EVALUATING THE EFFECT OF OPERATIONAL CONDITIONS AND PRACTICES ON WAREHOUSE PERFORMANCE

Andrew L. Johnson

Department of Industrial and Systems Engineering
Texas A&M University
College Station, TX 77843 USA

Abstract

Analyzing warehouse performance across different environments is critical to improving overall productivity and reducing costs. Although two-stage DEA estimators have been shown to be statistically consistent, the finite sample bias of DEA in the first stage carries over to the second-stage regression, which causes bias in the estimated coefficients of the contextual variables. The bias is particularly severe when the contextual variables are correlated with inputs. To address this shortcoming, we apply insights from Johnson and Kuosmanen (2010), who demonstrate that DEA can be formulated as a constrained special case of the Convex Nonparametric Least Squares (CNLS) regression to develop a new semi-parametric one-stage estimator. The new model is applied to a set of warehouses to illustrate its performance.

1. Introduction

Analyzing warehouse productivity involves identifying a set of inputs and outputs and estimating efficiency. Often, there are variables that are contextual to the warehouse process in the sense that they are neither inputs nor outputs, i.e. they characterize the production environment or the operational practices. These variables are observed (or determined) by the producer while the inputs are selected and thus may be correlated with inputs. Two-stage methods are commonly used to investigate the relationship between efficiency and the contextual variables. In the first stage, efficiency is estimated and in the second stage the estimates are regressed against the contextual variables. Estimation issues have been identified for the two-stage method when input variables (in an output oriented model) are correlated with contextual variables; however, the two-stage method

1

is still used when the parametric assumption in the first stage is believed to be overly restrictive and nonparametric efficiency estimation methods are preferred. This paper describes the use of a single-stage model for estimating the influence of contextual variables on efficiency estimates.

Understanding the relationship of contextual variables to performance levels is helpful for several reasons. For instance, when finding that characteristics of the operational environment are significant drivers of efficiency levels, it indicates the existence of factors beyond managerial control that impact the level of efficiency. When it is discovered that the operational policy is a significant driver of performance, it indicates that management should adopt these policies when considering new or expanded production facilities. Thus, the results of an accurate analysis of contextual variables are beneficial in identifying sources of heterogeneity between production processes and in identifying improvement strategies.

An early discussion of contextual variables and their effect on efficiency estimation was presented in Hall and Winsten (1959). Subsequent work then developed other approaches to address the issue, but note that contextual variables are addressed using a variety of terms, e.g., environmental, exogenous, non-discretionary, efficiency factors, attributes, and practices, each with subtle distinctions. The investigation of contextual variables using a two-stage method was pioneered by Timmer (1971), while Ray (1988) (1991) extended the approach by using the DEA estimator in the efficiency estimation and Simar et al. (1994) extended the two-stage method to stochastic frontier analysis (SFA) efficiency estimator in the first stage. There has also been extensive effort to incorporate contextual variables in the product models, particularly in the nonparametric setting. A list of the relevant papers can be found in Simar and Wilson (2007). If the production frontier assumes a parametric form, a single-stage method for identifying the frontier and estimating the impact of contextual variables has been developed by (Stevenson 1980). This allows efficiency and the relationship between efficiency and the contextual variables to be estimated in a single stage. Wang and Schmidt (2002) performed Monte Carlo simulation analysis to investigate the bias induced by estimating in two-stages. The result of their analysis indicated the bias can be quite significant and it is the basis for their recommendation of single-stage methods in parametric models.

There has been an ongoing debate regarding the use of parametric or nonparametric models. Recently, Simar and Wilson (2007) and Banker and Natarajan (2008) argued in favor of the two-stage method when nonparametric efficiency models are used to estimate efficiency in the first stage. Simar and Wilson emphasize the importance of a coherent data generation process and the issues of serial correlations among the efficiency estimates generated by nonparametric efficiency models. The authors develop a two-stage bootstrapping method to address the serial correlation issue inherit in a two-stage model and note that the slow convergence rate of non-parametric efficiency estimators limits the approaches available for asymptotic analysis in these models.

Banker and Natarajan have also looked at two-stage models applying DEA to estimate efficiency followed by ordinary least squares (OLS) or maximum likelihood estimation in the second stage. Under assumptions similar to Gstach (1998) and Greene

(1980) regarding the structure of noise, Banker and Natarajan showed that DEA followed by either OLS or maximum likelihood estimation produces consistent estimates in two-stage models. The authors benchmarked the two-stage approach against both one-stage and two-stage methods. Their results indicate that if the exact functional form of the underlying data generation process is used in the parametric methods, a one-stage parametric model outperforms the non-parametric two-stage methods. However, if the translog or Cobb-Douglas functional forms are used, the misspecification of the functional form introduces significant error, which makes a two-stage non-parametric method preferable. Banker and Natarajan's results provide evidence of the superior performance of two-stage nonparametric models.

Both sets of authors clearly stated that the two-stage model is only valid if the input variables are uncorrelated with the contextual variables. Note, however, that they use a standard production function and thus an output-oriented model. When using an input-oriented model similar assumptions are necessary, namely the output variables are uncorrelated with the contextual variables. This is a strong assumption, given that most contextual variables are observable to the decision-maker when determining either the input or output levels. If the contextual variables are correlated with efficiency and the decision-maker desires to maximize efficiency, the decision-maker must consider the effect of the contextual variables and select the input levels that maximize the production process's technical efficiency. Further, Banker and Natarajan showed that the performance of the two-stage estimate deteriorated significantly as the correlation between inputs and the contextual variables increased.

Two non-parametric methods for modeling contextual variables in a single state have been introduced in Johnson and Kuosmanen (2010). The authors build on the results of Kuosmanen and Johnson (2010) who showed DEA can be formulated as a nonparametric least squares problem. Previously, if nonparametric efficiency estimation was used to estimate efficiency, it was not possible to consistently estimate efficiency and the impact of contextual variables when inputs and contextual variables are correlated. One method is shown to be equivalent to DEA and the other method, called Corrected Convex Non-parametric Least Squares (C^2NLS), is based on a two-stage shifting method inspired by corrected ordinary least squares (COLS). The models including contextual variables are termed 1-stage DEA and C^2NLS in the presence of contextual variables.

The warehousing literature has shown some interest in applying efficiency models to analyze performance, Hackman et al. (2001), De Koster and Warffemius (2005), De Koster and Balk (2008), and Johnson and McGinnis (2010). All of these studies have used various analyses to quantify the effect of various contextual variables. However, all use a two-stage approach and thus their results include the small sample bias induced by using the DEA estimator in the first stage.

The paper is organized as follows. Section 2 introduces the necessary notation and the models that will be used for comparison. Section 3 introduces the one-stage nonparametric model, and Section 4 presents application results to demonstrate the method's superior performance. Section 5 gives some concluding remarks.

2. Models for estimating the effects of contextual variables

Consider the standard multiple input, single output, cross-sectional model in production economics:

$$y_i = f(\mathbf{x}_i) + \varepsilon_i \ \forall i = 1,...,n, \tag{1}$$

where y_i denotes the output of firm i, $f : R_+^m \to R_+$ is the production function that characterizes the production technology, $\mathbf{x}_i = (x_{i1}...x_{im})'$ is the input vector of firm i and represents the deviation of firm i from the frontier. The error term interacts with the production function additively.

Different models of efficiency analysis can be classified according to how the production function f is specified. Parametric models hypothesize a functional form and proceed to estimate the parameters, while nonparametric models assume certain regularity axioms such as monotonicity or regularity. However, monotonicity and convexity are widely considered minimum axiomatic properties, Shephard (1953). On this basis, advocates of nonparametric methods argue against the use of parametric functional form to characterize production.

In a production process, it is often assumed that factors exist which limit the possibility for the process to be efficient. The literature has labeled these variables as environmental, exogenous, non-distrectionary, efficiency factors, attributes, practices, fixed inputs, etc., and all have subtle distinctions. We use the term, contextual variables and denote them as $\mathbf{z}_i = (z_{i1}...z_{ir})'$, where there are r components of the $\mathbf{z}_i$ vector for firm i.

The underlying assumption of methods that consider contextual variables is that both the inputs and the contextual variables will affect the position of the efficient frontier. The two-stage methods implicitly assume the input levels and the contextual variables are uncorrelated, Kumbhakar and Lovell (2000); Wang and Schmidt (2002); and Coelli et al. (2005), among others. Further, using the two-stage method when there is correlation between the input levels and the contextual variables can lead to significant bias in small sample estimates, Banker and Natarajan (2008); We will use the model presented in Johnson and Kuosmanen (2010) to estimate the effects of a contextual variables in a warehouse data set.

Several approaches are available in the productivity literature. However, non-parametric methods suffer from the curse of dimensionality and reduce the reference set identifying a local comparison group. In addition, such approaches do not use the information available in all observations to determine both the shape and the level of the frontier. This loss of information can be critical considering the high data requirements for non-parametric methods. Two examples of this approach are Ruggiero (1996) in which a harsher environment in terms of the contextual variables was determined a priori and a given observation could only be compared to other observations in an environment

no better than its own. A similar alternative was suggested by Daraio and Simar (2007) in which an "optimal" local neighborhood definition in terms of the z variable and a nearest neighbor criterion determined the peer group.

3. One-state nonparametric model including contextual variables

3.1 Least squares formulation

A significant body of work exists for parametric extensions of SFA to address contextual variables (a brief introduction appears in Coelli et al. (2005) section 10.7). However, it has long been held that non-parametric methods such as DEA were fundamentally different then regression based methods such as SFA, see for example Thanassoulis (1993); Cooper et al. (2007). The considerable conceptual and philosophical gap between DEA and the regression-based methods has led to the separate development of methods to analyze contextual variables. Recently, Kuosmanen and Johnson (2010) established a formal connection between DEA and least squares regression by showing that the standard output-oriented DEA model can be equivalently formulated as a nonparametric least squares regression subject to monotonicity and concavity constraints on the frontier and sign constraints on the inefficiency terms for the single output multi-input case and the output-oriented efficiency measure. This finding has at least two significant benefits: 1. it helps to bridge the conceptual and methodological gap and paves the way to a unified framework of frontier analysis; 2. the least squares interpretation provides a platform for new methodological developments. We will use the platform to investigate a warehouse data set via a nonparametric regression-type model which allows the introduction of contextual variables while still maintaining the spirit of the original DEA, and lets the data speak for itself.

Kuosmanen and Johnson consider a non-parametric least squares method subject to continuity, monotonicity, and concavity constraints arising from Hildreth (1954). As mentioned above this method is termed Concave Nonparametric Least Squares (CNLS), Kuosmanen (2006). CNLS is based on the assumptions that the function f to be estimated belongs to the set of continuous, monotonically increasing and globally concave functions. Compared to the kernel regression and spline smoothing techniques, CNLS does not require specification of any smoothing or bandwidth parameters. To estimate the CNLS formulation in the general multi-input setting, Kuosmanen (2008) showed that the set of continuous, monotonically increasing and globally concave functions can be equivalently represented by a family of piece-wise linear functions characterized by the celebrated Afriat's Theorem (Afriat (1967); Afriat (1972)). Applying these insights, the following quadratic programming problem is used to estimate the CNLS measure:

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\varepsilon}}\left\{\sum_{I=1}^{n}\varepsilon_i^2 \left| \begin{array}{l} y_i = \alpha_i + \boldsymbol{\beta}'_i\mathbf{x}_i + \varepsilon_i \ \forall i = 1,...,n; \\ \alpha_i + \boldsymbol{\beta}'_i\mathbf{x}_i \leq \alpha_h + \boldsymbol{\beta}'_h\mathbf{x}_i \ \forall h,i = 1,...,n; \\ \boldsymbol{\beta}_i \geq \mathbf{0} \ \forall i = 1,...,n \end{array} \right. \right\} \tag{2}$$

In formulation (2), the first constraint estimates α_i and $\boldsymbol{\beta}_i$ parameters for each observation; thus n different regression lines are estimated instead of fitting one regression line to the cloud of observed points, as in OLS. The standard OLS problem is obtained by replacing the constraint $\boldsymbol{\beta}_i \geq \mathbf{0} \ \forall i$ by constraints $\alpha_i = \alpha_j, \boldsymbol{\beta}_i = \boldsymbol{\beta}_j \ \forall i, j = 1,...,n$. These n estimated lines can be interpreted as tangent lines to the unknown production function f. The slope coefficients $\boldsymbol{\beta}_i$ represent the marginal products of inputs (i.e., the sub-gradients $\nabla f(\mathbf{x}_i)$). The second constraint imposes concavity by applying a system of Afriat inequalities which are the key to modeling concavity constraints in the general multiple regressor setting. The third constraint imposes monotonicity.

CNLS can be extended to estimate a frontier production frontier in two ways. The first, inspired by Aigner and Chu's (1968) Parametric Programming approach (PP), is to restrict the error terms of the CNLS model to be one-sided; the error terms can thus be interpreted as inefficiency. This sign-constrained variant of the CNLS problem is stated formally as:

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\varepsilon}}\left\{\sum_{I=1}^{n}\varepsilon_i^2 \left| \begin{array}{l} \varepsilon_i \leq 0 \ \forall i = 1,...,n; \\ y_i = \alpha_i + \boldsymbol{\beta}'_i\mathbf{x}_i + \varepsilon_i \ \forall i = 1,...,n; \\ \alpha_i + \boldsymbol{\beta}'_i\mathbf{x}_i \leq \alpha_h + \boldsymbol{\beta}'_h\mathbf{x}_i \ \forall h,i = 1,...,n; \\ \boldsymbol{\beta}_i \geq \mathbf{0} \ \forall i = 1,...,n \end{array} \right. \right\}. \tag{3}$$

Kuosmanen and Johnson prove the hybrid model of PP and CNLS is equivalent to the standard DEA model.

Second, CNLS can be used in a two-stage shifting method, which we term Corrected Nonparametric Least Squares (C[2]NLS). It is a nonparametric variant of the Corrected Ordinary Least Squares (COLS) Winsten (1957); Greene (1980) model in which CNLS replaces the first-stage parametric OLS regression. Like COLS, the C[2]NLS method is implemented in two stages, which can be stated as follows:

Stage 1: Estimate $E(y_i|\mathbf{x}_i)$ by solving the CNLS problem (3). Denote the CNLS residuals by ε_i^{CNLS}.

Stage 2: Shift the residuals analogous to the COLS procedure; the C[2]NLS efficiency estimator is:

$$\hat{\varepsilon}_i^{C2NLS} = \varepsilon_i^{CNLS} - \max_h \varepsilon_h^{CNLS}, \tag{4}$$

where values of $\hat{\varepsilon}_i^{C2NLS}$ range from $[0, -\infty]$ with 0 indicating efficient performance. Similarly, adjust the CNLS intercepts α_i as:

$$\hat{\alpha}_i^{C2NLS} = \alpha_i^{CNLS} + \max_h \varepsilon_h^{CNLS}, \tag{5}$$

where α_i^{CNLS} is the optimal intercept for firm i and $\hat{\alpha}_i^{C2NLS}$ is the C^{2}NLS estimator. Slope coefficients $\hat{\boldsymbol{\beta}}_i$ for C^{2}NLS are obtained directly as the optimal solution to (3).

Note that in the formulation above the error term (or inefficiency term if we are working in a strictly deterministic setting) interacts with the production function in an additive manner. Heteroskedasticity is a common problem that can arise in econometric estimation. Kuosmanen and Johnson prove the consistency of the least squares formulation of the DEA problem. As with OLS, the slope coefficients $\boldsymbol{\beta}_i$ from the nonparametric least squares formulation of DEA are consistent even in the presence of heteroskedasticity. However, the consistency of the efficiency estimates will be affected by heteroskedasticity. Several standard tests in the literature are available to identify if a particular data set has heteroskedastic correlations between variables, e.g., Koenker's version of Breusch-Pagan test, or the White test.

3.1 Models including contextual variables

Now, we can extend the least squares formulation of DEA to include contextual variables, termed the 1-stage DEA model in the presence of contextual variables. We can also C^{2}NLS to handle contextual variables. Starting from a DEA model now formulated as a regression-type optimization problem allows us to incorporate the contextual variables which are often considered in a two-stage analysis into the first stage using any number of methods. The methods used for stochastic frontier analysis (SFA) can be used in this setting, e.g., Kumbhakar and Lovell (2000); Coelli et al. (2005). We assume that traditional interpretation that deviations from the frontier can be explained by non-discretionary variables and inefficiency:

$$\varepsilon_i = \sum_{i=1}^r \delta_i z_i + \theta_i \tag{6}$$

where θ_i is the remaining inefficiency for the ith firm that cannot be explained by contextual variables. This interpretation is important because we can further decompose

θ_i in the spirit of Jondrow et al. (1982) to allow for the measurement of efficiency in the presence of noise. Thus:

$$\min_{\alpha,\beta,\varepsilon} \left\{ \sum_{I=1}^{n} \theta_i^2 \left| \begin{array}{l} \theta_i \leq 0 \ \forall i = 1,...,n; \\ y_i = \alpha_i + \boldsymbol{\beta}_i' \mathbf{x}_i + \mathbf{z}_i \boldsymbol{\delta} + \theta_i \ \forall i = 1,...,n; \\ \alpha_i + \boldsymbol{\beta}_i' \mathbf{x}_i \leq \alpha_h + \boldsymbol{\beta}_h' \mathbf{x}_i \ \forall h,i = 1,...,n; \\ \boldsymbol{\beta}_i \geq \mathbf{0} \ \forall i = 1,...,n \end{array} \right. \right\}. \tag{7}$$

Under a basic set of assumptions the one-stage DEA model can be shown to produce consistent estimators of both θ_i and $\boldsymbol{\delta}_i$. See Johnson and Kuosmanen (2010) for details.

4. Warehouse analysis

Our data set of 216 warehouse records collected over a five-year period is treated as a cross-section, because the technical progress during the time period is believed to be minimal. The data were collected via the iDEAs-W Web site (http://ise.tamu.edu/ideas) and each data record summarizes the performance of a warehouse for a one-year period. iDEAs-W provides a browser-based interface which lets users enter data and receive an efficiency estimate based on the warehouse data collected to that point. Because not all respondents provided data for the contextual variables, the analysis below includes a subset of the data using only the respondents who provided the contextual variable information. A three-input (labor; space; investment), single-output (total lines shipped) model is specified to characterize warehouse performance. For more details regarding the data set or collection process, see Johnson and McGinnis (2010) and Johnson et al. (2010).

For the purposes of this analysis, warehouses attributes beyond the control of the warehouse manager are investigated. The eight contextual variables are: Number of Replenishments; Inventory Turns; Number of Stock-keeping Units (SKUs); SKU Churn; Seasonality; Average Weight per Order; Labor Turnover; and Percentage of Temporary Labor. While some of these variables may seem to be under the control of the warehouse manager, we argue that they are responses to industry demands, or characteristics of the market or the surrounding environment. We define them as:

Number of Replenishments: a count of the number of shipments received from all suppliers. Clearly, the warehouse manager can control this factor, but this would typically be done in response to customer demands, characteristics of the products, or shipping cost. Thus, this variable characterizes the product and its market.

Inventory Turns: number of times the entire inventory of a warehouse is sold and replaced in a year. This variable characterizes the amount of inventory the warehouse must hold to ensure a particular service level and thus is a proxy for the service level requirements of the industry.

Number of SKUs: a count of the number of different types of items stocked in the warehouse. Its use in our analysis distinguishes warehouses that handle a limited set of products in particularly high volumes. This variable characterizes the array of products the warehouse must provide to its customers.

SKU Churn: percentage of SKUs that changes from year to year, calculated as (SKUs dropped last year+ SKUs added last year)/ (beginning total SKUs last year). If new products are frequently introduced and some produced are considered obsolete, it could impact operational efficiency. This variable characterizes the dynamics of the industry which the warehouse serves.

Seasonality: defined as (volume in the peak month / average volume per month), where volume s based on items (pieces, or units). The seasonality of the demand for products has a significant impact of the efficiency of warehouse operations.

Average Weight per Order: size of the orders handled by the warehouse. This characteristic may indicate a particular warehouse can or cannot use particular material handling strategies, thus narrowing operational choices and impacting productivity.

Labor Turnover: percentage of workers that changes from year to year, calculated as (workers who quit or were fired last year+ workers hired last year)/ (beginning number of workers last year). Workforce dynamics can be due to salary limitations, local workforce force population, availability, skills, etc.

Percentage of Temporary Labor: the percentage of direct labor hours performed by workers who either do not work for the firm that owns and operates the warehouse, or are hired with the expectation of working less than 1 year.

The analysis presented next was performed by estimating (7) for each contextual variable separately due to data limitations. Table 1 reports the delta coefficients estimated and their significance level. Note that the number of total replenishments has a statistically significant and negative effect on the total number of lines shipped. On average an increase of 100 additional shipments leads to 2 fewer lines shipped. Note, too, that seasonality and temporary labor has negative effects on the number of lines shipped; these coefficients are significant at the 5 percent level. Increased inventory turns positively effect the number of lines shipped, however, the coefficient is only significant at the 10 percent level. Other factors, i.e. Number of SKUs, SKU Churn, Average Order Weight, and Labor Turnover are insignificant in predicting the total number of lines shipped.

We expect that warehouse operating in industries with high seasonality, requiring high service levels, and using significant amounts of temporary labor will be at a significant efficiency disadvantage while those operating in industries that allow lower inventory levels will have an advantage.

Table 1: List of warehouse attributes investigated; estimate of their effect on lines shipped; and significance level.

Contextual variable	Coefficients and Significance
Number of replenishments	-0.022***
Inventory turns	220,000*
Number of SKUs	0
SKU churn	2,120
Seasonality	-274,000**
Average weight per order	-43.4
Labor turnover	-73,700
Percentage of temporary labor	-20,600**

The significance of the level of the coefficients at the 1, 5, and 10 percent levels is indicated by ***,**, and *, respectively.

5. Conclusion

Comparing warehouse operations in different environments is an important issue when benchmarking warehouse performance. This paper addresses the concern by using a one-stage method introduced by Johnson and Kuosmanen (2010), which jointly estimates both the production function and the effect of the context. The new semi-parametric model estimates the production function consistent with axioms taken from microeconomic theory while estimating first-order parametric effects of the production context.

References

Afriat, S. N. (1967). The Construction of a Utility Function from Expenditure Data. International Economic Review 8: 67-77.

Afriat, S. N. (1972). Efficiency estimation of production functions. International Economic Review 13(3): 568-598.

Aigner, D. and S. Chu (1968). On estimating the industry production function. American Economic Review 58: 826-839.

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

Coelli, T., D. S. P. Rao and G. E. Battese (2005). An introduction to efficiency and productivity analysis. New York, Springer.

Cooper, W. W., L. M. Seiford, and K. Tone (2007). "Data envelopment analysis a comprehensive text with models, applications, references and DEA-solver software." 2nd.

Daraio, C. and L. Simar (2007). Conditional nonparametric frontier models for convex and nonconvex technologies: a unifying approach. Journal of Productivity Analysis 28 (1-2): 13-32.

De Koster, M. B. M. and B. M. Balk (2008). Benchmarking and monitoring international warehouse operations in Europe. Production and Operations Management 17(2): 175-183.

De Koster, M. B. M. and P. M. J. Warffemius (2005). American, Asian and third-party international warehouse operations in Europe: A comparison. International Journal of Operations and Production Management 25(8): 762-780.

Greene, W. (1980). Maximum likelihood estimation of econometric frontier functions. Journal of Econometrics 13: 26-57.

Gstach, D. (1998). Another approach to data envelopment analysis in noisy environments: DEA+. Journal of Productivity Analysis 9(2): 161-176.

Hackman, S. T., E. H. Frazelle, P. M. Griffin, S. O. Griffin and D. A. Vlasta (2001). Benchmarking warehousing and distribution operations: An input-output approach. Journal of Productivity Analysis 16: 79-100.

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

Hildreth, C. (1954). Pont estimates of ordinates of concave functions. Journal of the American Statistical Association 49(267): 598-619.

Johnson, A. L. and T. Kuosmanen (2010). How Operational Conditions and Practices Affect Productive Performance? Efficient Semi-parametric One-Stage Estimators. SSRN: http://ssrn.com/abstract=1485733.

Johnson, A.L. and L.F. McGinnis (2010). Productivity measurement in the warehousing industry. Accepted at IIE Transactions.

Johnson, A. L., W.-C. Chen and L. F. McGinnis (2010). Internet-Based Benchmarking for Warehouse Operations. Computers in Industry 61:280-286.

Jondrow, J., C. A. K. Lovell, I. S. Materov and P. Schmidt (1982). On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model. Journal of Econometrics 19(2-3): 233-238.

Kumbhakar, S. and C. A. K. Lovell (2000). Stochastic frontier analysis. Cambridge [England]; New York, Cambridge University Press.

Kuosmanen, T. (2006). Stochastic Nonparametric Envelopment of Data: Combining Virtues of SFA and DEA in a Unified Framework. MTT Disucssion Papers. Helsinki, Finland.

Kuosmanen, T. (2008). Representation theorem for convex nonparametric least squares. The Econometrics Journal 11: 308-325.

Kuosmanen, T. and A. L. Johnson (2010). Data envelopment analysis as nonparametric least-squares regression. Operations Research 58(1), pp. 149-160.

Ray, S. C. (1988). Data envelopment analysis, nondiscretionary inputs and efficiency: an alternative interpretation. Socio-Economic Planning Science 22(4): 167-176.

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

Ruggiero, J. (1996). On the measurement of technical efficiency in the public sector. European Journal of Operational Research 90(3): 553-565.

Shephard, R. W. (1953). Cost and Production Functions. New Jersey, Princeton University Press.

Simar, L., C. A. K. Lovell and P. Vanden Eeckaut (1994). Stochastic Frontiers Incorporating Exogenous Influences on Efficiency. Discussion Paper. Louvain-la-Neuve, Belgium, Institut de Statistique, Universite Catholique de Louvain.

Simar, L. and P. W. Wilson (2007). Estimation and inference in two-stage, semi-parametric models of production processes. Journal of Econometrics 136(1): 31-64.

Stevenson, R. E. (1980). Likelihood functions for generalized stochastic frontier estimation. Journal of Econometrics 13(1): 57-66.

Thanassoulis, E. (1993). A Comparison of Regression-Analysis and Data Envelopment Analysis as Alternative Methods for Performance Assessments. Journal of the Operational Research Society 44(11): 1129-1144.

Timmer, C. P. (1971). Using a probabilistic frontier production function to measure technical efficiency. Journal of Political Economy 79: 767-794.

Varian, H. R. (1992). Microeconomic Analysis. New York, Norton.

Wang, H. J. and P. Schmidt (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.

Winsten, C. B. (1957). Discussion on Mr. Farrell's Paper. Journal of the Royal Statistical Society Series a-Statistics in Society 120(3): 282-284.