

Assessing Low Bid Deviation from Engineer's Estimate in Highway Construction Projects

Minsoo Baek, Ph.D. and Baabak Ashuri, Ph.D.
Georgia Institute of Technology
Atlanta, Georgia

A significant deviation between the agency's estimated construction costs and winning bids submitted for construction projects can result in a loss of public funds, project delay, project cancellation, or inefficient budget allocation. However, few empirical studies are focused on analyzing the low bid deviation. Thus, the objective of this research is to identify factors affecting the low bid deviation from the agency's engineer's estimate. To achieve this objective, this study uses several important variables with the potential to explain the deviation, for example, variables representing project characteristics and market condition factors. Historical cost information for highway construction projects let in the State of Louisiana between 2011 and 2015 are utilized to build the explanatory model. Logit regression analysis is conducted for measuring the effects of influential factors on the low bid deviation. The results of this study indicate that the competition of bidding process, the number of activities in the contract, crude oil price, and value of construction put in place of pavement projects have significant impacts on the low bid deviation. This study contributes to the body of knowledge through the creation of a logit regression model that provides valuable insights into significant factors impacting the low bid deviation. It is anticipated that transportation professionals benefit from the findings of this research to develop more accurate bids and making reliable investment decisions for highway projects.

Key Words: Artificial Intelligence, Big Data, Data Mining, Logit Regression, Low Bid Deviation

Introduction

Low bid deviation is defined as the difference between the agency's estimated construction costs and the lowest bids (or winning bids) submitted by construction contractors for projects. A significant deviation between the agency's estimated construction costs and the winning bid costs can create considerable financial risks for both owners and contractors. For instance, upward deviation leads to inefficient budget allocation of public fund and project cancellation for owners. Downward deviation results in significant cost overrun because of changes and claims during construction for owners and contractors. Cost deviation has been a problematic issue in the construction industry. According to Flyvbjerg et al. (2003), a cost increase has not decreased over the past 70 years. According to the Government Accountability Office (1997), 77% of highway projects in the United States experienced a cost increase. Flyvbjerg et al. (2002) concluded the actual cost for road projects is 20% higher than estimated costs and nine out of ten transportation infrastructure projects experience cost escalation. Cost deviation can cause undesirable consequences such as postponing or canceling scheduled projects, reducing project scope, and losing public faith (Alavi and Tavares 2009). Thus, a proper explanation of cost deviation from the engineer's estimate is essential to prevent the waste of public funds and achieve the best economic result for the public (Carr 2005).

Previous research has been conducted to investigate the issue of a cost increase in construction projects. For example, the study conducted by Hinze et al. (1992) evaluated the submitted bids for 468 construction project in the State of Washington and identified a relationship between the range of bid and cost overrun. The authors concluded that the wide dispersion of bids was likely to lead to cost overrun. Gkritza and Labi studied the likelihood of discrepancies between the highway project final costs and the contract award amount using 1,957 contracted let in the state Indiana. The authors found that project characteristics such as the contract award amount have critical impacts on the cost increase for highway contracts. Love et al. (2013) studied a cost overrun from a contract's award using 276 construction and engineering projects. The authors found that the projects have 12.22% of the average cost overruns and identified the best-fit distributions, including the Cauchy, Wakeby, and four-parameter Burr, to calculate cost overrun probabilities by contract size. Another study conducted by Anastasopoulos et al. (2014) conducted regression analysis to study cost overrun in the transportation project and identify significant factors affecting the cost overrun. The authors concluded that project characteristics such as the number of work activities in the contract, project duration, and length, significantly impact the cost overrun of the projects. The literature revealed that few have focused on the deviation between the agency's estimated cost and the winning bid cost of highway projects. Thus, this study aims to analyze the low bid deviation using highway pavement projects and external factors related to market conditions.

Research Objectives

The main objective of this study was to explain the low bid deviation for highway pavement projects by incorporating external factors. To achieve this main objective, the sub-objectives of this research were to:

1. Develop a logit model to explain the low bid deviation for highway pavement project,
2. Identify factors which are related to project-specific and market conditions, and
3. Analyze the relationship between the low bid deviation and the potential factors

Research Methodology and Data

The cost deviation was measured by the submitted lowest bid to the agency's estimated cost as follows:

$$LD_i = \frac{C_{AEi} - C_{WBi}}{C_{WBi}}$$

where LD is the rate of the low bid deviation of the project i, relative to the corresponding winning bid cost (or lowest bid cost), C_{WB}, of the project, and C_{AE} is the agency's estimated cost for the project. The positive rate of LD indicates a cost increase where the winning bid cost is higher than the agency's estimated costs, while the negative rate of LD is cost decrease where the winning bid cost is less than agency's estimated costs. As the outcome variable is binary, discrete outcome models are possible approaches to estimate the probability of one of these two discrete outcomes (i.e., a cost increase and decrease). Thus, this study considers the binary logit model for analyzing low bid deviation from the agency's estimated cost. An overview of the binary logit modeling process is depicted in Figure 1. First, the study inspects the data to identify outliers and influential observations because they can make specification errors in the model estimation. The significance of the relationships between the low bid deviation and potential factors is measured using correlation analysis. Next, the best set of variables is identified to develop a logit model. Lastly, the results of the developed model are interpreted based on the relationships between the low bid deviation and identified variables.

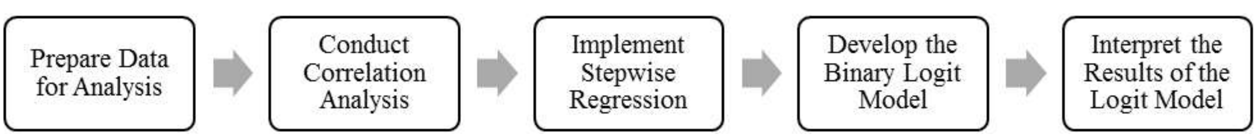

Figure 1

An Overview of the Logit Regression Modeling Process

For logistic regression, the dependent is the probability (P) that resulting outcome indicates the presence of a condition (Washington et al. 2010). The increase of low bid deviation (a cost increase) is coded as 1 in this study. To compare the likelihood of two events, the odds ratio, called the occurrence ratio, is used.

$$Odds\ Ratio = \frac{P_i}{1 - P_i}$$

The logistic regression model can be developed by using the following equation (Washington et al. 2010):

$$Y_i = \log it\ (P_i) = LN\left(\frac{P_i}{1 - P_i}\right) = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \cdots + \beta_k X_{k,i}$$

where β_0 is the model constant and the $\beta_1, \ldots, \beta_k$ are the unknown parameters corresponding with the explanatory variables ($X_k, k = 1, \ldots, k$ the set of independent variables). To estimate the unknown parameters, maximum likelihood methods are used. The estimated parameters are used to estimate the probability that the outcome takes the value 1 as follows (Washington et al. 2010):

$$P_i = \frac{EXP[\beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \cdots + \beta_k X_{k,i}]}{1 + EXP[\beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \cdots + \beta_k X_{k,i}]}$$

In addition, this study uses two model criterion, the -2 log-likelihood function (-2 Log L) and the Akaike information criterion (AIC) to identify the best-fit model. The smaller values of the -2 Log L and AIC indicate the model fits the data better.

Hypothesis Test of the Binary Logit Model:

The significance of the relationship between the low bid deviation and important variables are examined using parameters coefficients and their P-values at 95% significance level in the developed logit model. The following hypothesis is tested in this study.

Null hypothesis (H_0): there is no statistically significant relationship between the low bid deviation and important variables.

Alternative hypothesis (H_1): there is a statistically significant relationship between the low bid deviation and important variables.

Dataset

This study collected cost data of 959 highway pavement projects, which consisted of the agency's estimate cost and the winning bid cost. The highway pavement projects were let in the state of Louisiana between 2011 and 2015. The frequency of low bid deviation for highway pavement projects is described in Table 1.

Table 1

Frequency of Low Bid Deviation for Highway Pavement Projects

Low Bid Deviation	Frequency	Percent	Cumulative Frequency	Cumulative Percent
No Cost Increase (0)	624	65.07	624	65.07
Cost Increase (1)	335	34.93	959	100.00

Transportation agencies often encounter with a significant difference between the engineer's estimates and the low bids. Since the construction costs are dependent not only on project characteristics, but also on market conditions such as construction market, macroeconomic, and energy market conditions, critical knowledge is required to estimate construction costs for highway projects. Thus, this study collected explanatory variables, related the project and market conditions, for developing discrete outcome models. Variables related to project characteristics include the number of bidders, number of activities in the contract, and total contract price. Variables related to market conditions consist of crude oil price West Texas Intermediate (WTI), number of letting projects at the same month in the state of Louisiana, and value of construction put in place of pavement projects in the West South Central regions (i.e., Arkansas, Louisianan, Oklahoma, and Texas). According to Alavi and Tavares (2009), changes in market conditions can cause a significant project cost increase. The descriptive statistics of the input variables are provided in Table 2. As Table 2 suggests, Number of bidders and number of activities in the contract are used as binary variables for their classifications. The remaining variables show significant variations that can represent the nature of the project and market conditions.

Table 2

Descriptive statistics of Input Variables for regression analysis

Variable	Mean	Standard Deviation	Minimum	Maximum
Low Bid Deviation	0.349	0.477	0	1
Number of Bidders (Less than 3)	0.252	0.435	0	1
Number of Bidders (Between 3 and 5)	0.570	0.495	0	1
Number of Bidders (Greater than 5)	0.177	0.382	0	1
Number of Activities in the Contract (Less than 30)	0.233	0.423	0	1
Number of Activities in the Contract (Between 30 and 60)	0.590	0.492	0	1
Number of Activities in the Contract (Greater than 60)	0.177	0.382	0	1
Total Contract Price ($)	2960138.800	6620532.57	32902	146614876
Crude Oil Price WTI ($ per barrel)	87.307	18.808	37.189	109.533
Number of Letting Projects	30.989	11.830	4	63
value of construction put in place of pavement project (Millions of dollars)	4098.060	1492.150	1849	6566

Results and Discussions

The likelihood of low bid deviation for highway pavement projects, resulting in a cost increase, was investigated using the binary logit models. Through a stepwise selection process, this study identified five significant variables, including, number of bidders, number of activities in the contract, and crude oil price WTI, the value of construction put in place of pavement projects. The results of the model fit statistics for the developed binary logit model are presented in Table 3. Based on the smallest AIC (1151.542) and -2 log likelihood (L) (1139.542), the model with the intercept and covariates shows significantly better performance than the model with the intercept only.

Table 3

Model Fit Statistics for Binary Logit Model

Criterion	Intercept Only	Intercept and Covariates
AIC	1242.996	1151.542
-2 Log L	1240.996	1139.542

The results of the model specifications are provided in Table 4. All variables in the model are statistically significant at 95% significance level (i.e., P-value > 0.05). With regard to the relationship between the likelihood of low bid deviation and the identified variables, the findings are insightful. For instance, the number of bidders (less than 3 and between 3 and 5) is identified and positively related to the low bid deviation, while holding the other variables constant in the logit model. This finding indicates that compared to the number of bidders greater than 5, the likelihood of low bid deviation with the number of bidders less than 3 and between 3 and 5 tend to be greater (i.e., higher probability of a cost increase). The number of activities in the contract (greater than 60) has a positive relation with the low bid deviation, given the other variables are held constant in the logit model. It indicates that the likelihood of low bid deviation with number of activities in the contract greater than 60 is higher than the number of activities (less than 30 and between 30 and 60). In addition, crude oil price WTI is negatively related to the low bid deviation, while holding other variables constant in the model. This finding indicates that the increase of crude oil price decreases the likelihood of the low bid deviation between the agency's estimated cost and the winning bid cost. Lastly, the value of construction put in place of pavement project has a positive relation with the likelihood of low bid deviation, given the other variables are held constant in the logit model. This finding shows that the increase in values of pavement construction can increase the likelihood of low bid deviation for highway pavement projects.

Table 4

Model Estimation Results for the Likelihood of Low Bid Deviation for Pavement Projects

Logit Model	Estimate	Standard Error	t Value	P-value
Constant	-1.786	0.220	-8.110	<.0001
Number of Bidders (Less than 3)	1.972	0.253	7.800	<.0001
Number of Bidders (Between 3 and 5)	0.887	0.233	3.810	0.000
Number of Bidders (Greater than 5)	.	.	.	.
Number of Activities in the Contract (Less than 30)	.	.	.	.
Number of Activities in the Contract (Between 30 and 60)	.	.	.	.
Number of Activities in the Contract (Greater than 60)	0.501	0.187	2.680	0.007
Crude Oil Price WTI	-0.177	0.070	-2.520	0.012
Value of Construction Put in Place of Pavement Projects	0.162	0.072	2.230	0.026

Table 5 shows the marginal effects of the identified variables, which measure the changes in the conditional mean of the dependent variable in response to unit changes in the explanatory variables. The second row in Table 5 shows the averages of the 959 individual estimates of the marginal effects for the identified variables. For instance, the pavement projects with the number of bidders (less than 3) are 40.2% more likely to have the low bid deviation, compared to the pavement projects with number of bidders (greater than 5). The pavement projects with the number of bidders (between 3 and 5) are 18.1% more likely to have the low bid deviation, compared to the pavement projects with the number of bidders (greater than 5). Thus, it is found that less competition in the pavement project contracts leads to an increase of low bid deviation.

The pavement projects with the number of activities in the contract greater than 60 are 10.2% more likely to have a cost increase, compared to the pavement projects with the number of activities in the contract less than or equal to 60. As the larger number of activities in the contract indicates a more complex project, it can be concluded that the low bid deviation tends to be greater in the more complex pavement projects.

In addition, one unit increase of crude oil price results in an average 3.6% decrease in the probability that the pavement projects in the state of Louisiana would have the deviation between the agency's estimated cost and the winning bid cost. This finding is contrary to the prevailing understanding of the relationship between material prices and submitted bids by contractors. The finding of this study indicates that an increase in the energy price makes the winning bids for pavement contracts closer to the agency's estimated costs.

Moreover, one unit increase of the value of construction put in place for pavement projects in the West South Central regions results in an average 3.3% increase in the probability that the pavement projects in the state of Louisiana would have the deviation between the estimated cost and the winning bid cost. Thus, an increase in demand for pavement construction in the market leads to an increase of the low bid deviation.

Therefore, this study has the following major implications of the findings:

- Lack of competition in the pavement project contracts can cause a larger deviation between the agency's estimated cost and the winning bid. Thus, the highway agency should evaluate the expected level of competition for their contracts and make an effort to increase the level of competition.
- The larger the number of activities in the contract the higher the low bid deviation. Thus, the highway agency should pay attention to complex projects for developing their engineer estimates.
- The energy market has a significant impact on the low bid deviation. The agency may monitor/track the energy market conditions to develop more reliable construction costs and receive accurate bids from the contractors.
- The boom in construction market leads to an increase in the deviation between the agency's estimate cost and winning bids. The highway agency should evaluate the level of activities in the pavement construction market to avoid significant low bid deviation from the engineer estimate.

Lastly, this study measured the predictability of the developed logit model by comparing the actual frequency of low bid deviation in the data. As shown in Table 6, the predictability of the developed logit model is 35% that is exactly the same as the number of the frequency of 35% having low bid deviation between the agency's estimated costs and the winning bid costs.

Table 5

Marginal Effect of the Identified Variables

Variable	Mean	Standard Deviation
Number of Bidders (Less than 3)	0.402	0.085
Number of Bidders (Between 3 and 5)	0.181	0.038
Number of Activities in the Contract (Greater than 60)	0.102	0.022
Crude Oil Price WTI	-0.036	0.008
Value of Construction Put in Place of Pavement Projects	0.033	0.007

Table 6

Predictability of the Developed Logit Model

Label	N	Mean	Standard Deviation	Minimum	Maximum
Actual Frequency of Low Bid Deviation	959.000	0.349	0.477	0.000	1.000
Estimated Probability (Logit Model)	959.000	0.349	0.153	0.103	0.798

Conclusions

The primary objective of this study is to develop an explanatory model that explain low bid deviation from the engineer's estimate, collected in the State of Louisiana between 2011 and 2015. This study analyzed engineer's estimates and the winning bids submitted for 959 highway pavement contracts. This paper utilized the binary logit regression to explain the low bid deviation and identified several important variables, including the number of bidders, number of activities in the contracts, crude oil price, and value of construction put in place of pavement projects. The model fit statistics (i.e., AIC and -2 Log L) indicated that the developed logit model is significantly better than the model with intercept only.

This study contributes to the body of knowledge through the examination of influential variables on the low bid deviation in highway pavement projects. For example, the number of bidders (less than 3 and between 3 and 5), number of activities in the contract (greater than 60), and value of construction put in place of pavement projects are positively related to the low bid deviation and the crude oil price WTI has a negative relation with the low bid deviation. The findings indicate that the lower the degree of competition in the bidding process for pavement projects is the higher the low bid deviation between the agency's estimated costs and the winning bids is. In addition, the larger the number of activities in the pavement contracts is the lower the low bid deviation from the agency's estimated cost is.

This study also found that the oil price and construction market condition have profound impacts on the low bid deviations for pavement projects. More specifically, the higher crude oil price is an indication of declining the deviation between the agency's estimated costs and the winning bid costs. The increased demand in the local construction market leads to the increased deviation between the estimated cost and the winning bid costs for pavement projects. The findings of this study help transportation agencies in estimating or adjusting highway construction costs. For instance, the agencies can monitor or take into account the identified variables (level of competition, oil market, complexity of a project, and construction market conditions) to adjust their estimated cost for managing a cost increase resulting from market changes.

Reference

Alavi, S., and Tavares, M. P. (2009). Highway Project Cost Estimating and Management. Helena, MT.

Anastasopoulos, P. C., Haddock, J. E., & Peeta, S. (2014). Cost overrun in public-private partnerships: Toward sustainable highway maintenance and rehabilitation. Journal of Construction Engineering and Management, 140(6), 04014018.

Anderson, S., Molenaar, K., and Schexnayder, C. (2006). *Procedures for Estimation and Management for Highway Projects During Planning, Programming, and Preconstruction*. College Station, Texas.

Carr, P. G. (2005). Investigation of Bid Price Competition Measured through Prebid Project Estimates, Actual Bid Prices, and Number of Bidders. *Journal of Construction Engineering & Management*, *131*(November), 1165–1172.

Flyvbjerg, B., Holm, M. S., and Buhl, S. (2002). Cost Underestimation in Public Works Projects: Error or Lie? Journal of the American Planning Association.

Flyvbjerg, B., Skamris Holm, M. K., and Buhl, S. L. (2003). How common and how large are cost overruns in transport infrastructure projects?. Transport reviews, 23(1), 71-88.

Gkritza, K., and Labi, S. (2008). Estimating cost discrepancies in highway contracts: Multistep econometric approach. Journal of Construction Engineering and Management, 134(12), 953-962.

Government Accountability Office. (1997). "Managing the Costs of Large-Dollar Highway Projects." (February).

Hinze, J., Selstead, G., and Mahoney, J. P. (1992). Cost overruns on the state of Washington construction contracts. Transportation Research Record, 1351, 87.

Kaliba, C., Muya, M., and Mumba, K. (2009). Cost escalation and schedule delays in road construction projects in Zambia. International journal of project management, 27(5), 522-531.

Love, P. E., Wang, X., Sing, C. P., and Tiong, R. L. (2012). Determining the probability of project cost overruns. Journal of Construction Engineering and Management, 139(3), 321-330.

Washington, S. P., Karlaftis, M. G., and Mannering, F. (2010). Statistical and econometric methods for transportation data analysis. Chapman and Hall/CRC.