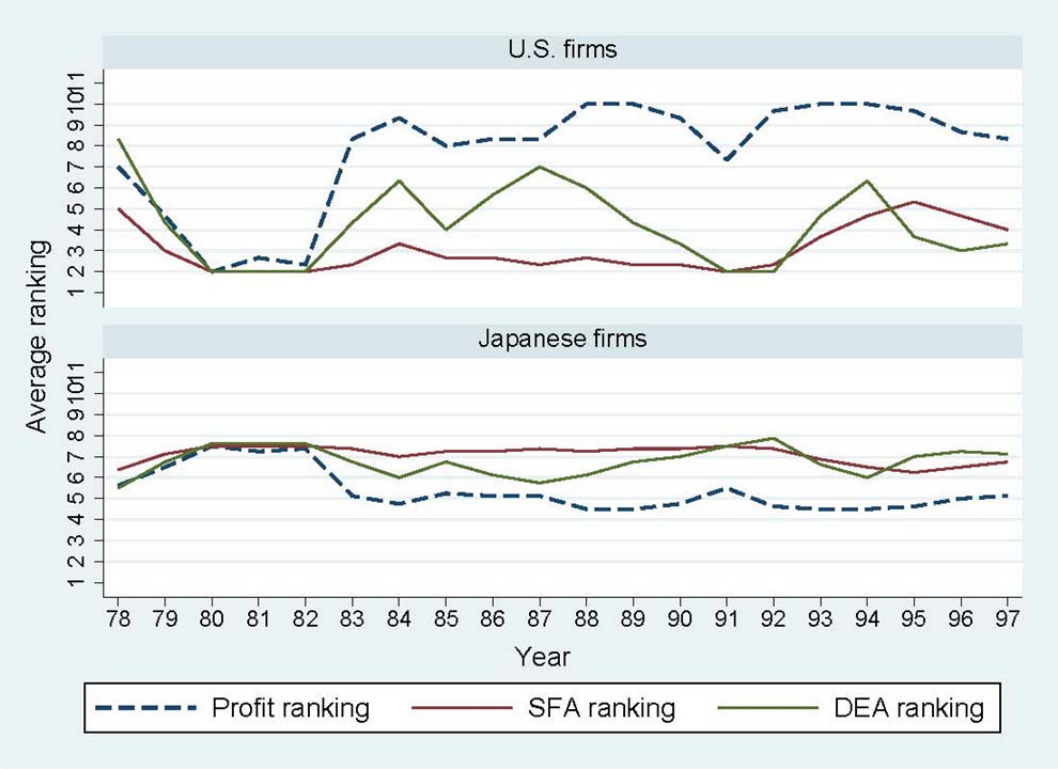

Table 1 Comparing DEA and SFA

Implementation aspects	DEA	SFA
Frontier shape	The DEA frontier is a piece-wise linear surface.	The SFA frontier follows a specific functional form (e.g., Cobb-Douglas, translog)
Applicable to multiple outputs	DEA allows for multiple outputs in the production function. However, including additional outputs may decrease the discrimination power	Production frontier model requires that output be specified as a single measure; cost frontier model can accommodate multiple outputs.
Statistical assumptions	DEA is a deterministic approach and therefore does not require assumptions about the probability distributions of parameters.	SFA requires a priori specification of the model, including the distribution form of the inefficiency term
Sampling errors	The DEA efficiency score is confounded with both statistical noise and inefficiency; it is also more susceptible to the influence of sampling errors and outliers.	SFA incorporates a statistical error term in the formulation.
Panel data structure	Panel data can be incorporated with assumptions on total productivity changes.	SFA can make use of the panel data structure.
Hypothesis tests for the impacts of inputs and exogenous factors on outputs	DEA generates the efficiency score only. To estimate the impacts (coefficients) of inputs and exogenous factors on outputs, it is necessary to fit an auxiliary regression model that uses the DEA efficiency score as the dependent variable.	SFA can estimate the marginal influence of each input and exogenous factors on the output.
Computation	The DEA efficiency score can be easily obtained by solving a number of linear programming problems.	SFA relies on maximum likelihood estimation; ill-structured data or misspecification of the SFA model can lead to numerical problems when estimating the coefficients.

Table 2 Descriptive statistics

	Obs.	Mean	Std. Dev.	Min	Max	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
[1] Value added (V, in 1982 yen)	231	1259776	1909814	42433	7336333	1							
[2] Capital stock (K, in 1982 yen)	231	2530595	4018201	80327.3	1.68E+07	0.96	1						
[3] Number of employees (L)	231	135375.3	223921.1	7890	846000	0.97	0.94	1					
[4] Work-in-progress/sales (WIP/sales)	231	0.02	0.02	0.00	0.12	0.18	0.18	0.22	1				
[5] Value-added/sales (V/sales)	231	0.22	0.10	0.12	0.49	0.15	0.11	0.16	0.41	1			
[6] Production volume per plant (Q/N)	231	313766.3	172986.8	89926.52	842474.6	-0.12	-0.09	-0.10	-0.27	-0.49	1		
[7] Number of trucks produced	231	579512.5	491572.6	88300	2200886	0.86	0.84	0.76	0.13	0.13	-0.12	1	
[8] Number of cars produced	231	1183930	1114576	25532	5285700	0.75	0.65	0.73	0.15	0.19	-0.06	0.63	1

Table 3 Parameters estimates of the SFA model

	Dependent variable: value-added		
	Independent variables	Coefficient	Std. error.
	Production function model		
	Time parameter (β_1)	0.025**	0.004
	No. of employees (β_2)	0.910**	0.053
	Capital stock (β_3)	0.126*	0.058
	Inefficiency model		
	Constant (δ_0)	1.414	0.440
	No. cars produced per plant (δ_1)	-0.819**	0.116
	Value-added/sales (lag 1 year)	-0.018	0.302
(δ_2)			
	WIP/sales (lag 1 year) (δ_3)	-0.045	0.094
	Variance estimates		
	$\sigma^2 = \sigma_u^2 + \sigma_v^2$	0.263	0.109
	$\lambda = \sigma_u^2/\sigma_v^2$	1.69	0.193
	Log likelihood	76.9	
	No. of observations	220	

** significant at the 1% level; * significant at the 5% level

Table 4 Summary of efficiency scores and financial rate of return

- Full sample

Model	Input variables	Output variables	Obs.	Mean	Std. Dev.	Min	Max	[1]	[2]	[3]
[1] DEA	Capital stock, and No. of employees	Value-added	220	0.797	0.161	0.348	1	1		
[2] SFA	**SFA**	Value-added	220	0.886	0.1259	0.412	1	0.562**	1	
	Definition									
[3] Financial rate of returns	operating profit/net property plant and equipment)		220	0.1684	0.1838	-0.53	0.86	0.439**	0.294**	1

*,** Correlation coefficients significant at the 5%, 1% significance level, respectively.

APPENDIX A: MATHEMATICAL FORMULATIONS OF DATA ENVELOPMENT ANALYSIS

(5) Max θ_1^{output}	(6) Min θ_1^{input}
Subject to $\sum_{j=1}^{n} \lambda_j x_{ji} \leq x_{1i}$ for i=1,...,m	Subject to $\sum_{j=1}^{n} \lambda_j x_{ji} \leq \theta_1^{input} x_{1i}$ for i=1,...,m
$\sum_{j=1}^{n} \lambda_j y_{jr} \geq \theta_1^{output} y_{1r}$ for r=1,...,s	$\sum_{j=1}^{n} \lambda_j y_{jr} \geq y_{1r}$ for r=1,...,s
$\lambda_j \geq 0 \; for \; j = 1, \dots, n$	$\lambda_j \geq 0 \; for \; j = 1, \dots, n$

The DEA formulations are shown in models (5) and (6). In the models x_{ji} and y_{jr} denote the i-th input and r-th output of firm *j*, respectively. Depending on whether the efficiency pertains to inputs or outputs, we have a choice between the output- or input-oriented DEA models. Model (5) is called the *output-oriented* DEA model, because the efficiency score θ_1 is attached to the outputs of firm #1. For firm #1, for example, Model (5) attempts to increase firm #1's outputs by maximizing θ_1^{output}, given the inputs $x_{1i}, i = 1, \dots, m$. An output-oriented inefficient firm can increase its output levels with its current input consumption level (i.e., the optimal value of θ_1^{output} will be larger than 1). The efficiency scores of efficient firms are equal to one.

Model (6) is the *input-oriented* DEA model. In Model (6), the efficiency score θ_1 is attached to the input variables, and the objective function seeks to reduce inputs by minimizing θ_1^{input}, given a firm's current output level. An input-oriented inefficient firm can reduce its input usage while maintaining its current output level (i.e., the optimal value of θ_1^{input} is less than 1). The efficiency scores of efficient firms in the input-oriented model are also equal to one.

We assume that the production technology in Models (5) and (6) has constant returns-to-scale (CRS), which implies that the marginal rate of transformation between inputs and outputs is constant. Therefore the optimal value of θ_1^{output} and θ_1^{input} in the CRS model are reciprocals of each other (i.e., $\theta_1^{output} = 1/\theta_1^{input}$). The variable returns-to-scale (VRS) DEA model can be implemented by adding an additional constraint $\sum_{j=1}^{n} \lambda_j = 1$ to the CRS DEA model (Banker *et al.*, 1984). This additional constraint makes the

VRS efficiency score closer to one than the CRS score and thus firms are closer to the efficient frontier in the VRS model.

Solving Models (5) or (6) will yield the efficiency score of one firm. Thus, to obtain the efficiency scores for all n firms, we need to solve the model for n iterations, and in each iteration the constants on the right-hand-side of the constraints are updated.

APPENDIX B: SOFTWARE PACKAGES FOR DEA AND SFA

Below is a partial list of the software programs for implementing DEA and SFA models. Since DEA models are linear programming problems, the DEA algorithm can be implemented in most optimization programs with the basic programming functionality.

DEA	Developer	Website
FEAR: Frontier Efficiency Analysis with R*	Paul Wilson	http://www.clemson.edu/economics/faculty/wilson/Software/FEAR/fear.html
DEAFrontier (add-in for Microsoft Excel)	Joe Zhu	http://www.deafrontier.net/
DEAP*	Tim Coelli	http://www.uq.edu.au/economics/cepa/software.htm
Frontier Analyst	Banxia software	http://www.banxia.com/frontier/
Limdep	Econometric Software, Inc.	http://www.limdep.com/
SFA		
R*	Freeware	http://www.r-project.org/
Stata	Stata Corp.	http://www.stata.com/
Limdep	Econometric Software, Inc.	http://www.limdep.com/
SAS	SAS Institute Inc.	http://www.sas.com/

***** see Bogetoft and Otto (2011) for detail about implementing SFA and DEA in R

APPENDIX C. STUDIES UTILIZING DEA & SFA IN THE STRATEGIC MANAGEMENT RESEARCH (SMJ, AMJ, AND MGMT SCI ARTICLES)

Authors/ Journals	Usage *	Regression model	Sample size and type	Types of inputs	Type of outputs	Sample	Avg. efficiency	Std dev.
DEA Studies								
Schefczyk (SMJ 1993).	A,B	OLS regression	N=15; cross-section	(1) Aircraft capacity, (2) operating costs, (3) Non-flight assets.	(1) Passenger-related revenue, (2) Non-passenger revenue.	Airlines	0.878	0.097
Howard and Miller (AMJ 1993).	C	N/A	N=443; cross-section	14 baseball performance indicators	Players' salary	Major league baseball players.	N/A	N/A
Cummins, Weiss, and Zi (Mgmt Sci 1999)	A,C	OLS regression	N=417(pooled); 10 years panel	(1) Labor, (2) Business services, (3) Debt capital, (4) Equity capital.	(1) Present value of losses, (2) Total invested assets.	Insurers	0.901 (pooled)	--
Majumdar (SMJ 1998)	C	N/A	N=39. Cross-section.	(1) total number of switches, (2) total number of access lines, and (3) total number of employees.	(1) local revenues, (2) toll revenues, and (3) other revenues.	Telecom firms	0.903	0.104
Majumdar and Venkataraman (SMJ 1998)	B	Tobit regression	N=40. 5-year panel	(1) number of switches, (2) number of lines, and (3) number of employees	(1) local revenues, (2) toll revenues, and (3) access and misc. revenues.	Telecom firms	5.12	1.34
Majumdar and Marcus (AMJ 2001)	A	Tobit regression	N=150, cross-section	(1) production expenses, (2) transmission expenses, (3) distribution expenses, (4) the total number of	(1) total sales, (2) dispositions of energy in megawatt hours	Electric utility firms	0.78	.158

Authors/ Journals	Usage *	Regression model	Sample size and type	Types of inputs	Type of outputs	Sample	Avg. efficiency	Std dev.
				employees, (5) amount of purchased power.				
Thursby and Thursby (Mgmt Sci 2002)	C	N/A	N=64 from 1993 to1997 (balanced)	(1) faculty size, (2) research funds, (3) the number of full-time equivalent personnel	Number of patents	Universities	N/A	N/A
Durand and Vargas (SMJ 2003)	A	ANCOVA and MANCOVA	N=162. Cross-section	(1) total fixed assets; expenditures in (2) R&D, (3) marketing, and (4) education	(1) gross profits; (2) sales	Printing, auto parts, chemical firms	0.552	0.253
Delmas and Tokat (SMJ 2005)	A	Tobit regression	N=707 (pooled). 4-year panel.	(1) labor cost, (2) plant value, (3) production expenses, (4) transmission expenses, (5) distribution expenses, (6) sales, (7) admin. and general expenses, (8) electricity purchases	(1) low-voltage sales (2) high voltage sales (3) sales for resale	Electric utility firms	0.86	0.18
Delmas, Russo, and Montes-Sancho (SMJ 2008)	B	OLS regression	N=177; 3-year panel.	Same as above	(1) low-voltage sales (2) high voltage sales (3) sales for resale	Electric utility firms	0.926	0.09
Delmas and Montes-Sancho (SMJ 2010)	B	Two-stage logit model	N=132. 6-year panel	Same as above	(1) low-voltage sales (2) high voltage sales (3) sales for resale	Electric utility firms	0.88	0.15
SFA Studies								
Miller and	C	SFA	N=1300; 7-	(1) total loans, (2)	Profits	Banks	N/A	N/A

Authors/ Journals	Usage *	Regression model	Sample size and type	Types of inputs	Type of outputs	Sample	Avg. efficiency	Std dev.
Parkhe (SMJ 2001)			year panel	earning assets, (3) labor, (4) physical capital, (5) funds and deposits				
Dutta, Narasimhan, and Rajiv (SMJ 2005)	C	SFA	N=64; 9-year panel	(1) Cumulative R&D expenditure, (2) Cumulative marketing expenditure	R&D patent counts weighted with citations	Semi-conductor firms	N/A	N/A
Lieberman and Dhawan (Mgmt Sci 2005)	C	SFA	N=11; 32-year panel	(1) Employee, (2) Capital	Economic value-added	Auto producers	N/A	N/A
Knott and Posen (SMJ 2005)	B	SFA (cost efficiency model)	170,859 firm–year observations over 14 years	(1) labor price, (2) physical capital price, (3) mortgage loans, (4) non-mortgage loans, (5) investment securities	Total costs	Banks	0.81	N/A
Wu and Knott (Mgmt Sci 2006)	B	SFA (cost efficiency model)	Same as above.	Same as above.	Total costs	Banks	0.81	N/A
Knott, Posen, and Wu (Mgmt Sci 2009)	A	SFA (cost efficiency model)	Same as above.	Same as above.	Total costs	Banks	0.81	N/A

*A: Efficiency scores used as the dependent variable in the regression model. B: Efficiency scores used as the independent variable in the regression model. C: Efficiency scores are not used in regression.