STRATEGIES TO WIN CMAR AND DB PROPOSALS: Which Evaluation Criteria Best Differentiate between Competitors?

by
Simplar

Abstract

The procurement processes used in the alternative contracting methods of design-build (D-B) and construction manager at risk (CMAR) are heavily focused on best-value and qualifications-based selection. However, previous research had not examined the effectiveness of owners' evaluation criteria in differentiating between competing bidders. The objective of this study was to document the selection outcomes of the bidders in D-B and CMAR projects and identify which evaluation criteria have the greatest differentiation in scores for competing bidders. The results were compared with previous research on the procurement of architectural and engineering consultants and design-bid-build (D-B-B) contractors. The study sample consisted of 362 bidders for 63 D-B and CMAR projects from the United States and Canada. The statistical analysis results show that scores on interviews and technical proposals have the greatest differentiation, while cost proposal scores have minimal differentiation. These findings provide practical guidance for owners and bidders regarding how to prioritize evaluation criteria and how to respond to them.

Research Details

INTRODUCTION

Design-bid-build (D-B-B) remains the predominant delivery method for construction services in the United States, with low bid procurement being the predominant method for selecting construction firms (El Asmar et al. 2010). However, an increasing number of owners are inclined to adopt alternative contracting methods, such as design-build (D-B) and construction manager at risk (CMAR), for better project performance (El Asmar et al. 2016; Sullivan et al. 2017; Carpenter and Bausman 2016). Because procuring a qualified contractor greatly affects the performance of the project (El Wardani et al. 2006; Chini et al. 2018), owners employing alternative contracting methods typically use qualifications-based selection (QBS) or best-value (BV) selection. A major reason that owners choose alternate procurement models of QBS and BV is that the scope of work, drawings, and specifications provided in the request for proposal (RFP) are incomplete. Consequently, it is difficult for bidders to develop hard bids (Xia et al. 2012a). On the contrary, a well-defined scope can greatly increase project cost, schedule, and quality performance (Xia et al. 2016). Moreover, these selection models provide the opportunity for owners to evaluate the qualifications and expertise of bidders.

Numerous studies have documented the advantages of using selection models that incorporate qualifications-based criteria to procure a construction management and a design-build firms (Shrestha and Fernane, 2017; Tran et al. 2018). Furthermore, using appropriate evaluation criteria and weightings can help owners select the most qualified team for a project (AGC and National Association of State Facilities Administrators, 2008). Because of the benefits, BV and QBS have been endorsed as the preferred selection methods for alternate contracting methods by many professional organizations, including the Design Build Institute of America (DBIA), Construction Management Association of America (CMAA), and Associated General Contractors of America (AGC).

During the past decade, much research has been conducted on the performance of alternative contracting methods, prequalification models, and multicriteria selection models, with RFP content analysis and statistical models used to evaluate the performance. In contrast, little research is available on evaluation criteria scores and the differentiation (i.e., degree of spread) in scores when using the BV and QBS methods of alternative contracting methods. The first objective of this study was to determine the evaluation characteristics of the selected Construction Manager Team (CMT), which refers to the selected construction manager firm hired under the CMAR contract, and selected Design-Build Team (DBT), which refers to the selected design-build firm hired under the D-B contract. Furthermore, the second objective was to determine the differentiation resulting from owners' use of evaluation criteria during project procurement. For the third objective, the results of this study were then compared to the results of previous research on the procurement of architectural and engineering (AE) consultants and D-B-B contractors (hereafter referred to as Contractor).

This study's data came from 63 projects in the United States and Canada that were executed via alternative contracting methods and procured using the BV and QBS approaches. The data set included 30 CMAR project and 33 D-B projects. From the 63 projects, a total of 362 proposing bidder evaluation scores regarding commonly used evaluation criteria were analyzed; descriptive statistics were used to determine the evaluation characteristics of the selected CMTs and DBTs. The Kruskal-Wallis H test was used to determine the differentiation in scores for each evaluation criterion. The results show that in BV-procured projects, the lowest CMT and DBT bidders were selected half of the time. Furthermore, the selected CMTs and DBTs typically have lower costs and slower schedules than the overall average for bidders. The greatest differentiation in scores was found in the categories of technical proposal and schedule. Additionally, across different prime vendors, all evaluation criteria except for schedule and past performance were found to have varying degrees of differentiation.

LITERATURE REVIEW

Although low bid procurement has been the predominant selection method for construction services in the United States, it has several noted disadvantages. At times, low bid results in the selection of contractors that submit unrealistically low bids, which can result in change orders, schedule delays, and disputes during the construction phase (Ioannou and Awwad 2010; Rosenfeld 2014). Similarly, owners in the United States are increasingly reluctant to use low bid procurement for some projects, which have become ever more complex in structure and design. Therefore, more owners are focusing on qualifications-based evaluation criteria (Yu et al. 2013) which has led to increasing use of D-B and CMAR

(Sindhu et al. 2018; Bilbo et al. 2015) to achieve better outcomes regarding cost, time, and quality (Konchar and Sanvido 1998; Hale et al. 2009; Xia et al. 2012b). Owners have also started using BV selection methods, which consider not only price but also expertise and other qualifications required to complete a project (Perrenoud et al. 2017; Ballesteros-Pérez et al. 2015). The BV approach provides an alternate to the low bid procurement by focusing on multiple evaluation criteria, such as time, operation and maintenance, quality-related aspects, and environmental aspects (Gransberg and Shane 2015). As with the BV approach, the QBS approach evaluates multiple criteria, and this approach has yielded lower cost growth (El Wardani et al. 2006) and fewer time overruns (Chen et al. 2016) than any other methods.

Many professional organizations have endorsed the use of QBS criteria and have strongly opposed the use of the low bid procurement (Christodoulou et al. 2004). For example, the DBIA (2012) has stressed the importance of using the QBS approach to procure DBTs because the approach leads to selecting a team that is proactive and collaborative. Likewise, the CMAA (2017), which considered construction management to be a professional service, advocated the use of the QBS approach for the public and private industries. Similarly, the AGC Project Delivery Committee QBS Working Group (2009) endorsed the use of the QBS and BV approaches for procuring contractors. State transportation departments are also increasing their use of the BV and QBS approaches (Gransberg and Shane 2015). Moreover, previous research shows that different contractors possess different qualifications and expertise; therefore, because the BV selection model considers cost and qualifications, it is advantageous for owners (Abdelrahman et al. 2008; Akintoye et al. 2003; Perng et al. 2006). Similarly, researchers recommend using the BV approach when selecting CMTs and DBTs, because the approach allows more weight to be assigned to qualification criteria than to cost (Gransberg and Shane 2015; Xia et al. 2013; Del Puerto et al. 2008).

Previous research on alternative contracting methods has mainly focused on project performance and multicriteria decision models. Many of these studies have compared the costs, project durations, and quality performance of alternative contracting methods to the traditional D-B-B system (Hyun et al. 2008; Hale et al. 2009; Rosner et al. 2009; Minchin et al. 2013). Other studies have analyzed the costs, project durations, and quality performance of alternative contracting methods in relation to different project scopes, project types, contract types, and selection methods (El Wardani et al. 2006; Xia et al. 2012a; Bogus et al. 2013; Chen et al. 2016). Furthermore, researchers have developed multicriteria decision models for procuring bidders for projects involving alternative contracting methods (Molenaar and Songer 1998; Abdelrahman et al. 2008; El Sayegh 2009; El Asmar et al. 2010).

Limited research is available on the evaluation characteristics of the selected bidders in CMAR and D-B projects and the differentiation identified through using evaluation criteria for the BV and QBS approaches for alternative contracting methods. This article examines the evaluation characteristics of selected CMTs and DBTs and the differentiation identified through using multiple evaluation criteria. The study sample included 63 projects, which received a total of 362 bids, from the United States and Canada. The bidders' scores and the evaluation criteria weights were analyzed using descriptive statistics. The Kruskal-Wallis H test was used to identify differentiation in evaluation criteria scores. Furthermore, common evaluation criteria scores were compared to identify differentiation among the prime vendors in the areas of architecture, engineering, and construction (AEC); AE; CMT; DBT; and Contractor.

RESEARCH OBJECTIVES AND HYPOTHESES

Point of Departure

Little research has been conducted to determine the evaluation characteristics of the selected bidders in CMAR and D-B projects or to identify the differentiation between the competing bidders for various evaluation criteria. Existing research has mainly focused on content analysis of owner RFPs, such as the weighting schemes of evaluation criteria, multicriteria selection procedures, prequalification processes, and project performance after contract award. In the few studies on the evaluation characteristics of selected bidders and their ranking across evaluation criteria (Christodoulou et al. 2004; Chinowsky and Kingsley 2009), the samples were relatively small or limited to a specific geographic area.

This study addresses the gap in the literature by analyzing data from a relatively large sample: 63 projects with 362 bidders from the United States and Canada. The projects were owned by public agencies at the state, municipal, and university levels across the United States and Canada. The project data were analyzed to determine the evaluation characteristics of the selected CMTs and DBTs, as well as to identify the differentiation in evaluation scores achieved by the evaluation criteria. The results of this study were then compared to previous research of selection outcomes for AE consultants and Contractors to determine similarities and differences within these contexts.

Research Question 1

The first research objective was to analyze the CMT and DBT bidders that were selected for the 63 projects. This objective led to the development of the first research question,

RQ1) What were the evaluation characteristics of selected CMTs and DBTs relative to competing bidders?

Answering RQ1 involved examining how the selected CMTs and DBTs ranked in terms of the evaluation criteria, as well as what qualifications and expertise the selected CMTs and DBTs brought to a project compared to the bidders with the lowest prices and with average prices. These objectives were achieved by using descriptive statistics to evaluate the data for the lowest-bidders, best-qualified bidders, and best-evaluation-score bidders that were selected. Best-qualified bidder refers to bidders who were best in qualification and were selected, and best-evaluation-score bidders were those who emerged as the top ranked bidder in the evaluation process. Descriptive statistics were also used to determine, for each evaluation criterion, the selected bidders' rankings, selected bidders' scores, average bidder scores, differentials from average bidder scores, and differentials from lowest-bidder scores. The differentials indicate the added value that selected bidders brought to projects, compared to average bidders and lowest-bidders.

Research Question 2

Because owners use a combination of qualifications and cost criteria for procuring CMTs and DBTs, it is important to determine the differentiation in scores for different evaluation criteria. This information can be used to identify which evaluation criteria have the largest and smallest spread in scores. Therefore, the

second research objective was to determine this information, and RQ2 and an accompanying hypothesis were developed:

RQ2: What is the amount of differentiation in scores for individual evaluation criteria, and which evaluation criterion is associated with the greatest differentiation?

H10: Different evaluation criteria have the same amount of differentiation in scores for CMTs and DBTs.

The coefficient of variation (CV) was used to measure the spread in evaluation criteria scores. The Kruskal-Wallis H test was used to determine the difference in CV values for a specific evaluation criterion and to order the criteria based on the amount of differentiation identified.

Research Question 3

The third research objective was to compare the results of this study with previous research on BV procurement of Contractors (Nguyen et al. 2018) and AE consultants (Lines and Shalwani 2017). This objective led to the development of RQ3 and its accompanying hypothesis:

RQ3: Is differentiation in evaluation criteria scores similar for different categories of prime vendors? If not, which evaluation criterion has the greatest differentiation?

H20: Differentiation in evaluation criteria scores is similar in different categories of prime vendors.

The Kruskal-Wallis H test was used to determine the differences in the CV values for the evaluation criteria for each set of prime vendors.

METHODOLOGY

Data Collection

Procurement data were collected from 63 public agencies in the United States and Canada. Thirty of the projects were CMAR and the remaining 33 were D-B. All projects were procured by public owners in the institutional sector, including states, municipalities, and elementary, secondary, and post-secondary school systems. Of the 63 projects, 28 were procured using the QBS approach, and the other 35 were procured using the two-envelope BV approach, with the qualifications proposal separated from the cost proposal. Regardless of the procurement method, the evaluation criteria specified in the solicitation documents were similar across the projects.

Within the sample size, it was ensured that the projects had similar construction scopes, facility types, and procurement procedures. The projects in the data only comprised of vertical construction in public institutional facilities. The total cost of all 63 projects were a little over $610 million, with an average construction cost $10 million. The data collected for each project included the evaluation criteria weights, owner-evaluation-team scores for each evaluation criterion for each bidder, and the mathematical model

used to select the best-evaluation-score bidder for the project. Later, the project data were compared to data from previous similar studies on AE consultants and Contractors.

A review of the projects' solicitation documents indicates that the procurement process starts with the release of an RFP or a request for qualifications, followed shortly thereafter by a presubmittal meeting. Once the proposals are submitted, a technical proposal with related experience is sent to the evaluation team for evaluation. After the evaluation team scores all the bidders, the top three to five scorers (based on criteria including cost and schedule) are invited to participate in the interview process. The bidder with the highest score at the end of the interview process is selected for the project. For QBS projects, the cost is negotiated with the highest scorer.

Variables

The variables used in this study were the evaluation method and the owner-evaluation-team scores for each of the evaluation criteria. CV was used to measure the spread in scores for each of the evaluation criteria, and the Kruskal-Wallis H test was used to statistically test the difference in CV values for all evaluation criteria.

Evaluation Criteria

This study focused on six common evaluation criteria: cost, schedule, technical proposal, past performance, interviews, and related experience. Scores for each of these criteria were based on a scale of 1–10, with 1 denoting the lowest score possible.

Cost. Cost was evaluated on the basis of the lowest bid. The bidder with the lowest cost was assigned the evaluation score of 10; the other bidders were rated using inverse linear proportions.

Schedule. The schedule criterion regarded the expected duration of the project. The score of 10 was assigned to the bidder with the shortest schedule; the other bidders received linearly prorated evaluation scores.

Technical proposal. Each bidder's technical proposal consisted of a brief written summary of the proposed execution plan, including the means and method, potential project risks, and value engineering options. Evaluation team was responsible to score technical proposal as part of the procurement process.
Past performance. Past performance regards the bidders' previous experience with similar projects and clients. For this study, past performance was evaluated in terms of previous clients' satisfaction with the bidders.

Interviews. Interviews were conducted with key personnel in each shortlisted bidder's proposed project team. Typically, the short lists included the top three to five bidders. Evaluation team was responsible to score interviews as part of the procurement process.

Related experience. Each bidder submitted a summary of its previous projects that were most similar to the project being bid on. Usually, bidders listed projects that were renowned, large, or for notable clients. Evaluation team was responsible to score related experience as part of the procurement process.

Other criteria. Some owners used additional evaluation criteria, such as regarding safety prequalifications, Disadvantaged/Minority/Women's Business Enterprise program (D/M/WBE), and specific team qualifications. These evaluation criteria were not included in this study's analysis because they were not considered in all projects and, when considered, were typically assigned fairly low weightings.

Coefficient of Variation

CV was used to measure the differentiation, or variability relative to the mean, in evaluation criteria scores. Mathematically, CV is the ratio of the standard deviation to the mean. Because CV does not have a specific unit of measurement, the differentiation for various evaluation criteria (which have varying units of measurement) could be compared. The higher the CV value, the more the spread in scores.

METHOD OF ANALYSIS

Descriptive Analysis of CMTs and DBTs

Descriptive statistics were used to answer RQ1, which focused on determining the evaluation characteristics of the selected CMTs and DBTs, based the evaluation criteria scores. The evaluation scores for all projects were reviewed to identify how frequently the lowest-bidder, the best-qualified bidder, or the best-evaluation-score bidder was selected. The differentials between the selected CMTs and DBTs compared to the average bidders and the lowest-bidders were also evaluated through descriptive statistics.

Kruskal-Wallis H Test for Differentiation

The Kruskal-Wallis H test was used to answer RQ2 and RQ3, which regarded the level of differentiation in scores for individual evaluation criteria and whether differentiation varied by type of prime vendor. The Kruskal-Wallis H test, which is also known as the one-way ANOVA ranks test, is a rank-based nonparametric test. Kruskal-Wallis H test can be used when the data is not normally distributed (Arditi et al. 2009) and more than two groups of data are needed to be compared (Xia et al. 2013). Pairwise comparison was used as a post hoc test for the Kruskal-Wallis H test.

RESULTS

Research Objective One: Evaluation Characteristics of the Selected CMTs and DBTs

The descriptive statistics indicate that the results for the CMT projects and the DBT projects were similar with each other. Analysis of the owner solicitations indicates that overall, qualifications-based criteria

were more heavily weighted than were cost and schedule criteria. Table 1 shows the evaluation criteria weightings used in the 63 projects. Evaluation criteria weighting were published in the RFP released by the owners. Cost had a median weight of 20% for CMTs and DBTs, whereas qualifications-based criteria were usually assigned weights of 70%–80%. Generally, interviews were assigned the most weight, while schedule was assigned the least weight.

The evaluation scores for all projects were reviewed to identify how frequently the lowest-bidder or the best-qualified bidder was selected. Almost two-thirds of the time or more, the best-qualified CMT or DBT was selected. Table 2 shows the frequency with which the selected CMTs and DBTs were the lowest-bidders, best-qualified bidders, and best-evaluation-score bidders. Although, the owners designed their procurement processes to focus on selecting the best-evaluation-score CMT or DBT, owners did not always do that. Closer inspection of the data indicate that when these bidders were not selected, either they bid substantially higher costs or the owners used different mathematical models than usual for awarding contracts.

The differences in the evaluation rankings and scores for the selected CMTs and DBTs compared to the other bidders were also analyzed. Table 3 shows the rankings and scores for the selected CMTs and DBTs, the differentials from the average-bid scores, and the differentials from the lowest-biddder scores. On average, the selected CMTs ranked first in the interview category; second in cost, schedule, technical proposal, and related experience; and third in past performance. Similarly, the selected DBT were ranked, on average, first in schedule, interviews, and relevant experience, and second in the other three evaluation criteria. Further, the selected CMTs' proposed schedules were longer than the average bidders' schedules and the lowest-bidders' schedules. The selected DBTs' proposed schedules were shorter than the lowest-bidders' but were longer than the average bidder proposed schedule duration. The selected CMTs and DBTs bid lower costs than the average.

Table 1. Weighting Schemes of Evaluation Criteria for CMT and DBT Solicitation

Evaluation criteria	CMT						DBT					
	N	Mean (%)	Median (%)	Max. (%)	Min. (%)	Std. dev. (%)	*N*	Mean (%)	Median (%)	Max. (%)	Min. (%)	Std. dev. (%)
Cost	6	18.7	20.0	28.6	0.0	6.5	9	19.8	20.0	30.4	0.0	7.1
Schedule	13	6.9	6.3	12.5	5.0	2.5	21	9.8	5.0	25.0	2.5	7.1
Technical proposal	30	32.6	30.0	55.0	20.0	8.6	33	30.4	30.0	46.7	17.5	6.6
Past performance	30	19.9	20.0	30.0	5.0	9.5	33	21.1	20.0	45.0	5.0	10.7
Interview	23	36.0	40.0	45.0	20.0	8.2	23	32.4	35.0	45.0	12.5	10.0
Related experience	6	15.9	16.9	25.0	5.0	7.4	4	15.5	17.7	20.0	5.0	6.2
Other criteria	14	9.98	6.25	25	5	6.8	19	8.08	5	20	5	5.3

Table 2. Price and Qualifications Rankings of the Selected CMTs and DBTs

Prime vendor	Procurement type	*N*	Lowest-bidder (%)	Best-qualified bidders (%)	Best-evaluation-score bidder (%)
CMT	BV	6	50.0	66.7	100.0
	QBS	16	N/A	87.5	87.5
DBT	BV	9	44.4	88.9	88.9
	QBS	12	N/A	83.3	83.3

Note: For CMT, 33% of the time the lowest-bidder was also the best-qualified; for DBT, 44% of the time the lowest-bidder was also the best-qualified.

Table 3. Differences in Evaluation Scores of the Selected CMTs and DBTs vs. Competing Bidders

Evaluation criteria	Average ranking of selected bidders	Average evaluation score of selected bidders	Average evaluation score of nonselected bidders	Differential from avg. bidder (%)	Differential from lowest-bidder (%)
		CMT			
Cost	2	—	—	−11.9	9.1
Schedule	2	—	—	12.7	25.0
Technical proposal	2	70	60	16.9	22.0
Past performance	3	95	94	1.5	18.7
Interview	1	80	74	12.7	100.0
Related experience	2	90	81	11.6	15.0
		DBT			
Cost	2	—	—	−3.8	17.0
Schedule	1	—	—	13.0	−46.7
Technical proposal	2	74	62	20.8	26.2
Past performance	2	91	89	1.9	10.4
Interview	1	84	72	19.2	95.1
Related experience	1	100	80	19.7	29.0

Research Objective Two: Differentiation Regarding CMT and DBT Qualifications

The other objective of the study was to determine whether the amount of differentiation varies by evaluation criterion. To determine the differentiation regarding a criterion, median CV values were calculated. Table 4 shows the descriptive statistics regarding the CVs for the evaluation criteria used for CMTs and DBTs. For CMTs, the technical proposal criterion had the highest median CV value (18.5%), followed closely by schedule (18.4%). For DBTs, schedule had the highest median CV value (21.3%), followed by related experience (18.8%). Past performance had the least differentiation (2.7% for CMTs; 5.8% for DBTs).

The Kruskal-Wallis H test was conducted regarding the median CV values for CMTs and DBTs. Regarding the CMTs, the result was significant, $\chi 2 = 32.875$, $p = 0.000$. A post-hoc test using pairwise comparison was performed to determine which, if any of the evaluation criteria, had different differentials. Table 5 shows the results from the post hoc test for CMTs. The results indicate that the differential for past performance varies a statistically significant degree from technical proposal ($p = 0.000$) and schedule ($p = 0.001$). Based on these results, the evaluation criteria were categorized as having great, moderate, or minimal differentiation. Technical proposal and schedule had great score differentiation. Related experience, cost, and interview had moderate score differentiation. Past performance had minimal score differentiation.

The post hoc test for DBTs also had a significant result, $\chi 2 = 20.822$, $p = 0.001$. Again, past performance was found to have a statistically significant difference with technical proposal (p value = 0.003) and schedule ($p = 0.008$), as shown in Table 6. Based on the results, the evaluation criteria were categorized as having great, moderate, or minimal differentiation. Technical proposal and schedule had great score differentiation. Related experience, cost, and interview had moderate score differentiation. Past performance had minimal score differentiation.

Table 4. Variations in Evaluation Scores for Competing Bidders

Evaluation criteria	*N*	Per project average score	Min. (%)	Max. (%)	Mean (%)	Median (%)	Std. dev. (%)
			CMT				
Cost	6	—	0.1	31.8	14.8	14.4	13.8
schedule	13	—	0.0	36.9	18.5	18.4	10.5
Technical Proposal	30	70	2.1	37.5	17.7	18.5	9.3
Past performance	30	95	0.5	17.3	4.7	2.7	4.4
Interview	23	80	0.6	49.2	12.5	8.8	12.9
Related experience	6	90	0.0	31.8	15.5	12.9	12.6
			DBT				
Cost	9	—	0.7	16.9	9.3	10.6	6.0
schedule	21	—	0.0	69.0	24.0	21.3	19.8
Technical Proposal	33	74	4.8	46.7	18.3	15.2	11.4
Past performance	33	91	1.0	32.8	8.8	5.8	8.4
Interview	23	84	2.5	44.5	16.4	10.3	12.9
Related experience	4	100	5.6	38.0	20.3	18.8	15.2

Table 5. Post Hoc Test Results for CMT Evaluation Criteria Scores

Base evaluation criteria	Base median CV (%)	Comparison evaluation criteria	Comparison median CV (%)	*p* value
Related experience	12.90	Past performance	2.70	0.382
	12.90	Technical proposal	18.45	1.000
	12.90	Interview	8.82	1.000
	12.90	Cost	14.35	1.000
	12.90	Schedule	18.36	1.000
Past performance	2.70	Technical proposal	18.45	0.000*
	2.70	Interview	8.82	0.158
	2.70	Cost	14.35	1.000
	2.70	Schedule	18.36	0.001*
Technical proposal	18.45	Interview	8.82	0.635
	18.45	Cost	14.35	1.000
	18.45	Schedule	18.36	1.000
Interview	8.82	Cost	14.35	1.000
	8.82	Schedule	18.36	1.000
Cost	14.35	Schedule	18.36	1.000

Note: * p <.01.

Table 6. Post Hoc Test Results for DBT Evaluation Criteria Scores

Base evaluation criteria	Base median CV (%)	Comparison evaluation criteria	Comparison median CV (%)	*p* value
Related experience	18.75	Past performance	5.80	1.000
	18.75	Technical proposal	15.20	1.000
	18.75	Interview	10.26	1.000
	18.75	Cost	10.60	1.000
	18.75	Schedule	21.29	1.000
Past performance	5.80	Technical proposal	15.20	0.003*
	5.80	Interview	10.26	0.298
	5.80	Cost	10.60	1.000
	5.80	Schedule	21.29	0.008*
Technical proposal	15.20	Interview	10.26	1.000
	15.20	Cost	10.60	1.000
	15.20	Schedule	21.29	1.000
Interview	10.26	Cost	10.60	1.000
	10.26	Schedule	21.29	1.000
Cost	10.60	Schedule	21.29	0.953

Note: * p <.01.

Research Objective Three: Comparison with Previous Literature of Procurement Evaluations in Design and Construction

RQ3 was answered by comparing the differentiation in evaluation criteria scores for prime vendors, CMTs, and DBTs with scores for AE consultants (Lines and Shalwani 2017) and Contractors (Nguyen et al. 2018). The Kruskal-Wallis H test was conducted to examine five evaluation criteria—cost, schedule, technical proposal, past performance, and interview—for prime vendors. The criterion of related experience was excluded from the analysis because it was not commonly used when procuring contractors. The analysis results show that cost differentiation varies to a statistically significant degree among prime vendors, $\chi 2 = 52.496$, $p = 0.000$. As Table 7 shows, the median CV for AE consultants (22.5%) is significantly higher than the median CVs for DBT (10.6%) and Contractors (10.1%).

Additionally, the differentiation in scores for technical proposal varied among prime vendors to a statistically significant degree, $\chi 2 = 22.864$, $p = 0.000$. Further analysis, via a post hoc test, show that the difference is statistically significant only for AE consultants and Contractors. Contractors' median CVs (21.2%) were higher than AEs' median CVs (15.0%). The results also indicate that the differentiation in prime vendors' scores for the interviews was statistically different. Further analysis, via a post hoc test, showed that CMTs' score differentiation varied to a statistically significant degree from AE consultants' and Contractors' score differentiation. Contractors' median CVs (20.7%) and AE consultants' median CVs (17.1%) were higher than CMTs' median CVs (8.8%).

There was no statistically significant difference in schedule and past performance for prime vendors. However, all prime vendors' median CVs for schedule were generally high, while median CVs for past performance were generally low.

Table 7. Post Hoc Test Results for Statistically Significant Differences in Evaluation Criteria Differentials for Prime Vendors

Base prime vendor	Base median (%)	Comparison prime vendor	Comparison median (%)	*p* value
		Cost		
AE	22.5	Contractor	10.0	0.000*
	22.5	CMT	14.4	0.384
	22.5	DBT	10.6	0.005*
Contractor	10.0	CMT	14.4	1.000
	10.0	DBT	10.6	1.000
CMT	14.4	DBT	10.6	1.000
		Technical proposal		
AE	15.0	Contractor	21.2	0.000*
	15.0	CMT	18.5	1.000
	15.0	DBT	15.2	1.000
Contractor	21.2	CMT	18.5	0.256
	21.2	DBT	15.2	0.132
CMT	18.5	DBT	15.2	1.000
		Interview		
AE	17.1	Contractor	20.7	1.000
	17.1	CMT	8.8	0.035*
	17.1	DBT	10.3	0.756
Contractor	20.7	CMT	8.8	0.004*
	20.7	DBT	10.3	0.225
CMT	8.8	DBT	10.3	1.000

Note: * $p < .01$

DISCUSSION

Research Objective One: Evaluation Characteristics of the Selected CMTs and DBTs

The selected CMTs and DBTs were found to have substantially higher scores for qualifications-based criteria and also were cost competitive. Compared with the average scores for competitors, the selected CMTs and DBTs received considerably more favorable evaluations for technical proposal and interviews, while also being lower in price. In cases in which the lowest-bidder was interviewed but not selected, the selected CMT or DBT was evaluated twice as favorably as the lowest-bidder in the interviews and also received a better technical proposal score. These findings indicate that the selected CMTs and DBTs had a wide range of qualifications. Therefore, it can be concluded that when owners appropriately use evaluation criteria in the procurement process, they are well positioned to select the most qualified bidder.

Research Objective Two: Differentiation Regarding CMT and DBT Qualifications

In D-B and CMAR, not all qualifications-based criteria yield similar procurement evaluation results. The results of the Kruskal-Wallis H test show that different evaluation criteria had different amounts of differentiation for CMT and DBT projects. Therefore, the researchers rejected the null hypothesis H10; that different evaluation criteria have the same amount of differentiation in scores for CMTs and DBTs, and accepted the alternate hypothesis. For both CMTs and DBTs, the criteria regarding technical proposal and schedule had great differentiation, whereas cost, interviews, and related experience had moderate differentiation; past performance had minimal differentiation. Scoring well in the technical proposal and schedule categories require bidders to have expertise in planning project-specific deliverables and how they will be delivered. Bidders' varying levels of expertise in this area resulted in higher differentiation in evaluation scores for these criteria. In contrast, the majority of bidders are able to list relevant experience, which results in minimal differentiation regarding that criterion.

Previous studies of Design-Build RFPs have shown that owners often treat cost, past performance, and technical proposals as three of the top four most heavily weighted evaluation criteria (Xia et al. 2013; Xia et al. 2012b). Similar results were found for this study, which reveals that they were three of the top four highly weighted evaluation criteria. However, the research findings suggest that both cost and technical proposals achieve the substantial differentiation among competing bidders and therefore should be weighted heavily. Conversely, past performance was found to result in the least differentiates and therefore should be weighted with lower priority. Because of the similarity in results for differentiation in scores for CMT and DBT, the authors recommend the same approach for the selection of Construction Managers.

Research Objective Three: Comparison with Previous Literature of Procurement Evaluations in Design and Construction

During the procurement, the level of scope provided in each of the delivery methods has an influence on the effectiveness of the evaluation criteria. The median CVs for prime vendors indicate that differentiation varied between some evaluation criteria. Therefore, the researchers rejected the null hypothesis H20;

that differentiation in evaluation criteria scores is similar in different categories of prime vendors and accepted the alternate hypothesis. The differentiation in the cost category was greater for AE consultants than for DBTs and Contractors. The difference may be result from different degrees of scope definition at the time of procurement. At the time that AE consultants are procured, either there is no design or the design is in a preliminary stage; therefore, bidders' cost proposals have greater differentiation. Typically, DBTs are procured when the design is between 15% and 60% developed, which enables bidders to provide more tailored cost proposals. For Contractors, the cost proposal has less differentiation, perhaps because the complete design is available.

For the technical proposal criterion, differentiation was greater for Contractors than for AE consultants. A well-defined scope can help owners differentiate between the qualifications of Contractors; otherwise, the differentiation is assumed to be based entirely on cost. In contrast, when procuring AE consultants, the design is not complete enough to differentiate the qualifications of consultants. Differentiation in interview scores was lower for CMTs than for AE consultants and Contractors. Greater differentiation in Contractors scores further bolster the inference made from previous results, that competing Contractors have different ranges of qualifications. Additionally, schedule and past performance scores are not significantly different for different types of prime vendors.

CONCLUSIONS AND CONTRIBUTIONS.

Conclusions

Despite the widespread use of alternative contracting methods, little research prior to the current study analyzed the evaluation characteristics of selected CMTs and DBTs in CMAR and D-B projects. This study not only identified the evaluation characteristics of selected CMTs and DBTs but also analyzed the level of differentiation in evaluation criteria scores. These results were also compared to data from previous research on the procurement processes for AE consultants and Contractors.

The data analysis shows that the qualifications criteria in the 63 projects were assigned more weight than the cost criterion was. Furthermore, in BV procurement for CMT and DBT, two-thirds of the time the best-qualified bidder was selected. Further analysis revealed that selected CMTs and DBTs proposed lower costs and longer schedules than the average. The results of the Kruskal-Wallis H test and the post hoc tests show that differentiation varies to a statistically significant degree for evaluation criteria scores for CMTs and DBTs. The criteria of technical proposal and schedule have the greatest differentiation in scores; cost, related experience, and interviews have moderate differentiation; and past performance has minimal differentiation. Therefore, the researchers rejected the null hypothesis that H10; that different evaluation criteria have the same amount of differentiation in scores for CMTs and DBTs and accepted the alternate hypothesis.

The differentiation in scores for different prime vendors—CMTs, DBTs, AE consultants, and Contractors—were also examined. The results show that the amount of differentiation varies among prime vendors regarding the criteria of cost, schedule, and interviews. Therefore, the researchers rejected the null

hypothesis H20; that differentiation in evaluation criteria scores is similar in different categories of prime vendors and accepted the alternate hypothesis.

Contributions

This study contributes to the AEC industry and the body of knowledge regarding alternative contracting method procurement. Few studies have empirically analyzed the evaluation scores for the selected CMTs and DBTs or the differentiation in scores when using common evaluation criteria. Previous studies on CMAR and D-B procurement typically had small sample sizes or were limited to specific geographic locations (Christodoulou et al. 2004; Chinowsky and Kingsley 2009). This study analyzed evaluation criteria scores to identify the characteristics of selected the CMTs and DBTs, as well as the differentiation in the scores. Further, the study had a relatively large sample size and geographic scope. The 63 projects, with 362 bidders, were from throughout the United States and Canada.

Construction managers and design-builders can use the findings of the study to improve their proposal responses by focusing on evaluation criteria that typically result in greater differentiation and, therefore, provide the greatest opportunity to stand out from low-qualified bidders. This study is important for owners because the results show that evaluation criteria scores differ based on the type of prime vendor being evaluated. Owners should ensure their procurement processes include evaluation criteria that best differentiate between competing bidders for various project stages and delivery methods.

LIMITATIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

Several limitations were present in this study. Because only the BV and QBS approaches were analyzed in this study, the results cannot be generalized to other procurement methods. Similarly, the results of this research only depict the result for vertical institutional projects in the United States and Canada. Further, the researchers did not have project performance data (schedule growth, cost overrun, quality, etc.); therefore, the relationship between procurement criteria and project performance was not analyzed. Future research could be conducted to analyze the relationship between the identification of highly qualified bidders and the overall project performance in order to determine how well the method of selection met the project objectives.

This study examined evaluation criteria that were commonly used for evaluating various types of prime vendors, but the researchers were cognizant that many owners use other qualifications criteria for procurement. In the future, research could analyze other qualifications criteria that are used in the AEC industry. Within this study, the owner's solicitation documents were typically limited to 5-7 evaluation criteria. However, other sectors of the construction industry may include a greater quantity of criteria; therefore, future research is recommended to consider the effect that the total number of individual evaluation criteria has on selection outcomes. Moreover, horizontal projects can be also be analyzed using similar research methodology. Lastly, a study similar to this one could be conducted to examine the interview scores assigned to the individuals in the project team.

REFERENCES

Abdelrahman, M., Zayed, T., and Elyamany, A. (2008). "Best-value model based on project specific characteristics." J. Constr. Eng. Manage., 134(3), 179–188.

AGC (Associated General Contractors of America) and National Association of State Facilities Administrators. (2008). "Best practices for use of best value selections." AGC and National Association of State Facilities Administrators, Arlington, VA.

AGC (Associated General Contractors of America) Project Delivery Committee QBS Working Group. (2009). "Qualifications based selection of contractors." AGC, Arlington, VA.

Arditi, D., Ongkasuwan, D., and Committee on Management Practices in Construction of the ASCE Construction Institute. (2009). "Duties and responsibilities of construction managers: Perceptions of parties involved in construction." J. Constr. Eng. Manage., 135(12), 1370–1374.

Ballesteros-Pérez, P., Skitmore, M., Pellicer, E., and González-Cruz, M. C. (2015). "Scoring rules and abnormally low bids criteria in construction tenders: A taxonomic review." Constr. Manage. Econ., 33(4), 259–278.

Bilbo, D., Bigelow, B., Escamilla, E., and Lockwood, C. (2015). "Comparison of construction manager at risk and integrated project delivery performance on healthcare projects: A comparative case study." Int. J. Constr. Educ. Res., 11(1), 40–53.

Bogus, S. M., Migliaccio, G. C., and Jin, R. (2013). "Study of the relationship between procurement duration and project performance in design-build projects: Comparison between water/wastewater and transportation sectors." J. Manage. Eng., 29(4), 382–391.

Carpenter, N., and Bausman, D. C. (2016). "Project delivery method performance for public school construction: design-bid-build versus cm at risk." J. Constr. Eng. Manage., 142(10): 05016009.

Chen, Q., Jin, Z., Xia, B., Wu, P., and Skitmore, M. (2016). "Time and cost performance of design-build projects." J. Constr. Eng. Manage., 142(2), 04015074.

Chini, A., Ptschelinzew, L., Minchin, R. E., Jr., Zhang, Y., and Shah, D. (2018). "Industry attitudes towards alternative contracting for highway construction in Florida." J. Manage. Eng., 34(2), 04017055.

Chinowsky, P. S., and Kingsley, G. A. (2009). "An analysis of issues pertaining to qualifications-based selection." American Council of Engineering Companies and American Public Works Association, Washington, DC.

Christodoulou, S., Griffis, F., Barrett, L., and Okungbowa, M. (2004). "Qualifications-based selection of professional A/E services." J. Manage. Eng., 10.1061/(ASCE)0742-597X(2004)20:2(34), 34–41.

CMAA (Construction Management Association of America). (2017). "CMAA calls for 'new ways of thinking' in dot comments." DOT Release, http://cmaanet.org/dot-release (July. 2, 2017).

DBIA (Design Build Institute of America). (2012). DBIA position statement qualification based selection, DBIA, Washington, DC.

Del Puerto, C. L., Gransberg, D. D., and Shane, J. S. (2008). "Comparative analysis of owner goals for design/build projects." J. Manage. Eng., 24(1), 32–39.

El Asmar, M., Hanna, A., and Loh, W. (2016). "Evaluating integrated project delivery using the project quarterback rating." J. Constr. Eng. Manage., 10.1061/(ASCE)CO.1943-7862.0001015,04015046.

El Asmar, M., Lotfallah, W., Whited, G., and Hanna, A. S. (2010). "Quantitative methods for design-build team selection." J. Constr. Eng. Manage., 136(8), 904–912.

El Sayegh, S. M. (2009). "Multi-criteria decision support model for selecting the appropriate construction management at risk firm." Constr. Manage. Econ., 27(4), 385–398.

El Wardani, M., Messner, J., and Horman, M. (2006). "Comparing procurement methods for design-build projects." J. Constr. Eng. Manage., 132(3), 230–238.

Gransberg, D. D., and Shane, S. J. (2015). "Defining best value for construction manager/general contractor projects: the CMGC learning curve." J. Manage. Eng., 31(4), 04014060.

Hale, D. R., Shrestha, P. P., Gibson, G. E., Jr., and Migliaccio, G. C. (2009). "Empirical comparison of design/build and design/bid/build project delivery methods." J. Constr. Eng. Manage., 135(7), 579–587.

Hyun, C., Cho, K., Koo, K., Hong, T., and Moon, H. (2008). "Effect of delivery methods on design performance in multifamily housing projects." J. Constr. Eng. Manage., 134(7), 468–482.

Ioannou, P. G., and Awwad, R. E. (2010). "Below-average bidding method." J. Constr. Eng. Manage., 136(9), 936–946.

Konchar, M., and Sanvido, V. (1998). "Comparison of U.S. project delivery systems." J. Constr. Eng. Manage., 10.1061/(ASCE)0733-9364(1998)124:6(435), 435–444.

Lines, B. C., and Shalwani, A. S. (2017). "Best value procurement of architectural and engineering services: Selection characteristics and the relative influence of various evaluation criteria." AEI 2017: Resilience of the Integrated Building, ASCE, Oklahoma City, 794–805.

Minchin, R. E., Jr., Li, X., Issa, R. R., and Vargas, G. G. (2013). "Comparison of cost and time performance of design-build and design-bid-build delivery systems in Florida." J. Constr. Eng. Manage., 139(10), 04013007.

Molenaar, K. R., and Songer, A. D. (1998). "Model for public sector design-build project selection." J. Constr. Eng. Manage., 124(6), 467–479.

Nguyen, H. D. P., Lines, C. B., and Tran, D. Q. (2018). "Best-value procurement in design-bid-build construction projects: empirical analysis of selection outcomes." J. Constr. Eng. Manage., 144(10), 04018093.

Perrenoud, A., Lines, B, C., Savicky, J., and Sullivan, K, T. (2017). "Using best-value procurement to measure the impact of initial risk-management capability on qualitative construction performance." J. Manage. Eng., 33(5): 04017019.

Perng, Y. H., Juan, Y. K., and Chien, S. F. (2006). "Exploring the bidding situation for economically most advantageous tender projects using a bidding game." J. Constr. Eng. Manage., 132(10), 1037–1042.

Rosenfeld, Y. (2014). "Root-cause analysis of construction-cost overruns." J. Constr. Eng. Manage., 140(1): 04013039.

Rosner, J. W., Thal, A. E., Jr., and West, J. C. (2009). “Analysis of the design-build delivery method in air force construction projects.” J. Constr. Eng. Manage., 135(8), 710–717.

Sindhu, J., Choi, K., Lavy, S., Rybkowski, Z. K., Bigelow, B. F., and Li, W. (2018). “Effects of front-end planning under fast-tracked project delivery systems for industrial projects.” Int. J. Constr. Educ. Res., 1–16.

Shrestha, P. P., and Fernane, J. D. (2017). “Performance of design-build and design-bid-build projects for public universities.” J. Constr. Eng. Manage., 143(3), 04016101.

Sullivan, J., El Asmar, M., Chalhoub, J., and Obeid, H. (2017). “Two decades of performance comparisons for design-build, construction manager at risk, and design-bid-build: quantitative analysis of the state of knowledge on project cost, schedule, and quality.” J. Constr. Eng. Manage., 143(6): 04017009.

Tran, D. Q., Diraviam, G., and Minchin, R. E., Jr. (2018). “Performance of highway design-bid-build and design-build projects by work types.” J. Constr. Eng. Manage., 144(2), 04017112.

Xia, B., Chan, A., Molenaar, K., and Skitmore, M. (2012a). “Determining the appropriate proportion of owner-provided design in design-build contracts: content analysis approach.” J. Constr. Eng. Manage., 138(9), 1017–1022.

Xia, B., Chan, A., Zou, J., and Molenaar, K. (2013). “Analysis of selection criteria for design-builders through the analysis of requests for proposal.” J. Manage. Eng., 29(1), 19–24.

Xia, B., Skitmore, M., and Zou, J. (2012b). “Evaluation of design-builder qualifications through the analysis of requests for qualifications.” J. Manage. Eng., 28(3), 348–351.

Xia, B., Xiong, B., Skitmore, M., Wu, P., and Hu, F. (2016). “Investigating the impact of project definition clarity on project performance: structural equation modeling study.” J. Manage. Eng., 32(1): 04015022.

Yu, W., Wang, K., and Wang, M. (2013). “Pricing strategy for best value tender.” J. Constr. Eng. Manage., 139(6), 675–684.