Innovative Applications of O.R.

Measuring the efficiency of highway maintenance contracting strategies: A bootstrapped non-parametric meta-frontier approach

Saeideh Fallah-Fini\textsuperscript{a,}\textsuperscript{*}, Konstantinos Triantis\textsuperscript{b,}\textsuperscript{1,}\textsuperscript{4}, Jesus M. de la Garza\textsuperscript{c,}\textsuperscript{2}, William L. Seaver\textsuperscript{d,}\textsuperscript{3}

\textsuperscript{a}Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA 24060, USA
\textsuperscript{b}System Performance Laboratory, Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA, USA
\textsuperscript{c}Center for Highway Asset Management Programs (CHAMPS), Via Department of Civil and Environmental Engineering, Virginia Tech, Blacksburg, VA, USA
\textsuperscript{d}Department of Statistics, Operations, and Management Science, University of Tennessee at Knoxville, TN, USA

\textbf{A B S T R A C T}

Highly deteriorated US road infrastructure, major budgetary restrictions and the significant growth in traffic have led to an emerging need for improving performance of highway maintenance practices. Privatizing some portions of road maintenance operations by state Departments of Transportation (DOTs) under the auspices of performance-based contracts has been one of the innovative initiatives in response to such a need. This paper adapts the non-parametric meta frontier framework to the two-stage bootstrapping technique to develop an analytical approach for evaluating the relative efficiency of two highway maintenance contracting strategies. The first strategy pertains to the 180 miles of Virginia’s Interstate highways maintained by Virginia DOT using traditional maintenance practices. The second strategy pertains to the 250 miles of Virginia’s Interstate highways maintained via a Public Private Partnership using a performance-based maintenance approach. The meta-frontier approach accounts for the heterogeneity that exists among different types of highway maintenance contracts due to different limitations and regulations. The two-stage bootstrapping technique accounts for the large set of uncontrollable factors that affect the highway deterioration processes. The preliminary findings, based on the historical data for the state of Virginia, suggest that road authorities (counties) that have used traditional contracting for transforming the maintenance expenditures into the improvement of the road conditions seem to be more efficient than road authorities that have used the performance-based contracting. This paper recommends that road authorities use hybrid contracting approaches that include best practices of both traditional and performance-based highway maintenance contracting.

\textsuperscript{*}Corresponding author. Tel.: +1 617 894 0351; fax: +1 540 231 3322.
E-mail addresses: fallah@vt.edu (S. Fallah-Fini), triantis@vt.edu (K. Triantis), cherna@vt.edu (J.M. de la Garza), wseaver@utk.edu (W.L. Seaver).

1 Tel.: +1 703 538 8446.
2 Tel.: +1 540 231 5789.
3 Tel.: +1 865 974 6862.
4 This paper is based in part on work supported by the National Science Foundation.

\textbf{1. Introduction}

\textbf{1.1. Context and objectives}

In the past 20 years, the American Society of Civil Engineers (ASCE) has constantly rated the US road system as being in a poor condition (ASCE, 2009a). As pointed out by the ASCE, US road authorities have been facing a huge gap between the level of capital investment and the actual value that is needed to significantly improve the condition of the nation’s road system (ASCE, 2009b).

In view of this highly deteriorated road system, major budgetary restrictions, and the significant growth in traffic, there is a tremendous interest in the improvement of the efficiency and effectiveness of highway maintenance practices that preserve the road infrastructure so as to better support society’s needs.

One of the innovative initiatives undertaken in response to this need was the Virginia Public-Private Transportation Act of 1995 (PPTA). PPTA authorized the Virginia Department of Transportation (VDOT) to establish contracts with private entities for construction, maintenance and improvement of transportation facilities. In the first Public-Private Partnership undertaken under the auspices of the PPTA in 1996, a private contractor took the responsibility for the administration and maintenance of all assets within 250 miles (approximately 25\%) of Virginia’s Interstate highway system. The very important characteristic of the 10-year pilot contract was its performance-based nature. A performance-based contract (PBC) sets the minimum required conditions for roads, bridges and other assets without directing the contractor to specific maintenance methods that would help achieve the performance targets. The performance-based approach for contracting highway maintenance
projects was introduced as an alternative to the traditional approach. In the traditional approach, the tasks that need to be performed as well as the methods that should be used are specified in advance. The transition to performance-based contracting has been in response to a nationwide recommendation made by the Federal Highway Administration (FHWA, 2003).

Considering the increasing trend toward the application of PBCs for highway maintenance operations, there is an emerging need for frequent assessment of PBCs to make sure that the required level of service (LOS) has been met efficiently (Anastasopoulos et al., 2009; McCullough and Anastasopoulos, 2009). In a performance measurement system, the effectiveness dimension focuses on the achieved level of service and the efficiency dimension focuses on the amount of resources consumed to achieve a given level of service. Currently, most of the performance measurement systems developed and implemented by State Highway Agencies for road maintenance operations focus mainly on the overall improvement of the road condition (effectiveness) and do not adequately investigate the utilized resources and expenditures (efficiency) that have led to a given level of service (Ozbek et al., 2010a; TRB, 2006). Thus, the main objective of this paper is to utilize a combination of analytical non-parametric performance measurement techniques and develop an analytical approach to (i) evaluate and compare the efficiency of various types of highway maintenance contracts at the aggregate level (e.g., the efficiency of performance-based contracts (PBCs) versus the efficiency of traditional contracts) considering their achieved LOS; (ii) evaluate and compare the efficiency of different maintenance projects (contractors) that are under the same type of contract so as to recognize the contractors that are performing better in comparison with others in terms of both efficiency and effectiveness; (iii) identify the potential sources of inefficiency for the inefficient contractors, thus facilitating budget planning. This research contributes to the body of knowledge in the highway maintenance domain considering the limited number of studies (e.g., see Anastasopoulos et al., 2009; de la Garza et al., 2009) that assess and compare the efficiency of various types of highway maintenance contracts.

### 1.2. Points of departure

There are two broad types of methods in the efficiency literature for arriving at measures of relative efficiency, i.e., parametric and non-parametric methods (Thanassouliis, 1993). The parametric methods typically assume a functional form for the benchmark (frontier) and use data to estimate the parameters of that function. The estimated function is then used to arrive at estimates of the efficiencies of the units under analysis. The non-parametric methods use data and construct a piecewise-linear function that acts as a benchmark for measuring relative efficiency. Data Envelopment Analysis (DEA) (Charnes et al., 1978) is one of the popular non-parametric techniques that examines the relative efficiency of a set of similar decision-making-units (DMUs) (e.g., maintenance projects in a specific year) when a number of factors need to be considered. Several studies exist regarding applications of DEA models for measuring the performance of highway maintenance operations in Ontario, Canada (Cook et al., 1990, 1994; Kazakov et al., 1989), New Zealand (Rouse and Chiu, 2008; Rouse et al., 1997), and Virginia, USA (de la Garza et al., 2009; Fallah-Fini et al., 2009; Ozbek et al., 2010a; Ozbek, 2007). This paper focuses on a selection of recently developed non-parametric (DEA) performance measurement techniques and assesses their usefulness for measuring the efficiency of highway maintenance strategies. Thus, this research contributes to the literature of non-parametric efficiency measurement by introducing highway maintenance as a new engineering domain for the application of recent developments in the performance measurement field.

The first development relates to the difficulty that one encounters when comparing DMUs that may be heterogeneous. For example, highway maintenance projects under different types of contracts have different characteristics in terms of performance targets, methods, and resources. Thus, DEA models that estimate only one frontier for their evaluation may not be applicable. In order to fully capture the heterogeneity of highway maintenance contracts, this paper uses the non-parametric meta-frontier framework developed by a combination of papers by Battese, Prasada Rao, and O’Donnell (Battese and Prasada Rao, 2002; Battese et al., 2004; O’Donnell et al., 2008). The meta-frontier framework, first evaluates the efficiency of each DMU with respect to its group frontier, where DMUs in each group are assumed to have the same characteristics (e.g., use the same type of contract). In order to make an efficiency comparison across groups, a meta-frontier is then developed using best practices of all groups. Estimating the gap between each group frontier and the meta-frontier can help decision-makers by identifying performance improvement programs (O’Donnell et al., 2008). A literature review to date has not come across any efficiency measurement analysis that has used the meta-frontier concept for comparing the efficiency of different types of highway maintenance contracts.

The second development relates to the bias as well as the statistical properties of non-parametric efficiency scores. As has been stated by Simar and Wilson (2008), the non-parametric efficiency scores are biased by construction and the bias depends mainly on the sample size (the number of DMUs under analysis) and the dimension of the problem (the number of inputs and outputs). This is one of the potential problems in the meta-frontier literature, since classifying the DMUs into groups based on their characteristics may lead to a limited number of DMUs in each group. Starting with the work of Simar (1992), bootstrapping techniques were introduced into the efficiency measurement literature as an attractive paradigm for conducting various statistical inferences on the efficiency scores, including the correction for the bias. This paper adapts the meta-frontier framework to the recently developed bootstrapping techniques to correct for the shortcomings that may arise in the meta-frontier analysis. To date, there has been no study in the area of highway asset management that uses bootstrapping techniques for the statistical analysis of efficiency scores.

The third development relates to the integration of environmental and operational conditions into the efficiency analysis. Highway deterioration and maintenance is a process that is highly affected by uncontrollable environmental factors (e.g., climate condition) and operational conditions (e.g., traffic and load) (Ozbek et al., 2010a). The uncontrollable factors may account for the efficiency differences, thus special attention should be given to these factors. There are many methods for integrating uncontrollable factors into the efficiency analysis, each of which has its own advantages and drawbacks as will be discussed in Section 2.4. The two-stage semi-parametric bootstrapping technique by Simar and Wilson (2007) addresses many difficulties that exist in previously developed methodologies. In the first stage, the Simar and Wilson (2007) two-stage bootstrapping technique obtains the non-parametric efficiency scores using only controllable input and output variables, and in the second stage, the observed efficiency patterns are econometrically explained using the set of uncontrollable factors. Adopting this approach constitutes an additional contribution to the literature of highway maintenance performance measurement, since previous research in this area has not used this approach to integrate uncontrollable factors into the efficiency analysis. This paper uses a combination of these techniques to develop an analytical approach and applies it to an empirical dataset of pavement condition, traffic, climate condition, and maintenance expenditures provided by VDOT.
The remainder of this paper is organized as follows. Section 2 describes the proposed analytical approach as well as a description of the techniques that are used. The results and insights obtained from implementing the approach on the empirical data are discussed in Section 3. Conclusion and future directions are provided in Section 4.

2. Foundations (methodology)

2.1. Non-parametric estimation of the production frontier

A Decision Making Unit (DMU) whose performance is measured is generally regarded as the entity that uses a production process that converts multiple inputs into multiple outputs. The underlying production process is constrained by the “production possibility set” or “technology set” \( \mathcal{V} \), which is the set of all physically attainable points \((x,y)\), where \( x \in \mathbb{R}^m \) is the input vector and \( y \in \mathbb{R}^n \) is the output vector. For the purpose of efficiency analysis, the upper boundary of \( \mathcal{V} \) that is called the efficient boundary or “technology frontier” is of importance. The frontier is defined as the set of best performing DMUs that use the minimum input level to produce a given output level or generate the maximum output given a specific input level. For a given DMU with input and output variables \((x,y) \in \mathbb{R}^m \times \mathbb{R}^n\), the measures of technical efficiency can be defined respectively as:

\[
\mu(x,y) = \inf \{ \mu | (x,y) \in \mathcal{V} \},
\]

\[
\lambda(x,y) = \inf \{ \lambda | (x,y) \in \mathcal{V} \}.
\]

\( \mu(x,y) \) and \( \lambda(x,y) \) lie between zero and one. \( \mu(x,y) \) is the input-oriented measure of efficiency and represents the input reduction required by the DMU to become efficient holding the outputs constant. \( \lambda(x,y) \) is the output-oriented measure of efficiency. \( 1/\lambda(x,y) \) represents the output expansion required by the DMU to become efficient holding the inputs constant.

Data Envelopment Analysis (DEA) uses mathematical programming (i.e., is considered a non-parametric approach) to compute the frontier and the efficiency scores corresponding to all DMUs under analysis. DEA is considered as an appropriate approach for measuring performance of highway maintenance operations since establishing “production standards” and measuring absolute efficiency in this setting is hard, if not impossible (Cook et al., 1994; Kazakov et al., 1989). In addition, DEA allows for the consideration of different non-economic factors, such as traffic, load, climate conditions, etc. where each of the factors plays an important role in the efficiency analysis. Readers are referred to Ozbek et al. (2009) for a summary on the use of DEA models in the transportation field.

Assuming the information on input and output variables of \( n \) DMUs is available, the non-parametric estimation of the variable returns to scale technology set \( \mathcal{V} \) can be written as follows:

\[
\hat{\mathcal{V}}_{VRS} = \left\{ (x,y) \in \mathbb{R}^{m+n} | \sum_{i=1}^{n} \gamma_i y_i \geq \sum_{i=1}^{n} \gamma_i x_i \right\}
\]

for \( \gamma_1, \ldots, \gamma_n \) such that \( \sum_{i=1}^{n} \gamma_i = 1; \gamma_i \geq 0 \). \( i = 1, \ldots, n \).

The estimated technology set \( \hat{\mathcal{V}}_{VRS} \) is a convex set enveloping the data points and includes a bound inner approximation of the true (unobserved) and larger technology set \( \mathcal{V} \). Thus, the DEA estimators overestimate the true efficiency scores. That is why DEA estimators of efficiency, although consistent, are biased by construction (Simar and Wilson, 2008). The bias as well as rate of convergence of the DEA estimators to their true (unobserved) values mainly depends on the properties of the dataset under analysis, namely: (1) the number of observations (DMUs) in the dataset; (2) the number of input and output variables; and (3) the distribution of observations around the frontier (Borger et al., 2008). As the number of input/output variables (dimension of the space) increases, the Euclidean distance between the observations increases. Thus, there will be fewer nearby observations that can convey information about the portions of the efficient frontier which is of interest (Simar and Wilson, 2008). In addition, an increase in the number of input and output variables requires more observations (DMUs) for constructing the efficient frontier, leading to an increase in the bias of the estimated efficiency scores (Simar and Wilson, 2008). In the literature, this situation is referred to as “curse of dimensionality”.

Another shortcoming of non-parametric estimators is related to the assumption that DMUs under analysis have the same characteristics and are similar. Thus, a common frontier or benchmark is estimated for evaluating efficiency scores of all DMUs. The reality is that this is not the case in most of the interesting and practical problems and DMUs usually experience some heterogeneity. For example, imagine one is comparing the efficiency of production units that are operating under different regulations or are located in different countries, or have different ownerships (public versus private). Under these circumstances, separate efficiency frontiers need to be estimated for different groups of DMUs (O’Donnell et al., 2008). Sections 2.2 and 2.3 describe some of the recent remedies that have been suggested in the literature to address the heterogeneity and small sample bias issues, respectively.

2.2. Heterogeneity among DMUs

When the DMUs under analysis face different production opportunities, they need to make choices from different sets of feasible input-output combinations. The differences in production opportunities can be attributed to the physical, social, and economic environments in which the production process takes place (e.g., type of equipment) (O’Donnell et al., 2008). Under the stated conditions, the DMUs belong to different technology sets (groups), thus using traditional DEA models that estimate one common technology frontier for comparing all DMUs will increase the risk of having unreasonable estimates of the efficiency scores.

This situation is relevant to highway maintenance in that there is a difference among limitations, regulations, and maintenance operations that correspond to the various types of maintenance contracts. For example, traditional contracts are short-term and conservative in nature. Their main focus is on first cost (lowest bidder), thus the corresponding set of techniques, tools and materials to perform the required maintenance operations are chosen accordingly. Performance-based contracts (PBCs) are long-term and tend to optimize the cost over a project’s lifecycle. Long-term responsibility of the PBCs motivates the contractors to be innovative in their design (the tasks that should be performed and their specifications), as well as in their selection of tools and materials. Thus, contractors or road authorities under various types of contracts are working in different production environments with different regulations. As a result they may only have access to a restricted part of the production possibility set. The boundaries of these restricted production possibility sets form the group frontiers (O’Donnell et al., 2008). Thus, when comparing the efficiency of the PBCs versus the traditional contracts, the DMUs under each type of contract form separate groups, each of which has its specific operational environment.

The recently developed analytical meta-frontier approach (Battese and Prasada Rao, 2002; Battese et al., 2004; O’Donnell et al., 2008) provides an appropriate methodology to evaluate and compare the efficiency of DMUs that belong to different groups.
Assuming that there are L groups of DMUs, the meta-frontier framework first pools the observations of all groups and estimates a meta-technology set that contains all input-output combinations that are technologically possible. The boundary of this unrestricted technology set is called the meta-frontier and is used to measure the efficiency of each DMU assuming that technology is freely interchangeable and that the DMUs in all groups have potential access to the same technology. In the second step, the observations of each group are used separately to define group-specific technology sets. The boundaries of the group technology sets are called group frontiers. The distance from an input-output point to its group frontier is a measure of technical efficiency of that DMU, meaning how well a DMU is performing in comparison with the rest of the DMUs in its own group. Each group frontier represents the boundary of a restricted technology set, where the restrictions are defined from resource, regulatory, or other constraints (O’Donnell et al., 2008). Analyzing the gap between a group frontier and the meta-frontier indicates the potential improvements that can be made in the efficiencies of the DMUs of that group when one removes the restrictions/regulations and uses the best practices that are provided by the meta-technology, which is defined as the technology of all input-output combinations associated with all DMUs in the sample (O’Donnell et al., 2008).

To provide a better understanding, Fig. 1 provides a graphical illustration of the meta-frontier for a simple example with one input and one output variable. DMUs under analysis belong to two heterogeneous groups, thus two group frontiers represented by XX′ and YY′ are computed. Consider a specific DMU operating at the input-output combination labeled by A. The output-oriented technical efficiency of DMU A with respect to its group frontier XX′ and meta-frontier MM′ are calculated respectively as:

$$TE_{XX′}(A) = \frac{OB}{OC}, \quad TE_{MM′}(A) = \frac{OB}{OD}.$$  \hspace{1cm} (4)

To analyze the gap between the group frontier XX′ and the meta-frontier MM′, the meta-technology ratio (MTR) of DMU A is defined as:

$$MTR_{XX′}(A) = \frac{TE_{MM′}(A)}{TE_{XX′}(A)} = \frac{OB/OD}{OC} = \frac{OC}{OD}. \quad \hspace{1cm} (5)$$

The meta-technology ratio basically measures how close a group frontier is to the meta-frontier. Moreover, Eq. (6) as a reconstruction of Eq. (5) implies that the technical efficiency of DMU A measured with respect to the meta-frontier can be decomposed into the product of technical efficiency with respect to the group frontier (representing the characteristics of the group and its state of knowledge), and the meta-technology ratio for group XX′ (representing how close the group frontier is to the meta-frontier) (O’Donnell et al., 2008).

$$TE_{MM′}(A) = TE_{XX′}(A) \cdot MTR_{XX′}(A). \quad \hspace{1cm} (6)$$

Once data on input and output variables of DMUs from all groups are available, the group frontiers and the meta-frontier can be estimated using either parametric or non-parametric methods. As it was mentioned in Section 2.1, the non-parametric DEA method is used in this paper for estimating the group frontiers and meta-frontier.

Note that the meta-frontier approach can be seen as similar to the conditional frontier approach (Daraio and Simar, 2005, 2007) in which efficiency score of each production units is evaluated with respect to a frontier that has been constructed conditioned on uncontrollable factors that production unit under analysis is facing and to the fuzzy clustering approach (Seaver and Triantis, 1992; Triantis et al., 2010) where clusters are created conditioned by the representation of the environment. Conditional frontier approach can be applied to our problem by defining an uncontrollable dummy variable characterizing group belonging. Conditioning on this dummy variable, separate frontiers can be estimated for various groups. Alternatively, one can use multivariate methods (Triantis et al., 2010) as an alternative modeling approach to identify relatively homogeneous populations for which the performance evaluation can be performed in the subpopulations as well as in the whole population.

To analyze the effects of uncontrollable factors on the production process, Daraio and Simar (2005, 2007) develop a non-parametric regression line of the ratios between the conditional and unconditional efficiency measures on uncontrollable factors. An increasing smoothing nonparametric regression line implies a negative effect of the uncontrollable factors on the production process; while a decreasing nonparametric regression line shows a positive effect of uncontrollable factors on the production process (Daraio and Simar, 2007). In the case of the multivariate framework (Triantis et al., 2010), an outcome of the DEA analysis is the construction of an environmental dependency index (EDI) defined as the ratio of the local efficiency (with respect to the DMUs in a cluster) to the global efficiency measure (with respect to all DMUs in the sample). The ratio between the conditional and unconditional efficiency measures developed in the conditional frontier approach and the EDI index in the multivariate framework is conceptually similar to the meta-technology ratio developed in the meta-frontier framework.

2.3. The bias associated with the estimated efficiency scores

The lack of elaboration on statistical properties of the non-parametric point estimates of efficiency scores early in the literature did not provide the opportunity for performing statistical analysis of the essential properties of the efficiency results (Borger et al., 2008). If probability distributions of the non-parametric estimators are known, then the construction of confidence intervals, the correction for the bias, or conducting statistical inferences on efficiency scores are possible. Considering the very limited number of general analytic derivations for asymptotic sampling distributions of non-parametric frontier estimators, the introduction of the bootstrapping technique in the efficiency measurement literature in the past decade has led to important breakthroughs in approximating the sampling distributions of efficiency estimators. Simar and Wilson (2007) describe in detail the bootstrapping algorithms for estimating the empirical distribution of efficiency scores, correction for their bias, and construction of confidence intervals. The fundamentals of the bootstrapping methodology are as below.

Bootstrapping starts by obtaining a consistent estimation of the data generating process (DGP), the process that has led to the input-output data corresponding to all DMUs in the sample. Re-
The three-stage approach is another technique developed by Ruggiero (Ruggiero, 1998) where a single environmental harshness factor that can best represent the effect of uncontrollable variables is developed. The environmental harshness factor is then used for grouping the DMUs such that each DMU is compared with those that are in the similar or worse environmental conditions. When more than one uncontrollable variable impacts output and when differences among the levels of those variables are subtle, defining a single environmental harshness factor to capture the effect of all uncontrollable variables needs further consideration (Triantis et al., 2010). As it was discussed in Section 2.2, the conditional efficiency measures (Daraio and Simar, 2005, 2007) or the multivariate framework (Triantis et al., 2010) are other approaches that can be potentially useful when incorporating environmental variables in efficiency analysis.

However, the focus of this research is on the innovative application of the bootstrapping techniques to address most of the difficulties associated with the standard two-stage approach explained before. Simar and Wilson (2007) suggests a bootstrapping algorithm where the bias-corrected efficiency scores are first constructed. The
bias-corrected estimates of efficiency scores are then used in the second stage-regression as the dependent variable. Bootstrapping on the second stage regression leads to consistent estimates of the parameters of the regression model.

The algorithm assumes that sample observations \((x_i, y_i, z_i)\) are i.i.d. observations from the random variable \((X, Y, Z)\) with probability density function \(f(x, y, z)\) where \(x\) and \(z\) are the input, output, and uncontrollable variables for DMU \(j\), respectively. In addition, based on the separability assumption between the space of input-output variables and the space of environmental variables, the support of the density function \(f(x, y, z)\) is assumed to be \(\Psi \times R^2\), where \(\Psi\) is the production possibility set constructed by only input-output variables \((x, y)\) and \(R\) represents the dimension of the uncontrollable variables. Then, the relation between efficiency scores \(\delta_j\) and the environmental variables \(z_j\) is assumed to be linear, i.e., \(\delta_j = z_j \beta + \epsilon_j\) where \(\beta\) represents the parameter vector and \(\epsilon_j\) distributed as \(N(0, \sigma)\), represents i.i.d. variables independent of \(z_j\) with left truncation at \(1 - z_j\), since the dependent variable \(\delta_j = \frac{1}{\delta_j} \geq 1\) (based on the Eqs. (1) and (2), inverse of measures of efficiency are greater than or equal to one).

2.5. The proposed analytical approach

After orderly arrangement of the methods and techniques described above, the methodological contribution of this paper can be described as follows:

1. A comprehensive set of key controllable input and output variables pertinent to the transformation process as well as a comprehensive set of external and uncontrollable factors that represent the environmental and operational conditions for the context under analysis (e.g., road maintenance operations) are developed.
2. DMUs under analysis are classified into different groups based on their characteristics as well as the objectives of the efficiency analysis.
3. Group frontiers are developed for each group of DMUs using the Simar and Wilson (2007) two-stage bootstrapping technique. Thus the DEA efficiency score of each DMU with respect to its own group frontier is obtained and corrected for the bias.
4. DMUs from all groups are pooled together to develop a meta-frontier using the Simar and Wilson (2007) two-stage bootstrapping technique.
5. Efficiency scores of DMUs with respect to their own group frontiers as well as the meta-frontier are compared via the meta-technology ratio (O’Donnell et al., 2008) to arrive at the measures of performance of different groups of DMUs and potential improvements each group can have to shift to the meta-frontier.

By performing steps 3 and 4 of this methodology, the bias-corrected efficiency scores as well as their corresponding confidence intervals are calculated. Moreover, using the second stage regression, one can also examine if the observed efficiency patterns can be explained based on the environmental and operational conditions. To illustrate the implementation of this methodology as well as the intuition behind each step, an empirical study using real data is presented in Section 3.

3. Empirical application: evaluating the performance of road maintenance operations

The proposed methodology is applied to an empirical dataset of pavement condition, traffic, climate condition, and maintenance expenditures for approximately 180 miles (lie within seven counties) of Virginia’s Interstate highways maintained using a “performance-based” maintenance strategy over the fiscal years 2002 to 2005. The PBC data belongs to the first three years of a 5-year contract. Due to the unavailability of cost data corresponding to the last two years of the performance-based contract, we couldn’t consider the whole planning horizon in the analysis. Given the mid-term nature of the performance-based contract used in this analysis, we proceed with the assumption that data of the first three years can represent the perspective of contractors over the whole planning horizon (i.e., five years). However, this could have not been a fair assumption if performance-based contracts had been set for longer periods such as 10 or 15 years. Applying the methodology described in Section 2.5 provides the possibility for measuring and comparing performance of the two contracting strategies (traditional versus performance-based). Detailed discussion on the implementation of each step of the methodology is provided next.

3.1. Factors considered in the analysis

In order to select the variables, emphasis was placed on the effects of maintenance activities as well as on a set of explanatory variables that have caused these effects. So, a set of input, output, and uncontrollable variables was obtained (de la Garza et al., 2009; Fallah-Fini et al., 2009). Following is a short description of these variables. For further discussion as to why they have been chosen and how they have been calculated can be found in Fallah-Fini (2010).

Lane-miles Served:
This variable represents the total lane-miles of the road sections that are maintained within each county and captures the extent of the workload each county has performed. Thus, it is considered as one of the outputs of the maintenance operations.

Change in Pavement Condition:
VDOT uses the Critical Condition Index (CCI) to represent the condition of a road section with respect to the load-related and non load-related distresses. The CCI varies between 0 and 100. In this paper, the change in the CCI of the maintained road sections is used to capture the improvement in the road condition due to maintenance operations. For example, change in CCI of a specific road section corresponding to year 2002 is obtained by deducting the CCI of the road section at year 2002 from the CCI of the same road section at year 2003. This variable is considered as one of the outputs.

Another data collected by VDOT is International Roughness Index (IRI), which is an indicator of overall pavement smoothness or ride quality. Change in IRI can potentially be considered as another output variable. However, this study only takes into account “change in CCI” as the output variable, since CCI is the main factor to identify deteriorated road sections and the required maintenance operations (JLARC, 2002).

Traffic:
In order to capture the effect of traffic on pavement deterioration, the Annual Average Daily Traffic (AADT) data is used. Large values for AADT can potentially increase the extent of pavement deterioration and require greater maintenance effort. Traffic is treated as an uncontrollable factor in this study.

Load:
In order to quantify the traffic load a pavement encounters, the concept of Equivalent Single Axle Load (ESAL) has been used. Load
is an uncontrollable factor and captures the extent of deterioration due to vehicle forces.

Climate Factors:
Climate condition factors are uncontrollable and mainly affect the deterioration of the paved lanes. Climate data, extracted from the National Climate Data Center (NCDC, 2010), contains four variables (i.e., minimum and maximum temperature, total rainfall, and total snowfall) for all the counties under analysis over the corresponding years. Based on the physical and elevation maps for the state of Virginia, the “Mountainous” dummy variable was also created to specify the counties that lie in a mountainous area and thus are exposed to more severe environmental and operational conditions. This variable takes the value of one if a county lies in a mountainous area and zero, otherwise.

Maintenance Expenditure:
This variable represents the cost data corresponding to the maintenance of the road using traditional or performance-based contracting. The “BHWA-Highway and Street Construction Cost Index” developed by the Bureau of Labor Statistics was used as an inflation/deflation rate to adjust the cost data of different years. Maintenance expenditure is the only controllable input variable used in this study.

Given the availability of the “maintenance expenditure” data at the county level, the definition of DMUs is limited to the counties of Virginia that encompass the sections of the Interstate system that are maintained using traditional or performance-based maintenance practices. Thus, each county for each fiscal year is considered as a DMU. This definition for DMUs enables us to analyze trends in the county efficiency scores over time.

Finally, not every county has performed maintenance operations at each year. Thus, after some data cleaning, 25 DMUs for the traditional approach and 26 DMUs for performance-based approach were used to form the two groups of DMUs (traditional strategy versus performance-based strategy) whose efficiency performance needed to be compared. The results of this comparison provide insights on the type of contracting strategy that has been more efficient when maintaining Virginia’s Interstate. Table 1 provides the descriptive statistics for the final dataset corresponding to the two groups of DMUs. Given the range of the values for the minimum temperature or snowfall in each group, it is obvious that DMUs (counties) are geographically spread across the state of Virginia and are experiencing different environmental conditions. Thus, it is hypothesized that climate conditions can potentially justify the differences among efficiency scores of DMUs within each group.

Before we start estimating the group- and meta-frontiers, we need to justify as why the separability assumption holds in the context of highway maintenance, and consequently, the two-stage approach is a valid technique to use. This requires justifying that uncontrollable factors such as climate condition or traffic load do not affect the shape of the production possibility set constructed by maintenance expenditure as input and lane-miles served as well as change in CCI as outputs. Note that condition of road infrastructure evolves due to interaction among deterioration and renewal processes. Uncontrollable factors such as precipitation mainly affect the deterioration process. To confirm this, we chose several road sections that had not been maintained over several years from both counties that are located in mountainous areas (with harsher environmental conditions) and the counties that are located in milder environmental conditions. Comparing the critical condition index (CCI) of these road sections showed that the ones that are located in harsher environmental conditions have faster deterioration over years as opposed to those road sections that are located in better environmental conditions (the counties toward east of Virginia).

However, the climate condition does not have a significant effect on the performance of maintenance operations. Based on the discussion we had with maintenance managers, the yearly maintenance operations are usually performed around late August till late October. The climate condition in the state of Virginia during these months is such that it should not affect the performance of workers who are in charge of performing maintenance operations. Most importantly, exploring our dataset showed that CCI of most of the road sections that have been maintained in each year has reached to 100, irrespective of the environmental conditions they belong to.

In sum, environmental condition in the context of highway maintenance affect the deterioration process, but it does not have a significant effect on the outputs of maintenance and renewal processes. Thus, counties that lie in harsher environmental condition can also achieve the best possible condition after performing required maintenance operations, and consequently, have access to the same production possibility set as those counties that are located in better environmental condition. It is just that road authorities in harsher environmental condition most probably need to spend more money to bring the road sections from a severe condition to an acceptable or perfect condition. More formally, one can perform the non-parametric test proposed by Daralo et al. (2010) for testing the separability assumption in any problem context.

3.2. Estimation of the group frontiers

In this section, group frontiers are constructed so that the performance of each road authority (i.e., county) can be compared with the performance of the rest of road authorities that are using the same type of contract. To estimate an appropriate non-parametric technology frontier for each group, the orientation (input or output oriented) and returns to scale (variable (VRS) or constant (CRS) return to scale) of the DEA model should be defined. Ozbek (2007) as well as Rouse et al. (1997) discusses a significant presence of scale effects in highway maintenance operations. Moreover, the lower and upper bound of CCI imposes a lower bound and upper bound for the output variable “Change in CCI”. There is also a limit for the maximum value that the variable “Lane-miles Served” can take and that is equal to the total area of the Interstate that lie in each county. Thus, a VRS frontier is needed to adjust for the upper and lower bounds of the output variables, since a CRS frontier continues extending linearly without taking any boundary constraint into account (Rouse and Chiu, 2008).

Choosing the orientation of the DEA model mainly depends on the objectives of the analysis. Note that each road authority is responsible for maintaining the highway network in its administrative area. In the real world, the whole highway network is divided into several road sections and maintenance treatments are defined for each one of these sections. Given the limited available budget, road authorities cannot maintain all road sections. Instead, they need to decide on road sections that have higher priority for performing maintenance operations. The result of this decision is captured by the output variable “Lane-miles Served”. Road authorities also need to decide on the type of maintenance operations that should be performed on those road sections. This decision affects the level of improvement in quality of the road sections and is captured by the output variable “change in pavement condition”. Given the limited available budget, higher spending for improving the condition of road sections (i.e., the quality of operations) may decrease the number of road sections that can be maintained (captured by lane-miles maintained). Thus, road authorities are the ones who decide as how the maintenance budget should be spent. This means that road authorities have control over outputs, both in...
terms of number of lane-miles that can be maintained and level of improvement that can be made. Based on the discussion that we had with several maintenance managers, they are mainly interested to know if they could use their limited budget more efficiently and maintain higher number of lane-miles (road sections) in better condition. Thus, we use the output-oriented BCC model in this analysis.

Table 2 shows the efficiency scores for some of the counties with respect to their group frontier. The shortfalls in the efficiency scores represent the magnitude of improvement that can be achieved in outputs without any additional investment in maintenance expenditure. Exploring the patterns of efficiency scores in different counties can have some important policy implications. For example, the Spotsylvania County has been 100% efficient over the years 2002 and 2003, but its efficiency score has decreased in the next two years. In contrast, Henrico County shows an improvement in their efficiency scores over the years 2002 to 2004. This may require maintenance managers of these counties to explore the changes in their policies and practices over the years under analysis as a potential source for the change in their performance. Some of the counties such as Augusta and Carroll Counties are of concern since their efficiency scores have been relatively low over all years.

The efficiency results presented in Table 2 do not show any correction for the bias that inherently exists in the non-parametric frontier estimations. To correct the bias, the Simar and Wilson (2007) bootstrapping algorithm was used for each group of DMUs. Table 3 shows a summary of the original (uncorrected) as well as the bias-corrected efficiency scores for a small selection of DMUs from the most efficient to the inefficient observations in both groups.

Analyzing the bias corrected results showed that the fully efficient DMUs that define the frontier have the largest value for bias, such as Alleghany (traditional approach) and Dinwiddie (performance-based approach) counties in year 2004. The large bias of the efficient DMUs can potentially be attributed to the fact that not enough observations in any of the groups are used for constructing the frontiers. Heterogeneity of observations within each group, due to the fact that DMUs (counties) belong to different environmental conditions, can be another important reason for the large bias that has been introduced in the computed efficiency scores (Borger et al., 2008). Moreover, for many observations, the original (uncorrected) efficiency scores are placed outside the constructed confidence intervals. This underlines the risk associated with using the uncorrected efficiency scores (Borger et al., 2008).

Finally, the bootstrapping results suggest that the DMUs whose uncorrected efficiency scores are 1.00 are not potentially 100% efficient. They seem to be efficient only due to the small sample sizes in both groups.

The next step is to test the hypothesis if any of the uncontrollable factors presented in Table 1 can significantly justify the difference in efficiency scores among the DMUs within each group. Based on the bootstrapping algorithm described in Section 2.4 and after exploring different functional forms on different sets of uncontrollable variables in both datasets, the inverse of the bias-corrected efficiency scores (i.e., the dependant variable) showed a significant relation with the maximum temperature and log of the snowfall in the traditional contracting dataset and with the maximum and minimum temperatures in the performance-based contracting dataset. The regression results and the confidence intervals for the estimated model parameters are depicted in Table 4. The bootstrapped confidence intervals for the estimated parameters do not contain zero, thus that the parameter estimates are significantly different from zero.

Note that in these models the inverse of the bias-corrected efficiency scores is used as the dependant variable, hence smaller values for the dependant variable mean better efficiency scores. The parameter values for performance-based contracting presented in Table 4 indicate that those road authorities whose minimum temperature is lower (representing a worse environmental condition) have worse efficiency scores (i.e., higher values for the inverse of efficiency scores). In terms of the traditional contracting approach, the higher values for the snowfall indicate a worse environmental condition (i.e., higher values for the inverse of efficiency scores). Thus, the observed positive sign for the coefficient of log(snowfall) is correctly expected. Moreover, in both models the variable maximum temperature has shown a significant relation and its coefficient has a negative sign. This means that the counties with higher maximum temperature have shown better efficiency scores. This may seem counter-intuitive initially, since based on
the physics of road deterioration, higher frequency and severity of extreme hot days leads to problems related to pavement softening as well as load-related rutting (NRCAN, 2010). The fact is that the stated scenario can potentially happen in states such as Texas or Arizona which experience severe hot days during the Summer with extreme temperatures much higher than 100 °F. As it was shown in Table 1, the averages of maximum temperature in both groups are around 85 °F. Thus, as our datasets show (and also with respect to the geographical location of the state of Virginia), the variable maximum temperature is not contributing as a deterioration factor. Instead, higher values for the maximum temperature mean that those counties are experiencing a better environmental condition overall and have been able to come up with better results after performing maintenance operations. Another possible explanation for this observation is that maintenance operations can be carried out with less maintenance expenditure, meaning more efficiently, in better environmental conditions.

Note that load did not show up as a significant factor in any of the regression models. Further analysis in this regard showed that counties (interstates) that are maintained using performance-based contracting group (i.e., I-81, I-95, and I-77) face similar traffic associated with heavy vehicles, and consequently, incur similar traffic load. Thus, the traffic load did not show up as a significant variable in the regression model associated with the performance-based group. On the other hand, exploring traffic data of the counties (interstates) in the traditional contracting group (i.e., I-81, I-64, and I-66) showed that these interstates experience a wider range of values for traffic load associated with heavy vehicles. However, traffic load still did not show up as a significant variable in the regression model. Our conjecture in this regard is that traffic load can be a significant factor particularly when road section experience extreme cold or hot weather. Extended heat and traffic load increases wear and tear on road section just like extended cold and traffic load. However, the state of Virginia does not regularly get the very high or very low temperatures for the extended time periods.

### Table 3

<table>
<thead>
<tr>
<th>County</th>
<th>Original eff. score</th>
<th>Bias-corrected eff. score</th>
<th>Lower bound (5%)</th>
<th>Upper bound (95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional contracting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alleghany (2004)</td>
<td>1.00</td>
<td>0.78</td>
<td>0.72</td>
<td>0.88</td>
</tr>
<tr>
<td>Spotsylvania (2004)</td>
<td>0.77</td>
<td>0.65</td>
<td>0.60</td>
<td>0.71</td>
</tr>
<tr>
<td>Roanoke (2004)</td>
<td>0.56</td>
<td>0.49</td>
<td>0.46</td>
<td>0.52</td>
</tr>
<tr>
<td>Performance-based contracting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dinwiddie (2004)</td>
<td>1.00</td>
<td>0.85</td>
<td>0.77</td>
<td>0.94</td>
</tr>
<tr>
<td>Chesterfield (2002)</td>
<td>0.77</td>
<td>0.71</td>
<td>0.67</td>
<td>0.75</td>
</tr>
<tr>
<td>Bland (2004)</td>
<td>0.64</td>
<td>0.56</td>
<td>0.53</td>
<td>0.60</td>
</tr>
</tbody>
</table>

### Table 4

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter value</th>
<th>Lower bound (2.5%)</th>
<th>Upper bound (97.5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>4.724</td>
<td>3.569</td>
<td>6.135</td>
</tr>
<tr>
<td>Max temp.</td>
<td>–0.030</td>
<td>–0.050</td>
<td>–0.014</td>
</tr>
<tr>
<td>Min temp.</td>
<td>–0.034</td>
<td>–0.058</td>
<td>–0.012</td>
</tr>
</tbody>
</table>

### Table 5

<table>
<thead>
<tr>
<th>Group</th>
<th>Technical efficiency wrt group frontier</th>
<th>Technical efficiency wrt meta-frontier</th>
<th>Meta-technology ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
<td>Mean</td>
</tr>
<tr>
<td>Traditional contracting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perf-based contracting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alleghany</td>
<td>0.71</td>
<td>0.18</td>
<td>0.70</td>
</tr>
<tr>
<td>Spotsylvania</td>
<td>0.82</td>
<td>0.13</td>
<td>0.61</td>
</tr>
</tbody>
</table>

*Table 5 reports the original efficiency scores. Using the bias-corrected efficiency scores leads to the same conclusions. However, differences in magnitude of the bias for each DMU with respect to the group and meta-frontiers occasionally results in meta-technology ratios greater than one.*

### 3.3. Estimation of the meta-frontier

In Section 3.2, the relative performance of each road authority within its group was evaluated. As a general rule, road authorities’ efficiency levels measured with respect to one frontier (e.g., traditional contracting frontier) cannot be directly compared with other road authorities’ efficiency levels measured with respect to another frontier (e.g., performance-based contracting frontier). In this section, based on the meta-frontier approach, a common meta-frontier for measuring the relative efficiency of road authorities across groups is computed. As it was described in Section 2.2, the meta-frontier envelopes the two group frontiers that were computed in Section 3.2. To construct the meta-frontier, the datasets corresponding to both traditional and performance-based contracting strategies were pooled. Next, the output-oriented BCC model was applied to the pooled dataset to measure the efficiency of the road authorities relative to the estimated meta-frontier, irrespective of their contract type. Table 5 reports the average of the efficiency scores of each group with respect to the group frontiers and meta-frontier, as well as the average of meta-technology ratios for each group.

Looking at the group level efficiency scores in Table 5 shows that the technical efficiency of the performance-based contracting group is 0.82 with respect to its group frontier. This means that, on average, road authorities that are using performance-based contracting can improve their outputs by 18% using the same amount of maintenance expenditure they have already expended. In addition, the technical efficiency of the performance-based contracting group with respect to the meta-frontier is 0.61. This means that based on the unrestricted meta-technology (with no limitations and regulations), the road authorities that are using performance-based contracting could improve their output by 39% (more than twice as the potential improvement they could achieve based on the performance-based contracting). The average meta-technology ratio of 0.73 for performance-based contracting group shows that on average, the maximum output that can be achieved by this group is 73% of the maximum output that can potentially be...
achieved using the unrestricted meta-technology (by removing limitations and regulations).

The large value of meta-technology ratio for the traditional contracting group shows that the counties that are using the traditional contracting approach are playing a more important role in constructing the meta-frontier than the counties that are using the performance-based contracting approach. In other words, the tangency of the meta-frontier and the traditional contracting group frontier is much more than the tangency between the meta-frontier and the performance-based contracting group frontier. This preliminary finding suggests that, based on our historical data for the state of Virginia, road authorities (counties) that have used traditional contracting for transforming the maintenance expenditures into the improvement of the road conditions seem to be more efficient than road authorities that have used the performance-based contracting approach. Note that traditional contracting focuses on the lowest-bid combined with method-based specifications. In contrast, performance-based contracting, in its purest form, focuses on maintaining (achieving) performance targets without detailing on how, when, and where the work should be performed (NCHRP, 2009). Our finding suggests that VDOT should not rely solely on LOS specifications in performance-based contracting, instead VDOT may want to use some hybrid approaches by bringing some of the features of traditional highway maintenance contracting into performance-based maintenance. This is in fact the expectation created by the meta-frontier in that all DMUs should have access to the best practices associated with each group. One plausible explanation for our finding is that traditional contracting has been used and tested many times with different road authorities, but this is VDOT’s first experience with performance-based contracting. As a result, there are many reasons that can affect the successful implementation of PBCs, such as contractor’s quality/capacity, the acquisition/award process, managing the cultural change inside the organization, the methods used for monitoring and evaluating the contractors, risk management process, etc. (NCHRP, 2009). It is very possible that VDOT’s shift to the needed PBC culture had not been fully developed during the execution of this first performance-based maintenance pilot project; in fact the culture needed to support PBC is perhaps 180 degrees apart from the traditional one. Furthermore, the road-builder’s contracting industry is not used to think in terms of performance-based models which are radically different from means-and-methods-based specifications. These and other dimensions should be considered while analyzing the numerical results from Table 5.

Just like in Section 3.2, the bootstrapping algorithm is used to see which of the uncontrollable (environmental and operational) factors can significantly justify the difference in DMUs’ efficiency scores evaluated with respect to the meta-frontier. Note that in the pooled datasets the observations (DMUs) belong to two different types of contracts. Thus, a dummy variable representing the contract type should be added to the set of independent variables. After exploring various functional forms, the inverse of the bias-corrected efficiency scores showed a significant relation with the dummy variables “Mountainous” and “Contract type”. The test results rejected the null hypothesis that the efficiency scores of road authorities under different types of contracting strategies are not significantly different at 5% confidence interval. This further justified the use of meta-frontier framework for evaluating and comparing efficiency of traditional versus performance-based contracting strategies.

Table 6

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter value</th>
<th>Lower bound (5%)</th>
<th>Upper bound (95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.82</td>
<td>1.60</td>
<td>2.02</td>
</tr>
<tr>
<td>Mountainous dummy variable</td>
<td>0.437</td>
<td>0.208</td>
<td>0.666</td>
</tr>
<tr>
<td>Contract Type dummy variable</td>
<td>-0.473</td>
<td>-0.699</td>
<td>-0.244</td>
</tr>
</tbody>
</table>

The type of maintenance contracts has a significant role in the difference among efficiency scores of different road authorities and they do not belong to the same technology (frontier). Note that this dummy variable takes the value of one if a DMU has used the traditional contracting approach and zero otherwise. Thus, the negative sign of the dummy variable implies that those road authorities that have used the traditional contracting approach have lower dependent variable (i.e., higher efficiency scores).

The Kruskal–Wallis test was also run to test if there is any significant statistical difference between the efficiency scores of road authorities that have used traditional contracting and those that have used performance-based contracting (where efficiency scores were obtained with respect to the meta-frontier). The test results rejected the null hypothesis that the efficiency scores of road authorities under different types of contracting strategies are not significantly different at 5% confidence interval. This further justified the use of meta-frontier framework for evaluating and comparing efficiency of traditional versus performance-based contracting strategies.

4. Conclusions

This paper utilizes recent developments on non-parametric (DEA) frontier estimations to develop an analytical approach for evaluating and comparing the performance of highway maintenance projects that are using different types of contracts. The developed approach is applied to an empirical dataset of pavement condition, traffic, climate condition, and maintenance expenditures for approximately 180 miles of Virginia’s Interstate highways maintained by VDOT using traditional maintenance practices and 250 miles of Virginia’s Interstate highways maintained using a performance-based maintenance strategy.

The non-parametric meta-frontier framework is exploited in this paper to account for the heterogeneity of DMUs (i.e., road authorities that are using traditional maintenance contracting and road authorities that are using performance-based maintenance contracting). The meta-technology ratio that is computed for each group depicts how close a group frontier is to the meta-frontier. Evaluation of this gap helps road authorities and decision makers by analyzing the effects of the economic and physical characteristics of contract types (e.g., regulations, limitations, size and quality of labor force, etc.) on the performance of road authorities. Thus, potential improvements in performance of road authorities resulting from any change in regulations and limitations correspondent to contract types can be assessed.

Comparing the meta-technology ratio of the traditional maintenance contracting group (98%) with that of the performance-based maintenance contracting group (73%) in the dataset under analysis showed that, based on our historical data for the state of Virginia, road authorities that have used traditional contracting for transforming the societal resources for the improvement of the road conditions seem more efficient than road authorities that have used the performance-based contracting approach. However, the
validity of this preliminary finding can be further evaluated by using a more comprehensive dataset that includes the whole planning horizon for the performance-based contracting. As it was discussed, since this is VDOT’s first experience in using performance-based contracting, many reasons can potentially justify the under-performance of PBCs. Moreover, using some hybrid contracting approaches by bringing some of the features/best practices of traditional highway maintenance contracting into performance-based maintenance can potentially be very helpful with improving the efficiency of highway maintenance contracting.

As it has been discussed in the literature, the non-parametric measures of efficiency are biased by construction. Thus, the Simar and Wilson (2007) bootstrapping technique was applied on the non-parametric (DEA) group frontiers as well as on the meta-frontier to correct for the bias in the estimated efficiency scores, and also to construct confidence intervals for efficiency scores. For many road authorities, the original (uncorrected) efficiency scores are located outside the constructed confidence intervals. This finding underlines the risk involved with using uncorrected efficiency scores.

The Simar and Wilson (2007) bootstrapping technique also revealed the patterns in the estimated efficiency scores with respect to group frontiers and the meta-frontier can significantly be explained by environmental factors, in particular minimum/maximum temperatures, snowfall, and characteristics such as being located in a mountainous area. Considering the important role played by the uncontrollable factors in the deterioration process of the Interstate, more extensive research needs to be done on finding alternative methods for incorporating these uncontrollable factors in performance measurement.

Moreover, there are many other explanatory variables (e.g., the geological characteristics of soil and rock types, vegetation, pavement age and thickness, construction type (Portland cement or bitumen), etc.) that affect the deterioration as well as outputs of maintenance operations. These factors have not been taken into account in this analysis due to lack of availability of data associated with those variables. However, a more comprehensive analysis of performance of contractors requires accounting for these factors in the analysis as well.

This study can also be further improved by taking into account the fact that the required maintenance expenditure not only depends on the amount of improvement in pavement condition, but also the initial condition of pavement (initial CCI). For example, an improvement of 20 units in CCI potentially needs different levels of effort if we start from CCI of 50 and go to CCI of 70 or start from CCI of 70 and go to CCI of 90. In this study, both these cases are treated equivalently. However, incorporating the initial condition into the DEA model would increase validity of the results.

Despite the data limitations this paper has, this line of research provides the decision-makers with proper knowledge of the efficiency level of different highway maintenance contracts or projects. This is crucial for guiding future decisions regarding the renewal of contracts, the pricing of these contracts, and the potential efficiency improvement opportunities, since past performance is one of the very important criteria that road authorities can consider when awarding new contracts.

Acknowledgment

We would like to acknowledge the assistance of the Virginia Department of Transportation for providing the data in this research. This research is funded by the National Science Foundation, Award # CMMI-0726789. Any opinions and/or findings are those of the authors and do not necessarily represent the views of the sponsors.

References


