

1-1-2003

Use of DEA to Evaluate Non-linear and Imprecise Information in Construction Contractor Performance

Gerald H. Williams

Portland State University

Timothy R. Anderson

Portland State University, tim.anderson@pdx.edu

Let us know how access to this document benefits you.

Follow this and additional works at: https://pdxscholar.library.pdx.edu/etm_fac



Part of the [Engineering Commons](#)

Citation Details

Williams, G. H. and Anderson, T. Use of DEA to Evaluate Non-linear and Imprecise Information in Construction Contractor Performance, Portland International Conference on Management of Engineering and Technology (PICMET), Portland, OR, 2003.

This Article is brought to you for free and open access. It has been accepted for inclusion in Engineering and Technology Management Faculty Publications and Presentations by an authorized administrator of PDXScholar. For more information, please contact pdxscholar@pdx.edu.

Use of DEA to Evaluate Non-linear and Imprecise Information in Construction Contractor Performance

Gerald H. Williams, Jr., Timothy R. Anderson

Department of Engineering Technology Management, Portland State University, USA

Abstract- This paper explores the use of Data Envelopment Analysis as a tool to evaluate the efficiency of building construction projects in Oregon where the inputs, including, “reported information” available prior to pricing and contractor selection is imprecise or at least, cannot be entirely or accurately captured by exogenous measures. This work builds on prior work in the field that evaluated missing, imprecise or non-existent information in DEA models and has broad application to the service sector of the economy where information is the prime input in the system.

I. INTRODUCTION

The generic economic production model used in the Data Envelopment Analysis (DEA) methodology assumes a linear relationship between inputs and outputs even though we are unaware of precisely how those inputs are translated to outputs. For example: if one were to use DEA to evaluate the efficiency of automobiles, the input might be gallons of gasoline and the output miles driven. The most output for a fixed input or the least input for a fixed output would be a measure of automobile efficiency. However, DEA researchers recognized a some time ago that not all automobiles are created equal and it is nonsensical to compare fuel efficiency of say an Austin Mini to Chevrolet Suburban; the auto’s have significantly different missions and what we may refer to as “economies of scale” or as Zhu [1] does, “Convexity.” Banker formulated a variable returns to scale (VRS) method to accommodate the economies of scale problem [2]. The VRS model allows for economic comparisons among a relatively narrow range, however the basic assumption that underlies Banker’s formulation remains the linear relationship between inputs and outputs.

While the VRS method would, most likely, exclude the comparison of the Mini to the Suburban, it almost certainly would not exclude the comparison of two Suburban’s with different amounts of inputs and outputs. And if there was a distinctly non-linear relationship between the amount of gas input and say the number of passenger-miles output, this would merely show up as inefficiency in one of the two vehicles (or perhaps even the same vehicle on two separate trials.) However, there may be a very good reason for the differences that are not based on “efficiency” but rather a non-linear relationship between inputs and outputs, in the example due to the different configurations of the automobile.

There are a number of commercial institutions where the economic transactions can be characterized by non-linear relationships between inputs and outputs – in particular, those where there are limiting boundaries for either measure or endogenous effects that cannot be fully captured. This is

particularly true with professional services where input resources like “information,” and output measures “time” and “cost” have definite fixed boundaries. An example that comes to mind almost immediately is weather prediction. Without virtually any information, it is easy to predict that the weather at any location on the planet by simply knowing the location and having some idea of the historical climate. For example, we can say that the temperature in Portland, Oregon tomorrow will range between 5° Fahrenheit and 105° F. We can say this because no matter what the time of the year, because we know historical climate data, that the extreme temperatures ever recorded in Portland fall within these boundaries. With little more information, like the day of the year requested, we can narrow that range significantly. But, no matter how much money we invest, we know that we cannot exactly predict the weather six months from now. Therefore there is a distinctly non-linear relationship between the amount of information paid for as an input, and the accuracy of the output prediction of temperature.

The same relationship can be said to be true in the construction sector of the economy. Without almost no information, we know that the cost of construction is finite, and exists within some approximate range, say between \$1 and \$5000 per square foot. The amount of information collected by the owner and transmitted to the contractor, narrows the range of costs considerably, but simply cannot eliminate the variability or range entirely, in part because the building environment exists in “weather” that cannot be reliably predicted. The question many owners want answered is, “how much information should I pay for in order to reduce the uncertainty in the pricing and how much information is simply a waste of resources, given the cost boundaries?”

In the public sector of the construction industry in Oregon, owners have the opportunity to select and contract with contractors relying on less than complete construction documents, which is nearly unique among states in the US [3] [4]. Beginning in the 1980’s, public agencies in Oregon started to use alternative construction contractor selection methods to hire builders for public projects. This process was authorized by law¹ as early as 1977, but little used prior to the mid to late 1980’s, particularly in the building sector of the construction industry. The common contractor selection method used since the public bidding laws were enacted in the 1930’s was by sealed competitive bid (commonly known as Design-Bid-Build, DBB or Lump Sum Bid, LSB method.) However, by the early 1980’s several public owners and construction contractors felt that the DBB method was a

¹ See Oregon Revised Statutes: ORS 279.015 (as amended)

principal factor in an ever increasing litigious market sector, delays in completion, increasing insurance rates and bond costs and overall lowering of the quality of work. As a remedy, first the Port of Portland² and next the Portland Development Commission³ authorized negotiated procurements on their respective construction projects (called CM/GC jobs for Construction Manager / General Contractor method.) From the 1980's to the year 2000, public agencies in Oregon have constructed more than five hundred public building projects that cost more than a million dollars (\$1,000,000) each⁴ and representing in excess of five billion dollars (\$5,000,000,000) in tax expenditures. At least one hundred thirty⁵ of those projects were negotiated procurements where contractor selection was based at least in part on prior performance of the firm (including overall project quality of performance metrics.) While there have been some attempts to analyze or audit specific CM/GC projects⁶, there have been no prior attempts to evaluate a large set of these projects and compare their collective performance against the more widely used competitive bid project delivery system (PDS.)

II. LITERATURE REVIEW

A. Information and Contractor Selection

“Winner’s Curse” or adverse selection is widely studied in field of competitive economics and it is of particular concern in the public sector of the construction industry where most owners continue to use DBB as their primary PDS. Kagel [5] and others have studied common value auctions and the winner’s curse in the construction industry and concluded that, while experienced construction bidders are subject to winners curse affects, the construction market attempts to correct for the curse by employing three strategies⁷: withdrawal of a low bid due to error, subcontractor buyout, and by overpricing change orders. Kagel also points out that there is a significant amount of “private information” in the bidding environment that is not accounted for by the plans and specifications, some of which

can be characterized as “experience” of the bidders and reliance on “rules-of-thumb” in bidding. One strategy that Kagel omits however, which might just be the most important, is the bonding of subcontractor’s, which serves to almost entirely eliminate the general contractor’s greatest risk: the risk of a major subcontractor failing to perform on their contract.

In Kagel’s brief review of the construction industry for his research, he concludes that contractor’s rely primarily on the plans and specifications as the primary information as a basis for their bid. In fact, that proposition is solidly embedded in both Federal and Oregon Contract Law.⁸

Information Model

The amount of information available to a construction contractor at the time pricing or selection is required is never complete. Figure 1 shows four assumed states of information: the lowest figure (arrow) represents all or “complete” information. Clearly, this ideal is never achieved, as indicated in the second from bottom figure in the set. This is the theoretical DBB competitive bid model of information upon which virtually all contracts and construction law are based. It provides that the basis for all competitive contracts or scope of work for all competitive contracts is based solely on information provided by the Owner (typically by and through their consultant architects and engineers) and the State of Nature. The State of Nature concept is based on an acceptable or typical performance of the States of Nature, usually meaning the “weather” but also includes such things as normally expected underground and hidden conditions. The amount of information regarding the State of Nature can be increased by expending time, effort and funds to make better predictions about the State of Nature, such as paying for the consultants to perform underground soils investigations. But some States of Nature, in particular the weather, are outside the bounds of investigation, particularly on long term projects which are the subject of this study. The standard in the construction industry is to use the “average” State of Nature as a benchmark for what the bidders should expect to encounter. Only if the actual State of Nature of the weather is abnormally adverse, will there be an amendment to the Contract.

The “Competitive Bid Contract Model” more accurately reflects the actual state of information in the competitive contract method, where the actual information provided at the time of pricing is less than Complete Information, by both the State of Nature and Errors and Omissions in the information provided by the Owner. This leads to the contractual provisions in most construction contracts that allow for amendments to the Contract based on errors or omissions.

² The first known competitive negotiation selection for a construction contractor was the then K-Wing of terminal south at the Portland International Airport. It is not known why this project received an exemption from public bidding, no records of this project currently exist.

³ The second known competitive negotiation selection for a construction contractor was for the Yamhill parking structure serving the Pioneer Place downtown development. The process was justified in that project because the PDC had bad experiences in the past with projects that did not finish on time and the parking structure had to finish prior to the opening of the mixed use development for holiday season. Increasing liquidated damages to cover the risk of developer lost profits was thought to be an uneconomical approach to ensuring the project opened on time.

⁴ Our study focused on projects costing more than \$5 million, but we did drop down to as low as \$1 million in order to obtain as many negotiated projects as we could.

⁵ There are 127 known negotiated procurements in the database, and another four dozen or more project that the selection method is not known (whether lump sum bid or negotiated procurement.)

⁶ Oregon’s Secretary of State’s Audit Division, audited three major prison projects.

⁷ Although not necessarily by these names.

⁸ See *In United States v. Spearin*, 248 U.S. 132, 39 S.Ct. 59, 63 L. Ed. 166 (1918), and *Oregon, A.H. Barbour & Sons, Inc. v. State Highway Commission*, 248 Or. 247, 433 P.2d 847 (1967); *General Construction Company v. Oregon State Fish Commission*, 26 Or. App. 577, 554 P2d 185 (1976).

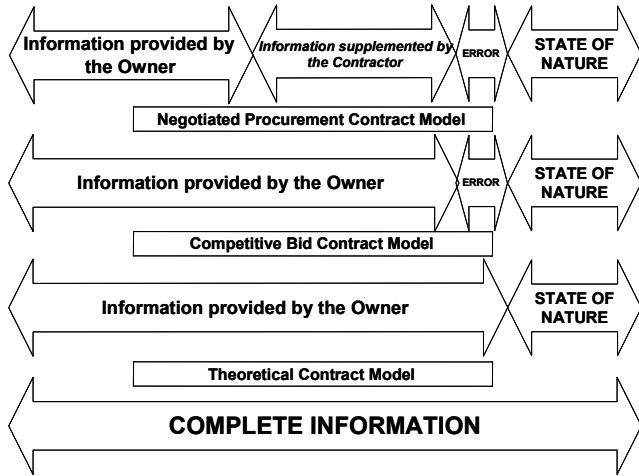


Figure 1: Various Construction Bidding Models

The top figure in the Information model is the CM/GC model where the amount of information presented at the time of selection or pricing is admittedly less than complete than provided in the DBB model. This model recognizes the construction contractor’s ability to exercise judgment and requires him to fill in the gaps between the amount of information given and what is required to derive a Contract price.

B. DEA Methodology

DEA is a widely studied methodology as demonstrated by Tavares [6] and Seiford [7]. For a complete discussion of the development of DEA and the different DEA Models the reader is referred to: [8], [9] and [1]. For the purposes of this paper, we present only a brief discussion to support our use of DEA in this research.

DEA is a non-parametric, linear programming based method for evaluating relative efficiencies of Decision Making Units or DMUs. DEA was originally devised by [10] as a method to derive relative efficiencies of different organizations or activities using multiple inputs and multiple outputs. The focus on comparative efficiency using multiple inputs and outputs was, in part generated by the desire to measure factors other than financial performance such as sales and profitability.

The linear programming method used in DEA evaluates a weighted sum of the various outputs divided by a weighted sum of the various inputs, in the model, while leaving the weighting scheme left up to the particular DMU, in order to maximize its DEA score. This “Ratio Model” is expressed mathematically, in the form as:

$$\text{Max: } h_0 = \frac{\sum_{r=1}^s u_r y_{r,0}}{\sum_{i=1}^m v_i x_{i,0}} ;$$

$$\text{Subject to: } \frac{\sum_{r=1}^s u_r y_{r,j}}{\sum_{i=1}^m v_i x_{i,j}} \leq 1 \quad j = 1, \dots, n;$$

$$u_r, v_i \geq 0; \quad r = 1, \dots, s; \quad i = 1, \dots, m.$$

In the model, the x_{ij}, y_{rj} are the known inputs and outputs of the j th DMU and the u_r 's and v_i 's are the variable (or criteria) weights to be determined by the solution of the linear program⁹. That is, DEA allows each DMU₀ to pick the weighting scheme that maximizes its efficiency score relative to all other DMUs, subject to the constraint that any other DMU, with an identical weighting scheme, cannot achieve an efficiency score greater than 1.0. The weighting scheme adopted by each individual DMU can be considered to reflect that specific DMU’s strategy for converting inputs into outputs. For example a given DMU strategy may be to put all of its weight on one particular input and one particular output, and if that DMU has the highest ratio of these two metrics it will be deemed efficient.

1. Current DEA Application Areas

Seiford [7] and Tavares [6] have traced the evolution of DEA and provided a comprehensive bibliography of DEA papers that includes more than 1,500 and 3,200 references respectively. DEA’s principal application has been to determine relative efficiencies using both financial and non-financial performance measures. The ability to include non-financial performance measures has lead to initial applications in the Education, Government and Healthcare fields. More recently, banking, warehousing, and the airline industries have also been major application areas where DEA has been used to combine both financial and non-financial performance measures.

2. DEA in the Construction Industry

The construction and building industry is often times called the single largest industry in the United States, and the largest single sector of the US economy representing between 10 and 12% of GDP. It would seem, given the industry size and impact on the economy, there would be plenty of Construction Industry applications of DEA; that is not however the case. DEA applications in the construction industry have been limited to a few papers including: building sector research [11], road construction vehicle management [12, 13], and nuclear power plant construction times [14]. These applications have not directly involved a broad analysis of construction project management of major capital construction.

Perhaps one of the reasons DEA has not been used to evaluate project management on major capital construction

⁹ Note that, as shown, the mathematical model is a non-linear formulation. Linearizing this formulation is discussed in the referenced DEA texts.

projects, or project management in general, is the problem of determining what inputs and outputs to use and how to measure them with any precision. In order to deal with imprecise measures of inputs and outputs, Cooper [15] and Zhu [16] have developed methods that use Ordinal and Categorical variables instead of direct measures. In this research, the direct measures are known and quantifiable, but fail to capture all of the influences on inputs in the system.

III. DATA COLLECTION AND ANALYSIS

A. Type of Data Required

The basic Construction Project model assumed for this research is shown in Figure 2:



Figure 2: Basic Construction Project Model

The model dictates the type of data that needed to be obtained: inputs, consisting of the kind of information the contractor required for production and; outputs: measures of performance. Measures of performance in “projects” generally, have been limited to measures of: Cost, Quality and Schedule performance. On the input side, we needed to obtain data that characterized the type of information the contractor would need as a basis for the contract price (either a competitive lump sum or a negotiated guaranteed maximum price.) Both the input and output metrics were determined by an Expert Panel of industry executives, representing Architects and Engineers, Contractors, and Owners. The Input and Output Metrics originally obtained included the following: Percentage Complete of the plans and specifications, Quality of the Plans, Owner’s prior experience in similar projects, Owner budget and schedule expectations, Contractor team experience, and Access to the site issues.

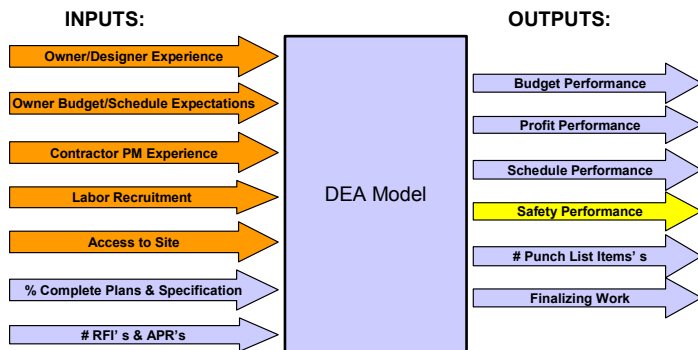


Figure 3: Original DEA Model from Expert Panel Work

Figure 3, illustrates the initial DEA model that resulted from two rounds of surveys and discussions with the industry experts. The Owner/Designer experience, Labor recruitment, budget expectations and Access to Site metrics were derived from subjective rankings based on a 1-5 scale. As a proxy for Construction Team Experience, the construction team project manager’s experience in number of years of construction experience (at any professional level) was used.

The percent complete of the “plans and specifications” was also something of a subjective measure because there is no exact definition or consistent measure of completeness used in the industry. Architects, according to the Experts, typically consider their work as a process moving from: Conceptual Design to Schematic Design, to Design Development or Preliminary Design, then proceeding to Final Design and Working Drawings. But, there is no consistent rule like 25% complete means “X” and 50% complete means “Y”. However, the data collected did tend to fall into certain ranges of percent complete and were not categorically distributed¹⁰ as is evident from the histogram presented in Figure 4.

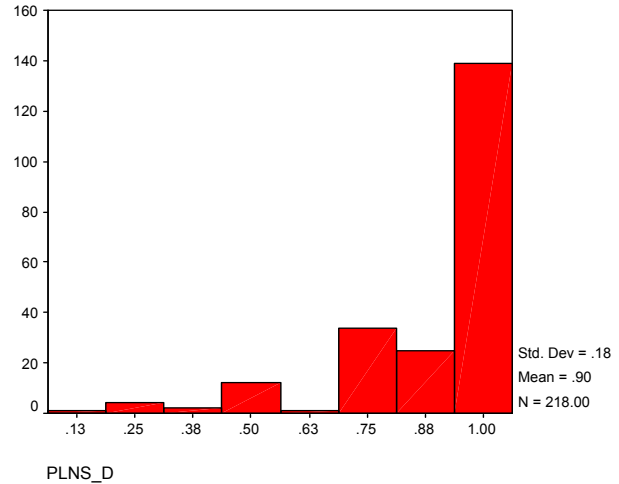


Figure 4 Histogram of %Complete Plans and Specifications

This data and further discussions with Architects on the Expert Panel, would guide our parsing of the data into different categories.

1. Data Collection

The collection of data for this research was a multi-phase process involving construction contractors, owners, architects, internet websites and the Portland and Seattle Daily Journal of Commerce (DJC) newspapers. Ultimately through all of the difference sources we were able to identify more than 500 public building projects, 407 of which generally fit into our research criteria. Of the 407 projects we obtained some cost information on 367 projects, ranging in

¹⁰ By this we mean that all of the data does not fall into specific groupings like: 25%, 50%, 75% and 100% complete plans and specifications. However, in spite of the fact that the data does not fall into discrete categories, they do generally fall into consistent ranges, which we understand from our Expert Panel represents the different levels of design completeness.

cost¹¹ from a \$175,591 CM/GC fire station project to a \$189,859,282, also CM/GC airport expansion project, in sum totaling \$5,050,962,407 of public construction.

A summary characterization of the projects by Project Delivery System is contained in Table 1.

TABLE 1 DESCRIPTION OF THE DATA

Project Delivery System Type	Number	Smallest	Largest	Number with Unknown Financial Information
CM/GC	136	\$175,591 1,400 SF	\$189,859,282 1,000,000 SF	8 with no \$ data 17 missing some \$ data
Design Bid Build	191	\$663,325 4,600 SF	\$112,497,017 569,000 SF	2 with no \$ data 89 missing some \$ data
Unknown	80	\$223,166 11,300 SF	\$35,211,963 270,000 SF	2 with no \$ data 78 missing some \$ data
Total	407			

Of these 407 projects we were only able to obtain enough data on 218 jobs that allowed us to use them in our DEA model¹². A total of 71 Construction Contractors¹³ and 70 Architecture firms¹⁴ are represented in the project database, ranking from the largest firms in their respective sectors in Oregon, to small, even single practitioner offices.

2. Analysis of the Data

The first analysis performed on the data obtained for this research was to evaluate the population using standard statistical methods. Figure 5 is a single frequency histogram of the output metric Budget Performance, which we defined above as the total final cost divided by the initial bid or contract guaranteed maximum price. Note the average Budget Performance is 0.947, with a standard deviation of 0.07. The figure indicates that the data is generally well distributed with one sole outlier at the extreme left hand side of the figure.

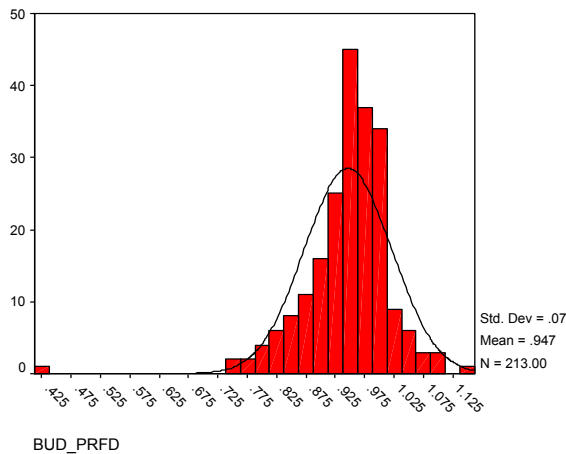


Figure 5: Budget Performance Metric Histogram

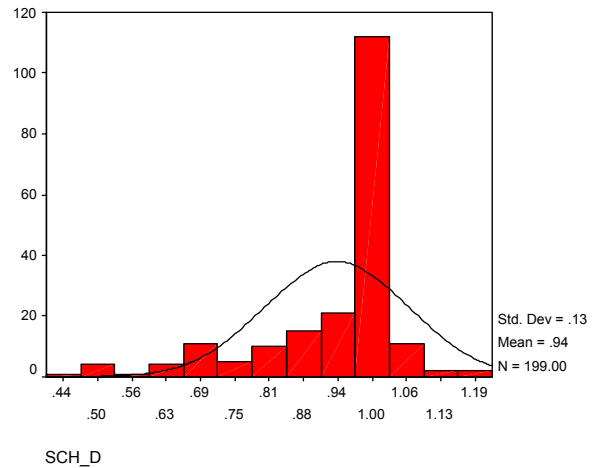


Figure 6: Schedule Performance Metric Histogram

Figure 6 is a single frequency histogram of the Schedule Performance metric, which indicates a much broader distribution than the Budget Performance metric. This stands to reason because time is a more available resource than money in most situations and exceeding planned performance time is more likely and less costly than exceeding the project financial resources. It's far more likely to allow a project to run longer by 100% of the original schedule, resulting in a Schedule Performance Metric of 0.50, than increase by double the cost. An increase in total cost of double would in fact only occur in projects that undergo a "Cardinal Change" and therefore, by definition, the original contract value is irrelevant because it does not reflect the intent of the contracting parties.

3. Population Independence

The principle purpose of this research was to determine if the different project delivery systems resulted in better projects based on certain output metrics as determined by the Expert Panel.

¹¹ These costs are adjusted to 2001 using Engineering News Record's Construction Cost Index obtained from their website at: <http://enr.construction.com/features/conEco/costIndexes/constIndexHist.asp>.

¹² Some of these would be eliminated during the model validation phase

¹³ This number includes two Joint Ventures as separate companies from their parents.

¹⁴ Note that some firms have changed names over the years; we have attempted to count a firm only once if it added or lost partners along the way. However, Joint Venture firms were counted separately from their parent organizations consistent with our treatment with contractors.

TABLE 2 GROUP STATISTICS¹⁵

Group Statistics					
	PDS	N	Mean	Std. Deviation	Std. Error Mean
BUD_PRFD	1	101	.93817	.076605	.007622
	2	112	.95411	.071808	.006785
SCH_D	1	97	.94000	.129317	.013130
	2	102	.93098	.132108	.013081

Table 2 presents the Group Statistics for the two principal populations: DBB (#1) and CM/GC (#2.) What this analysis tells us is that, while there are differences in the population, they are slight and not statistically significant. The important thing to note here is how little variance there is in the principal output metrics. Another way to look at this data is given in the two-dimensional plot given in Figure 7 .

Output Metric Plot

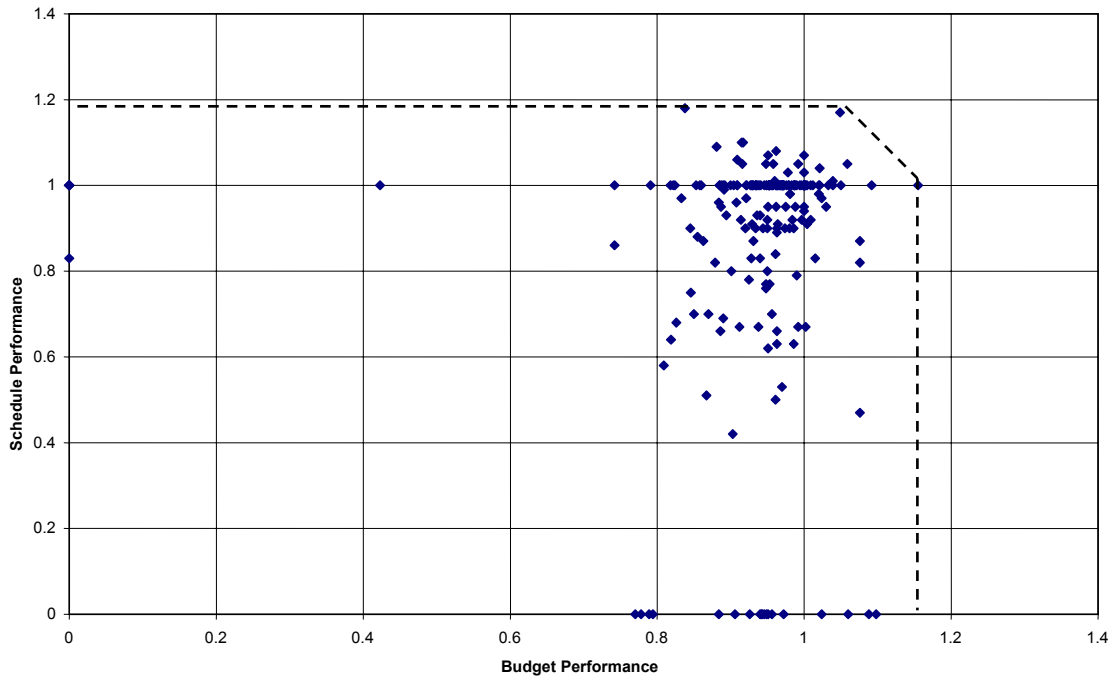


Figure 7 - Two Dimensional Plot of the Output Metrics

If there were no inputs, and these were the only outputs, then the envelopment of this data, shown as a dashed line above, would give us the technical efficiencies of the projects. From the plot, it appears that the minimum efficiency would be in the range of 0.70 or 70%. Again however, the important thing to note is that the data is not broadly distributed but instead, it is rather tightly grouped around the 1.0, 1.0 intersection. In fact, only just slightly more than 13% (29/218) of the projects with non-zero data points, fall outside of plus or minus 20% of the 1.0 measure in either direction. If this were the envelopment data, it appears that just three projects would form the efficiency frontier, but we do have inputs and other outputs for this data set.

4. DEA Model Analysis

Figure 7 presents a good representation of project outcomes if inputs are not considered. Our research was intended to consider both inputs and outputs. However, the important thing to note from Figure 7, and from our statistical analysis presented above, is that the project outputs are not dramatically dissimilar. In fact, while there is quite a range of outputs, particularly in the schedule performance metric, the vast majority of the projects had performance outputs that were substantially similar, with few truly outstanding and few truly horrible projects. Even those projects that did poorly on one of the two metrics, appear to have made up for it in the other. And, the reason the project performance results presented in Figure 7 are important, is because it directly impacts how the DEA analysis would be performed.

B. Initial DEA Model Analysis (CRS and VRS Models)

The normal method for evaluating DMUs in DEA is by application of either a Constant Returns to Scale (CRS) model or a Variable Returns to Scale (VRS) model. These models have been used throughout the literature and applied

¹⁵ The reader will note that the number of cases are not the same for each Metric, this is due to the fact that some information is missing in some of the cases and zero's have been omitted from this statistical analysis. In the case of the DEA analysis that follows, zero's are included in the output measure.

in a number of different industries and economic sectors. However, neither the traditional CRS or VRS models is well suited to the evaluation of the data set in this research. This is because of the peculiar fact that the inputs vary so substantially but the outputs do not. Traditionally, when a

researcher wanted to evaluate a data set of DMUs with extreme differences in input and output metric values; such as comparing grocery stores and including mini-marts, traditional mainstream stores, and warehouse stores in the same data set, the researcher would apply a VRS model.

VRS DEA Scores

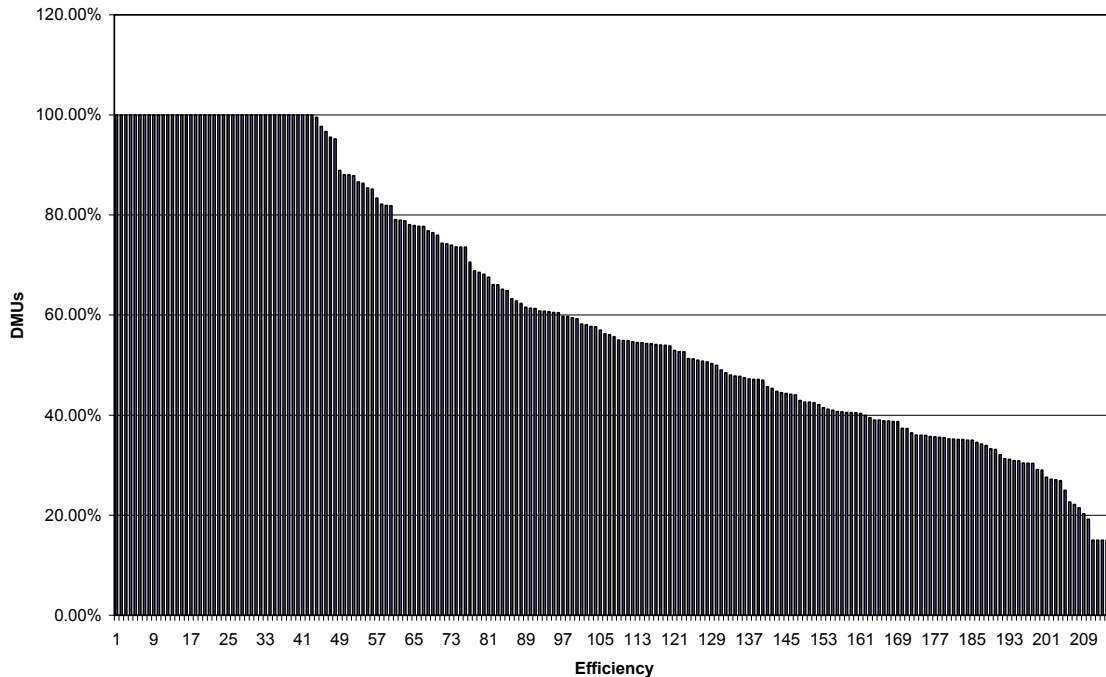


Figure 8 VRS DEA Scores - Base Model Ranked by Efficiency Score High to Low

Running the VRS Model results in only 130 of 215 total projects with greater than 50% efficiency scores and 48 of 215 with better than 90% efficiency score. These results do not correspond to either the results of the statistical analysis (Table 2 above) or the output plot (Figure 7) nor do they match the perception of the Expert Panel Members who uniformly maintained that, except in rare cases, projects should fall within a narrow range of results.

It is probably intuitive, but the results of a CRS model are even worse than the VRS model. The CRS model resulted in only 16 of 215 projects with efficiency scores above 90% and only 87 of 215 with scores better than 50%. It should be obvious from these results that the application of either a CRS or VRS model without significant modification does not shed any light on the analysis of this construction project data.

C. Modifying the DEA Model for Construction Project Data

It's really fairly obvious why the VRS and CRS models result in the distribution of scores as they do, since the outputs lay within a rather narrow range and the inputs vary from 0.25 to 1.0 (or 25% to 100% complete plans and specifications.) Nearly all of the projects with plans and

specifications of 100% complete will have efficiency scores below 50% - which is exactly what happens. The problem comes in the evaluation of those input metrics. Recall that above we stated that there is no fixed standard for evaluating percentage complete, and the method we used was simply to ask the various parties to assign a number based on their past experience,¹⁶ this results in extreme non-linear relationships between the principal inputs and outputs. While there is basic agreement that "conceptual design" is less complete in terms of "% complete plans and specifications" metric than "schematic design" or "preliminary design" the precise estimates of the values varies. In other words, one project managers 25% complete may be another's 35 or 40% complete.

In situations where the inputs and outputs are not directly linearly related and the exact relationship is unknown or not captured by known inputs, a possible approach would be to apply a "categorical" variable model. In that case the "conceptual design" complete projects do not compete directly with the "final design" complete jobs, but rather they compete among each other within categories. However, that

¹⁶ Note also that 100% complete was the default if no other information was given.

would render meaningless the point of this research, which is in part, to compare the PDS by outcome, against one another and determine if one type is significantly superior to the other.

It is important to note also, that the analysis set forth in Table 2 and described in Figure 7 support Kagel's point [5], and our conclusion that the amount of actual "Information" provided is not captured by the metric "% complete plans and

specifications" in spite of the fact that the Expert Panel recommended. This is because the metric fails to take into account the economic reality of the industry and the amount of training and experience of the estimators and managers of the construction companies involved, which is substantial¹⁷. This is perhaps better visualized in the Process Model suggested by Figure 9.

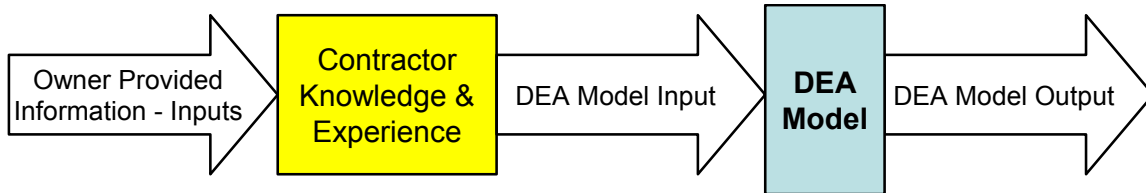


Figure 9 Process Model

We know, both from our own study and from industry standards that the cost of public buildings of virtually all types resides within a relatively small range. For example, RS Means *Building Construction Cost Data*, estimating guide provides a section that provides per square foot cost estimates for approximately sixty different type buildings from Apartments to Warehouses. This data includes estimates at the 25 percentile, Median, and 75 percentile ranges. Furthermore, while these data are determined by national averages, the guide also provides regional indexes to convert the average costs to a cost for a specific area. Portland, Oregon for example has a weighted average of about 1.06 times the national cost average for buildings, according to Means.

Armed with a commonly available estimating guide and experience in the local construction market, it is possible, easy in fact, for a construction estimator to narrow the range of possible costs far tighter than an input of 25% to 100% would imply. So, for the purposes of a DEA model, is it possible to account for this base of knowledge that is an additional "resource" (or enhances the Information provided resource) that results in "production" from the model? And the answer to that question is, probably, but probably not to a level of certainty that makes the evaluation meaningful. Also, while we did collect data on the different construction companies, none of the data we collected as a proxy for experience (including number of years in business, bonding capacity, project team experience and so on) could be reliably tied to a single input metric for "knowledge" that would differ significantly from company to company. This again, would make the input meaningless (if for example, all the companies had the same input value.)

The decision was made to apply a modified categorical model using successive data sets that included: 1) all project

data; 2) only projects with higher than 40% complete plans and specifications; 3) only projects with higher than 75% complete plans and specifications; and finally, 4) only projects with higher than 95% plans and specifications (in other words only those that had the design complete.) Mathematically this revision to the basic DEA formulation in vector form looks like:

$$\begin{aligned}
 &\text{Min: } \theta \\
 &\text{Subject to: } \quad \mathbf{Y}\lambda \geq \mathbf{Y}_0 \\
 &\quad \quad \quad \mathbf{X}\lambda \leq \theta\mathbf{X}_0 \\
 &\quad \quad \quad \lambda_i \geq 0 \\
 &\quad \quad \quad \lambda_i = 0 \quad \forall i \text{ such that } C_i < C_0
 \end{aligned}$$

- Define C_i = 1: Conceptual Design: <50% Complete
 2: Schematic Design: 50% to 75% Complete
 3: Preliminary Design: 75% to 95% Complete
 4: Final Design: 95% to 100% Complete

The controlling direction of the analysis was guided by state statute. In Oregon, as nearly every other state, on public building projects, the state gives the priority to open public bidding and discourages closed negotiated procurements except when it can be shown to be a substantial benefit to the public. The relevant portions of the particular Oregon statute, ORS 270.015 is as follows:

279.015 Competitive bidding; exceptions; exemptions.
 (1) Subject to the policies and provisions of ORS 279.005 and 279.007, all public contracts shall be based upon competitive bids or proposals except:

- (2) Subject to subsection (6)(b) of this section, the Director of the Oregon Department of Administrative Services or a local contract review board may exempt certain public contracts or classes of public contracts from the competitive bidding requirements of subsection (1) of this section upon approval of the following

¹⁷ One reason for limiting the projects to a certain size, larger than \$5 million, was because we knew that only construction firms with substantial resources and experience can qualify for Miller Act, performance and payment bonds, for that size of work.

findings submitted by the public contracting agency seeking the exemption:

(b) The awarding of public contracts pursuant to the exemption will result in substantial cost savings to the public contracting agency. In making such finding, the director or board may consider the type, cost, amount of the contract, number of persons available to bid and such other factors as may be deemed appropriate.

(3)(a) Before final adoption of the findings required by subsection (2) of this section exempting a contract for a public improvement from the requirement of competitive bidding, a public agency shall hold a public hearing.

It is clear from these sections of the statute that the Oregon Legislature intended to make the option of exempting from bidding and negotiating public building contracts a difficult and well reasoned alternative to open competitive public bidding, but certainly an option. Since the state has established the baseline PDS to be DBB, then the negotiated procurements, the CM/GC projects, must be considered the “challenger.” Therefore, the challenger, which is disfavored in the statute, has the burden to show it is superior against the baseline, whereas the baseline has no such burden¹⁸. It is reasonable to argue that this being the case, there is no need to evaluate the different levels of information in terms of the

Percent Complete metric, rather, simply run all CM/GC jobs categorically against all competitors, and then run only the DBB projects to establish their efficiency scores. While this obviously can be done, we wanted a finer break-out from the analysis.

D. Results from the Modified DEA Model

We did not recode the existing computer software instead the Modified DEA Model was run using EMS[®] Software from project data stored in an Excel[®] spreadsheet. Four passes were made on the data each using a CCR, Input Oriented Model¹⁹. The first pass included all data in the project database but was only used to calculate the DEA scores of the projects with less than 50% plans and specifications. The second pass included all the data for projects with plans and specifications more than 50% complete, but again was used only to calculate the DEA scores for projects with plans and specifications between 50% and 75% complete. The third pass included all projects with plans and specifications more than 75% complete and was used to calculate the DEA efficiency scores for projects with plans and specifications between 75% and 95% complete. Finally, the fourth pass included only projects with plans and specifications more than 95% complete and was used to calculate these project’s DEA efficiency scores.

The results of this process is depicted in Table 10 below:

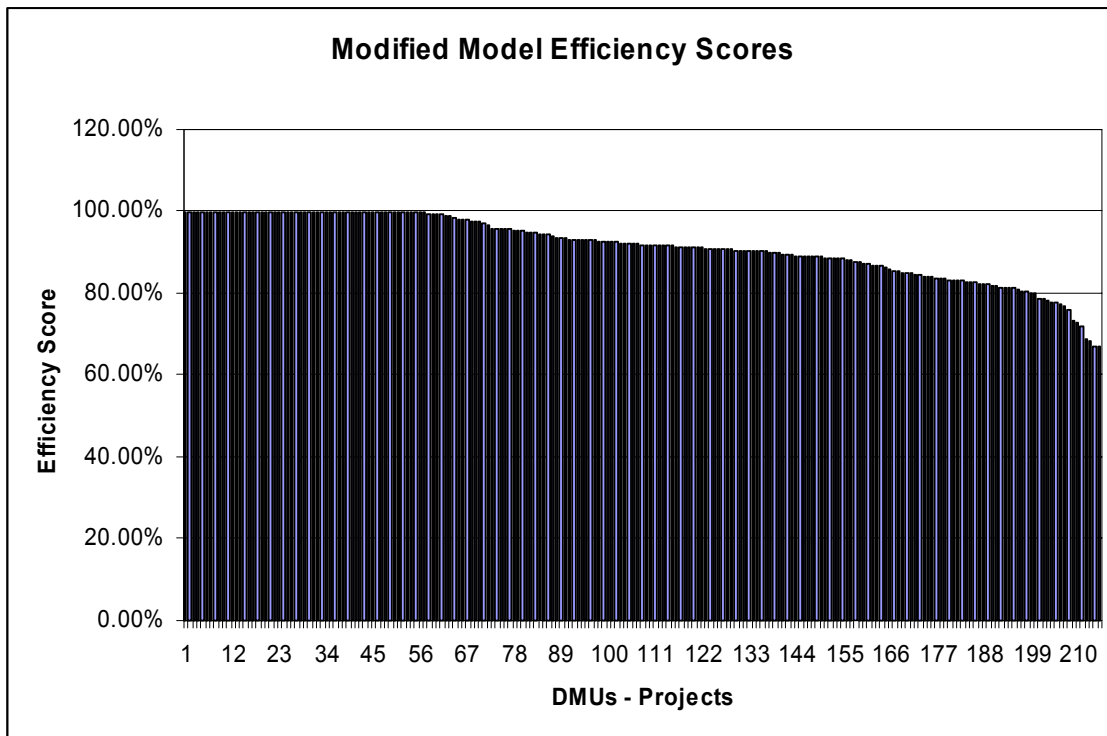


Figure 10 Modified Model Efficiency Scores

¹⁸ This argument is drawn directly from the plain language wording of §2(B) of ORS 279.015

¹⁹ All of these passes were made without regard to the quality of the information metric SF/(RFI+AR)

The DEA efficiency scores from the Modified Model, range from approximately 67% to 100% and have an arithmetic mean of 91.52%²⁰, as presented in Table 3 below. Also shown is a single frequency histogram of the DEA efficiency scores from the Modified Model, presented in Figure 11.

TABLE 3 DESCRIPTIVE STATISTICS OF DEA SCORES FROM MODIFIED MODEL

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
SCORE	215	67%	100%	91.52%	7.759%
Valid N (listwise)	215				

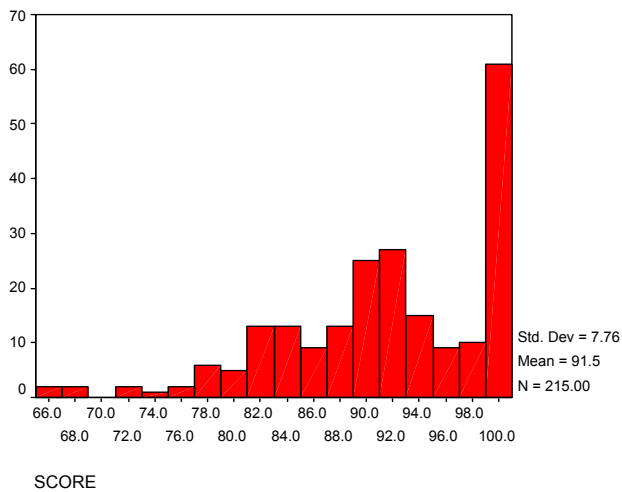


Figure 11 Histogram of all DEA efficiency Scores from the Modified Model

The results obtained by the Modified Model are more consistent with the Expert Panel’s intuitive understanding of construction project performance than the results obtained in the earlier DEA models depicted in Figure 8. Fifty-seven of the 215 projects in the final data set were determined to be 100% efficient, and 140 projects scored 90% or higher.

E. Evaluating the Project Delivery Systems

The next step in our analysis was to evaluate the two project delivery systems, DBB (Design Bid Build or Lump Sum Bid) and CM/GC (Construction Management/General Contractor the Negotiated Procurement method.) This was done by applying normal statistical methods, reserving of course, the same caveats about statistical analysis of DEA score distributions that was previously discussed. Both the group statistics and the tests for independence are presented in Table 4 and Table 5 respectively²¹.

TABLE 4 GROUP STATISTICS OF DEA SCORES FROM THE MODIFIED MODEL

Group Statistics					
	TYPE	N	Mean	Std. Deviation	Std. Error Mean
SCORE	C	111	92.50%	8.592%	.815%
	L	104	90.48%	6.643%	.651%

TABLE 5 TEST FOR INDEPENDENT POPULATIONS FOR PDS BASED ON DEA SCORES FROM THE MODIFIED MODEL

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
SCORE	Equal variances assumed	9.897	.002	1.921	213	.056	2.02%	1.052%	-.052%	4.096%
	Equal variances not assumed			1.937	205.718	.054	2.02%	1.044%	-.036%	4.080%

The statistical analysis of the DEA scores gives similar results as that produced on the base output data analysis presented in Table 2. That is, in spite of the fact the CM/GC negotiated procurements have a slightly higher average there is no statistically significant difference between the two PDS groups in terms of DEA scores from the Modified Model.

Since DEA is a non-parametric method and the DEA scores are distinctly not normally distributed, another way of evaluating any difference in the two populations is to actually look at the distributions and apply non-parametric techniques. Observing these two populations we note that the CM/GC projects have a higher proportion of 100% efficient projects than do the DBB jobs. In fact, 41 of 111 CM/GC projects were determined to be 100% efficient, whereas only 15 of the 104 DBB jobs were determined to be 100% efficient. However this difference becomes less distinct when you compare all projects with 90% or better efficiency scores; in that case, 64 of the 104 DBB jobs scored better than 90% efficient, while 74 of the 111 CM/GC projects scored 90% or

²⁰ DEA scores are known to be non-normally distributed however both the arithmetic mean and the standard deviation do provide us with valuable information about the distribution of the DEA data.

²¹ Note in Table 4, “C” is used for CM/GC projects and “L” is used for Lump Sum Bid jobs.

better. And, on the other end of the spectrum, six of the 111 CM/GC projects scored less than 75% efficient, whereas only one of the DBB jobs scored lower than 75%.

IV. SUMMARY & CONCLUSIONS

A. Summary of Analysis

This research sets forth a method to deal with data that is both imprecise but known to exist within certain limits and is non-linearly related to output performance in the Data Envelopment Method model. The research uses input information metrics to evaluate the performance of construction projects based on two different project delivery methods: Design Bid Build, the common method of open competitive lump sum low bid selection; and Construction Manager/General Contractor, a negotiated public construction procurement method used extensively in Oregon since 1986.

In order to compare the two methods the data was broken up into different input categories basically reflecting the different stages of design: Conceptual Design, Schematic Design, Preliminary Design, and Final Design, while keeping the precise estimates provide in the data set from the different data sources (Construction Contractor, Architect, and Owner, project management personnel assigned to the specific jobs.) Since public policy in Oregon favors open competitive lump sum bidding, the CM/GC projects were used as "Challengers" and competed against the bid projects to determine their DEA scores, whereas the bid projects generally competed against only other bid jobs.

The results of this research show that there is no statistically significant difference between the procurement methods based on an analysis of either the raw output metrics or the DEA efficiency scores from our Modified Model. However, analysis of the populations does indicate that the CM/GC projects have both a higher likelihood of being 100% efficient than the DBB projects, and a higher likelihood of being below 75% efficient than the DBB projects. Both methods appear to produce about the same proportion of projects in the 90% or better range, which is probably "acceptable" for public construction projects.

B. Conclusions

In a certain limited number of instances, the input metrics for a DEA model are not capable of properly capturing the total amount of resources available to produce performance outputs because they are either: unknown, not precisely measurable or highly non-linear, in terms of inputs to outputs. This is the case in the public construction contracting sector of the economy, where the inputs include not only what the Contractor is given by the Owner in terms of plans and specifications in order to determine a price, but also must include his knowledge and experience both in the type of work and the local market. And while it's probably true that the more experienced contractors with greater the knowledge are probably better able to convert information provided by the Owner into inputs for the DEA Construction

Project model, that process is neither well understood nor easily measured. Therefore a modified categorical variable approach was developed and used to overcome these shortcomings.

The principal conclusion that we draw from this research is that there is no significant difference between the two project delivery systems either in terms of their DEA scores or the statistical analysis of the raw output metrics for schedule and budget performance. Since the public policy of the State of Oregon asserts a preference for competitive open public bidding except in specific cases provided for under the exemption clause, namely a significant financial benefit to the public, this research under those conditions does not support the use of negotiated public procurements. However, it is not so much our view that the negotiated procurement method is bad, but rather that the policy is wrong. In fact, it is probably best to look at our analysis from the reverse side of the public policy, that is: since there is no significant difference between the two populations, the public pays no additional price for using the negotiated procurement method instead of the competitive open public bid method. That conclusion recommends a reexamination of the public policy exemption criteria, which most certainly needs to be changed.

C. Future Work

This research is neither the final word on this DEA application nor the public policy debate regarding the procurement of construction services for public projects. Clearly there is a need to better understand the process by which certain types of information are converted into inputs for DEA models in order to apply DEA to a broader set of service sector applications. The process construction contractor's undergo is similar to the process that any professional consultant undergoes in order to produce a product or outcome for an owner. Whether that person is an Accountant, Architect or Engineer, all are given a varying amount of information from their client, which they convert into useful inputs through some "process model" and eventually into outputs. The same perhaps can be said of physicians treating patients. Absent that understanding, the linkage between the information and the outputs in the DEA model, are far to non-linear to produce acceptable and reasonable results using DEA.

As for the policy debate regarding construction contracting project delivery methods, it is clear from our research that the legislature has to come to grips with the fact that the current policy is simply unworkable since it cannot be supported by either empirical or theoretical data analysis. The focus of this effort should be in better aligning the policy goals with the benefits that can be achieved by the different methods. The fact that the public pays no additional cost for starting a project at the conceptual or preliminary design phases is a clear indication of the advantages presented by the method that is totally ignored in the policy statement.

REFERENCES

- [1] J. Zhu, *Quantitative Models for Performance Evaluation and Benchmarking*, 1st ed. Norwell, Mass.: Kluwer Academic Publishers, 2003.
- [2] R. D. Banker, "Estimating most productive scale size using data envelopment analysis," *European Journal of Operational Research*, vol. 17, pp. 35-44, 1984.
- [3] S. J. Comen, "Basic Principles of Bidding," in *Construction Bidding Law*, vol. 1, *The Construction Law Library*, R. F. Cushman and W. J. Doyle, Eds. New York, NY: Aspen Law & Business, 1990, pp. 1-16.
- [4] M. S. Rhodes, *Construction Bidding Law - 1999 Cumulative Supplement*, 2nd ed. New York, NY: Aspen Law & Business, 1999.
- [5] J. H. Kagel and D. Levin, *Common Value Auctions and the Winner's Curse*. Princeton, NJ: Princeton University Press, 2002.
- [6] G. Tavares, "A Bibliography of Data Envelopment Analysis (1978-2001)," RUTCOR, Rutgers University, Piscataway, NJ RRR 01-02, JANUARY, 2002 2002.
- [7] L. M. Seiford, "A bibliography for Data Envelopment Analysis (1978-1996)," *Annals of Operations Research*, vol. 73, pp. 393-438, 1997.
- [8] A. Charnes, W. W. Cooper, A. Y. Lewin, and L. M. Seiford, "Data Envelopment Analysis: Theory, Methodology and Applications." Boston: Kluwer, 1994.
- [9] W. W. Cooper, L. M. Seiford, and K. Tone, *Data Envelopment Analysis - A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. Norwell, Mass: Kluwer Academic Publishers, 2000.
- [10] A. Charnes, W. W. Cooper, and E. Rhodes, "Measuring the efficiency of decision making units," *European Journal of Operational Research*, vol. 2, pp. 429-44, 1978.
- [11] N. Sarafoglou and K. E. Haynes, "Regional efficiencies of building sector research in Sweden: an introduction," *Computers, Environment and Urban Systems*, vol. 14, pp. 117-132, 1990.
- [12] L. Hjalmarsson and J. Odeck, "Efficiency of trucks in road construction and maintenance: An evaluation with data envelopment analysis," *Computers & Operations Research*, vol. 23, pp. 393-404, 1996.
- [13] J. O. Odeck, "Measuring Productivity Growth and Efficiency With Data Envelopment Analysis: an Application On the Norwegian Road Sector," Goteborgs Universitet (Sweden), 1993, pp. 186.
- [14] A. A. H. Shash, "A Probabilistic Model for U.S. Nuclear Power Construction Times," in *Civil Engineering*. Austin, TX: University of Texas, 1988.
- [15] W. W. Cooper, K. S. Park, and Y. Gang, "IDEA and AR-IDEA: Models Dealing with Imprecise Data in DEA," *Management Science*, vol. 45, pp. 597-607, 1999.
- [16] J. Zhu, "Imprecise Data Envelopment Analysis," *European Journal of Operations Research*, vol. Volume 144, pp. 459-688, 2003.