UNLV | University Libraries
University of Nevada, Las Vegas

UNLV Theses, Dissertations, Professional Papers, and Capstones

December 2017

Effects of Contract Procurement Factors on Performance of Transportation Projects

Ruiko Maharjan
University of Nevada, Las Vegas, ruikomaharjan@gmail.com

Follow this and additional works at: https://digitalscholarship.unlv.edu/thesesdissertations

Part of the Civil Engineering Commons

Repository Citation
Maharjan, Ruiko, "Effects of Contract Procurement Factors on Performance of Transportation Projects" (2017). *UNLV Theses, Dissertations, Professional Papers, and Capstones*. 3148.
https://digitalscholarship.unlv.edu/thesesdissertations/3148

This Dissertation is brought to you for free and open access by Digital Scholarship@UNLV. It has been accepted for inclusion in UNLV Theses, Dissertations, Professional Papers, and Capstones by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact digitalscholarship@unlv.edu.

المنارة للاستشارات

www.manaraa.com

EFFECTS OF CONTRACT PROCUREMENT FACTORS ON PERFORMANCE OF TRANSPORTATION PROJECTS

By

Ruiko Maharjan

Bachelor of Engineering - Civil Engineering
Tribhuvan University, Nepal
2011

Master of Science in Engineering - Civil and Environmental Engineering
University of Nevada, Las Vegas
2013

A dissertation submitted in partial fulfillment
of the requirements for the

Doctor of Philosophy - Civil and Environmental Engineering

Department of Civil and Environmental Engineering and Construction
Howard R. Hughes College of Engineering
The Graduate College

University of Nevada, Las Vegas
December 2017

www.manaraa.com

Copyright 2017 by Ruiko Maharjan

All Rights Reserved

www.manaraa.com

Dissertation Approval

The Graduate College
The University of Nevada, Las Vegas

September 12, 2017

This dissertation prepared by

Ruiko Maharjan

entitled

Effects of Contract Procurement Factors on Performance of Transportation Projects

is approved in partial fulfillment of the requirements for the degree of

Doctor of Philosophy – Civil and Environmental Engineering
Department of Civil and Environmental Engineering and Construction

Pramen Shrestha, Ph.D.
Examination Committee Chair

Hualiang Teng, Ph.D.
Examination Committee Member

Mohamed Kaseko, Ph.D.
Examination Committee Member

Jin Ouk Choi, Ph.D.
Examination Committee Member

Ashok K. Singh, Ph.D.
Graduate College Faculty Representative

Kathryn Hausbeck Korgan, Ph.D.
Graduate College Interim Dean

www.manaraa.com

ABSTRACT

Effects of Contract Procurement Factors on Performance of Transportation Projects

by Ruiko Maharjan

Cost and schedule savings are the main measures of a project's success. Several factors affect the cost and schedule performances in a construction project, such as design changes, material, labor and equipment shortages, unpredictable weather conditions, and errors & omissions in contract documents. Some studies have shown that either the construction cost or the schedule performance of a project was dependent on the procurement factors, namely: bid cost, number of bidders, the bid cost deviation between the first and second bidder, the liquidated damage rate per day, the type of a contract, and the project location. However, a comprehensive study on the combined effect of procurement factors on performance metrics has not been yet conducted. Therefore, this study collected all the available contract procurement factors to determine the combined effect of these factors on the construction cost and the schedule performances. In addition, the multiple linear regression models within the study were developed to predict the performance metrics based on these factors.

For this study, the project data completed between the year 2000 and 2016 were collected from two state department of transportations (DOTs): Texas and Florida. The results showed that not only cost growth but also schedule growth had a significant correlation between the liquidated damage rate per day, the type of a contract funding, the type of a contractor, and the location of a project. The validation process showed that the models developed during this study could predict project performance metrics accurately. Further research is recommended with more state DOTs data to check whether the relationships between the procurement factors and project performance metrics are similar to those found in this study.

www.manaraa.com

ACKNOWLEDGEMENTS

I am grateful to my supervisor, Dr. Pramen P. Shrestha, whose expertise and generous guidance made it possible to complete my doctorate degree. It was a pleasure working with him.

I would like to extend my sincere gratitude and appreciation to my dissertation committee members, Dr. Mohamed Kaseko, Dr. Ashok Singh, Dr. Jin Ouk Choi, and Dr. Hualiang Teng for their constructive and valuable suggestions to finalize my work.

I would also like to express my deep appreciation and gratitude to my lovely husband, Sanjib Mulepati, for being patient, motivative, and supportive throughout this journey. Without his guidance and persistent help, this dissertation would not have been possible. Last but not the least, I would like to thank all my family: my parents and to my brothers and sister for supporting me emotionally throughout my life in general.

www.manaraa.com

TABLE OF CONTENTS

المنارة للاستشارات

www.manaraa.com

المنارة للاستشارات

www.manaraa.com

المنارة للاستشارات

www.manaraa.com

المنارة للاستشارات
www.manaraa.com

LIST OF TABLES

المنارة للاستشارات

www.manaraa.com

www.manaraa.com

www.manaraa.com

LIST OF FIGURES

المنارة للاستشارات

www.manaraa.com

LIST OF ABBREVIATIONS

CG	Cost Growth
DOT	Department of Transportation
LD per Day	Liquidated Damage Rate per Day
MLR	Multiple Linear Regression
SG	Schedule Growth
UNLV	University of Nevada, Las Vegas
U.S.	United States of America
VIF	Variation Inflation Factor

www.manaraa.com

CHAPTER 1 INTRODUCTION

1.1 Background

Cost and time overruns in a project are widespread problems in the construction industry. Memon et al. (2012) found 89% and 92% of their survey respondents faced cost and time overruns respectively in their construction projects. These findings were based on 140 clients, consultants, and contractors involved in construction projects in Malaysia. Additionally, Harbuck (2004) mentions that an average contract cost overrun in a project depends on its geographical condition in the United States. The author found that the average cost overrun on a project in Texas and Florida state is 4.9% and 9.5% respectively. Of the several performance indicators, as seen in Table 1, the most important indicators are cost and time to define a project's success. Other frequent measures are quality, safety, and owner satisfaction, which play a vital role in determining the success of a completed project.

Table 1 Performance Measurement Framework used in Construction Industry

Sources	Framework	Performance Criteria
Chan and Chan (2004)	KPI	Cost, time, quality, health and safety, participant's satisfaction, user expectation, environmental performance, commercial value
Kagioglou et al. (2001)	Modified BSC	Financial, customer, internal business, project, supplier leadership
Yu et al. (2007)	BSC and KPI	Profitability, financial growth, financial stability, external/ internal customer satisfaction, market share, research and development, technological capability, business efficiency, human resource development, organization competency, access to information

Note: KPI =Key Performance Indicator BSC= Balanced Score Card

Numerous factors affect cost and time performance indicators. The factors that affect final cost are project scope, project delivery partners, operating region, project duration, and initial estimated cost (Ahiaga-Dagbui & Smith, 2014). Some uncertainties such as inattention to risks within the government, biases in decision-making in the evolution and use of information, and uncertainty in project management

www.manaraa.com

and administration also cause cost overruns (Jennings, 2012). Similarly, factors such as improper planning, lack of effective communication, and design error affect the overall schedule of the project (Tumi et al., 2009). The factors affecting cost and schedule in a project lead to claims and litigation between clients and contractors. Owner's often perceived claims as arising from the nature of the task performed and people’s deliberate practice, whereas contractor often perceives latent conditions as the major reason for claims and litigation in a project (Love et al., 2010). Before going to litigation, both parties frequently adopt the resolution techniques such as negotiation, arbitration, and mediation to solve the disputes between main parties in construction projects.

The construction of the project delivered through the traditional project delivery method has two stages after detailed design development: Procurement and Construction. In the first phase, there is the availability of various procurement data such as a number of bidders and their bid amounts, liquidated damage rate per day, the type of contract funding, the type of contractor, and the location of a project. In addition, the completion of the construction phase of a project provides final cost and schedule to complete a project. Table 2 shows some of the studies that determined the relationship of cost performance with procurement data. The past studies mainly focused on individual relationships between several factors and the cost growth. Some authors found a relationship between the cost growth and factors such as the size of a project, the particular contractor, the type of work, and the number of items.

www.manaraa.com

Table 2 Studies on Relationship between Cost Performance Metrics and Procurement Data

Cost Performance Metrics	Authors	Parameters	Findings
Construction Cost Growth	Jahren and Ashe (1990)	Size of project Award-Estimate Difference	- Found a significant influence on construction cost growth by size of the project when percentages were calculated for each group size.
Construction Cost Growth	Randolph et al. (1987)	Contractor Type of work # Unit cost items in bid Year the project is bid Job classification Funding availability Project designer Project size	- Higher change in project cost is found for smaller projects - Found significant correlation between construction cost growth and bid data information such as the type of work, particular contractor, and a number of items.
Final Project Cost	Williams (2003)	Low bid cost	- Rate of increase in final cost is higher for larger highway projects than small projects. - More than 50% of highway projects from New York, Texas, and New Jersey had predicted values within 10%
Total Cost Growth	Ahiaga-Dagbui and Smith (2014)	Purpose of project Project scope Project delivery partners Operating region Project duration Initial estimated cost	- Created an artificial neural network to predict final cost of project and found 87% of model predictions are within a range of ± 5%

Table 3 shows some studies findings related to schedule performance metrics. The researchers also investigated the relationship between schedule growth with construction cost and total gross floor area. The researchers found that there is a correlation between duration and both cost and floor area in building projects.

المنارة للاستشارات
www.manaraa.com

Table 3 Studies on Relationship between Schedule Performance Metrics and Procurement Data

Authors	Independent Variables	Regression function
Martin et al. (2006)	Construction Cost (C)	(D = a ± b C) * Duration increase linearly along with increase in the cost of a project for all building types except for catering buildings
Chan and Kumaraswamy (1995)	Construction Cost(C) Total Gross Floor Area(A)	(Log(D) = log K + B log C) ** (Log(D) = log K + B log A) ** Found a significant relationship between time and cost & time and floor area

Note: *a = y-intercept, b = slope of line, and D = construction duration
**K = constant describing general level of time performance for a project, B = constant indicating the effect of cost on time performance, and D = construction duration

The past studies suggest that the procurement data of a project has a correlation with the data available during the construction phase of a project. Therefore, the performance of the project can be predicted during the procurement phase of the project, if the procurement data can be used to build the regression models for the cost and schedule metrics. This study focused on collected comprehensive procurement data of transportation projects to build the multiple linear regression models.

1.2 Scope and Objectives

The scope of this project is limited to building multiple linear regression models for cost and schedule performance metrics based on the procurement data of a project. In this study, seven contract procurement data were collected to build two main project performance models. The objectives of this research study are as under:

i. Find the relationship between various contract procurement data and project performance metrics among TxDOT and FDOT projects.

ii. Develop multiple linear regression models for each of the performance metric based on available contract procurement data.

iii. Compare the difference between TxDOT, FDOT, and Combined DOT model findings.

iv. Validate the developed multiple linear regression models using independent project data.

المنارة للاستشارات
www.manaraa.com

1.3 Practical Implications

The model developed can be converted to a desktop package for predicting quick cost and schedule variation from data that are available during procurement phase of a project. This will help to find alternative solutions ahead if there is a cost overrun and schedule growth for a construction project. In addition, the model will help in reducing time and resources spend during the estimation process.

1.4 Limitation of the Study

i. The cost and schedule performance metrics are only included. The other success factors such as quality, safety, owner satisfaction, and so on are excluded from measuring project success.

ii. Only included transportation construction projects such as freeway, non-freeway, and bridge construction and upgrade.

iii. The projects such as maintenance and safety are excluded in the study.

iv. The transportation projects were completed in between the year 2000 and 2016.

www.manaraa.com

CHAPTER 2 LITERATURE REVIEW

The review of past literature is grouped into four sections. The first section covers the literature related to cost performance metrics. The second section explores the literature about schedule performance matrix. Then, the last section summarizes the gaps in literature reviews and need for this study.

2.1 Cost Performance Metrics

In the traditional method, the cost is the sole criterion for contractor selection used by owners, so it is the key factor to be considered. Under design bid build, the construction estimates are first prepared by owner's estimator and then bids are invited from the contractors in which lowest bidder wins the contract. The accuracy of cost estimation plays a vital role during project execution phase. If the cost estimated is not done correctly, then the contractor bid amount will deviate significantly from the estimated cost. Similarly, if the contractor bid estimate is not prepared accurately the completion cost of the project could be impacted. Table 4 shows some of the projects that have significantly gone over budget.

Table 4 List of Transportation Projects with Cost Overruns (Source: Edwards & Kaeding, 2015)

Transportation Projects	Cost Estimates and Date of Estimate		% Construction
	Original	Final	Cost Growth
Boston Big Dig	$2.6b (1985)	$14.6b (2005)	462
New York City East Side Access	$4.3b (1999)	$10.8b (2014)	151
San Francisco-Oakland Bay Bridge	$1.4b (1996)	$6.3b (2013)	350
Denver International Airport	$1.7b (1989)	$4.8b (1995)	182
New York City WTC Rail Station	$2.0b (2004)	$4.0b (2015)	100
Denver West Light Rail	$250m (1997)	$707m (2013)	183
Virginia Springfield interchange	$241m (1994)	$676m (2003)	180

Note: m = million, b = billion

Memon et al. (2012) state that 89% of construction project in Malaysia are facing cost overrun. This finding is based on 140 responses received from questionnaire survey among client, consultant, and

www.manaraa.com

contractor. Besides the accurate cost estimation, the final cost is affected by several factors (Table 4).

Table 5 Top Three Factors Influencing Cost Overrun

Authors	Major Factors of Cost Overrun	Respondents Type	Project Location
Memon et al. (2011)	Poor design & delays in design Unrealistic contract duration Lack of experience	15 (client, consultant, and contractor)	Malaysia
Kaming et al. (1997)	Increase in material cost Inaccurate material estimating Project complexity	31 project managers	Indonesia
Mansfield et al. (1994)	Price fluctuations Inaccurate estimates Delays	50 (contractor, consultant, and client organization)	Nigeria

Meeampol and Ogunlan (2006) write that the management of construction resources and budget, construction method, and communication are the key factors reducing cost overrun of the construction project. These findings were based on literature reviews and opinions of experienced engineers in Thailand. In addition, there are numerous factors that have a direct impact on the decision to bid any project. Some of the research conducted to determine the factors affecting the bidding decision are shown in Table 6. Carr (1983) found that a rational contractor will lower the markup when the number of bidder increases in a project. However, the increase in a number of bidders may impact the chances of getting a fair bid amount in a project. The authors developed a regression model that predicts the minimum number of bidders required to be in the competition for an owner to get a reasonable bid price (Ngai et al., 2002). The authors assumed a rate of change of Tender Price Index (TPI) as the main input data for the model because it is an indirect measure of the Hong Kong market conditions. This analysis is based on a sample of 229 Hong Kong building projects with 3,285 bids received in between the year 1990 to 1996.

المنارة للاستشارات
www.manaraa.com

Table 6 Top Three Factors Affecting Bidding Decision

Authors	Factors affecting Bidding Decision	Respondents Type	Project Location
Leśniak and Plebankiewicz (2013)	Type of work Experience in similar projects Contractual terms	62 contractors	Southern Poland
Bageis and Fortune (2009)	Contractor size Classification status of contractor Main client type	91 construction and maintenance contractors	Saudi Arabia
Chua (2000)	Competition Risk Company's position in bidding	153 top contractors	Singapore
Shash (1993)	Need for work Number of competitors tendering Amount of experience	300 top general contractors	United Kingdom
Ahmad and Minkarah (1988)	Type of job Need for work Owner	400 top general contractors	United States

The past studies determined the correlation between the number of bidder and deviation between bids received for a project. Li et al. (2008) found that as the number of bidders increased, the log of the percentage difference between the second-lowest and the lowest bid decreases. These results are based on a study of 927 building projects in Utah. The authors also found that by timing the projects to seasonal periods of construction slowdown, the effect of a number of bidders on the difference between the second-lowest and lowest bid can be reduced. In addition, Runeson and Skitmore (1999) also found that with the increase in the number of bidders average ratio of the lowest and second lowest bids reduces. The average difference between the ratios also reduces with increasing number of bidders. In addition, Carr (2005) found that increase in the number of bidders will result in a decrease in the bid prices in building projects with an awarded cost between 73 thousand and 14 million dollars. The decrease in bid price will lead to decrease in the gap between the engineer's estimate and the contract bid amount. In addition, it will reduce the gap between the bidders' bid. The study showed that the linear and

المنارة للاستشارات
www.manaraa.com

curvilinear correlation coefficient between a number of bidder and contract award cost growth was - 0.32 and - 0.62 respectively. The findings of this study are based on 19 major public work building projects in New York. The author also suggested the need for conducting future research to determine the impact of the number of bidders on the construction cost growth.

Several studies have been conducted to predict final construction cost of a project. Shrestha and Pradhananga (2010) conducted a study to determine the effect of competitive bidding on cost performance of the projects. The data for the study was based on 435 bids received on 113 street construction projects built by Clark County Department of Public Works (CCDPW) from 1991 to 2008 in Nevada. The major finding of the study is that for the larger projects (>5 million) there is a significant relationship between a number of bidder and contract award cost growth. This concludes that the bid price decreased with the increase in the number of bidders. Also, there is a significant relationship found between the lowest bid price and the final construction cost. Therefore, the authors also developed a regression model to predict the final construction cost by using the lowest bid price. The study, however, did not compare the effect of competitive bidding on schedule performance of the projects. Moreover, (Williams, 2003) found that the rate of increase in final cost is higher for large highway projects than small projects. This finding is based on linear regression between natural log transformation of low bid and final cost. More than fifty percent of highway projects from New York, Texas, and New Jersey had predicted values within ten percent of the actual completed cost.

In a study conducted on public works projects from year 1977 to 1985 in Michigan, Randolph et al. (1987) considered seven independent variables: bid year, type of work, number of unit cost items in bid, funding source of project, job classification, contractor code, and project designer for analysis of each of these variable effect on construction cost growth. The authors found significant high correlation among a number of the contractor, type of work, and a number of unit cost items in the bid. The amount of

المنارة للاستشارات

www.manaraa.com

funding available per year also had an influence on the rate of increase in project cost. The project type used for investigation ranged from the storm and sanitary sewer construction to sidewalk repair to road and bridge construction. Further, the comparison of projects on three group sizes: less than 50 thousand, between 50 and 250 thousand, and more than 250 thousand were done. It was found that the construction cost growth decreased as the size of the project increased.

In addition, Jahren and Ashe (1990) found that the size of the project and the difference between the low bid and engineer's estimate (in percentage) had an influence on construction cost growth. They found that the median construction cost growth increases as the project size increases. This finding was based on 1,576 construction projects that were administered by Naval facilities engineering command in California. In addition, the authors also found that the risk of high construction cost overrun rates is greater when the award amount is less than the government estimate. This is due to the competition during bidding phase of a project which compels contractors to lower their contingencies to win a bid and then making up the difference through changes and claims during construction of the project.

Based on the data of 1,600 water infrastructure and utility projects completed from 2004 through 2012, the authors created a model using the artificial neural network to predict cost overrun (Ahiaga-Dagbui & Smith, 2014). The projects included in the model had cost range of between £4,000 and £15 million and were newly built, upgraded, repaired or refurbished. An absolute percentage error achieved was 3.67 and 87% of the testing data predicted the values within a range of ± 5 percent. The authors have considered eight predictors for a final cost of a project: project purpose, project scope, project delivery partners, operating region, project duration, and initial estimated cost.

Akinci and Fischer (1998) described the specific factors related to construction, economic and political environment are the uncontrollable factors affecting the final cost of a project. The construction specific factors include unknown geological conditions, weather conditions, and client and subcontractor-

المنارة للاستشارات
www.manaraa.com

generated risk factors. The economic and political factors include economic and political risk factors. Besides this, the authors found the contract specific factors such as the type of contract and legal context of a contract such as liquidated damage rate per day clause increase the risk to the contractor in executing the project.

Shrestha et. al (2012) conducted an individual correlation between 21 input variables and two output variables using Pearson and Spearman's correlation tests. The 21 input variables are nature of construction, number of interchanges, bridge area, pavement type, liquidated damage rate per day, schedule performance bonus, specification type, partnering, environmental assessment, ROW assessment, value engineering, design hours per week, number of construction shifts, working days lost, responsibility of ROW procurement, number of ROW parcels, ROW procured by eminent domain, and ROW delay. The output variables are cost growth and schedule growth. Out of 21 variables, only working days lost have a positive significant correlation with cost growth. In addition, there is a positive correlation, though not significant, found between cost growth and liquidated damage rate per day. The study comprised a total of 22 large highway projects (>50 million) completed by using either design-build or design-bid-build project delivery method in TxDOT.

The construction cost of the highway project across U.S. states can differ for a number of reasons. In a study, conducted by (WSDOT, 2002) on a 1.2-mile long interchange project among 25 members of American Association of State Highway and Transportation Officials (AASHTO) from different U.S. states, the estimated project cost ranged from $4 million to $26.7 million. (LTAP, 2005) developed a regression model to estimate the cost of replacing a bridge based on bridge deck area in Indiana county. The model showed that the cost of constructing a bridge project with federal fund was higher than locally funded. A total of 377 bridges from 46 counties in Indiana completed during fiscal years 1997 to 2005 were reviewed for this study.

المنارة للاستشارات
www.manaraa.com

(Kishore & Abraham, 2009) conducted the study on the effect of contract funding type on the construction cost of highway projects and found that federally funded project cost higher in comparison to state-funded project. The reason for higher cost in federal projects was due to federal requirements such as meeting Disadvantaged Business Enterprise (DBE) goals, following standard specification for design and materials, environmental compliance documentation, etc. Based on the interviews of Indiana contractors, the level of risk involved is lower for locally funded jobs because of its lower investment at stake, decision-making being faster due to quick access to projects resulting in expedient decision-making by county highway personnel, the permitted use of economic and readily available materials, and lesser material testing requirements.

(Hinze & Selestead, 1991) found the cost overrun of a project was impacted by an individual variable such as a number of bidders, project size, districts, project type, precontract engineering effort, construction engineering effort, the frequency of awarding WSDOT contracts to a contractor. This study is based on the information of 433 highway projects constructed between 1985 and 1989 in Washington state. The analysis of the data showed that as the number of bidders increased, the cost overrun rate also increased. In addition, the median cost overrun rate for projects greater than $2.5 million was greater than projects with $0.25 million median cost overrun rate. Moreover, the median cost overrun rate was higher in district 4 in comparison to other five districts considered in the study. The author found that for increase in both precontract engineering and construction engineering cost, the project becomes more complex increasing the cost overrun rate of a project. Bordat et al. (2004) found that the percentage of cost overruns of a project were different based on the different state in U.S. as shown in Table 7.

المنارة للاستشارات
www.manaraa.com

Table 7 Percentage of Cost Overrun in Different States of U.S.

State	Period	Cost Overruns
Idaho	1997 to 2001	55% - 67%
Indiana	1996 to 2002	55%
Missouri	1999 to 2002	60% - 64%
New Mexico	2002	62%
Ohio	1994 to 2001	80% - 92%
Oregon	1998 to 2002	18% - 33%
Tennessee	1998 to 2002	61%
Texas	1998 to 2002	66% - 75%

There is various effective cost estimating methods that could be applied to predict the final cost of the project. Traditionally, the estimation of the construction cost of any project is done based on case-based reasoning (CBR) of retrieved cases. However, it is very difficult to find a similar type of projects to predict the cost of the new project. So, Kim et al. (2012) developed a revised case-based reasoning model based on the regression analysis model for construction cost estimation in the early phase of the railroad-bridge construction projects in South Korea. Kim (2011) created a cost estimation model for irrigation-type river facility construction based on case-based reasoning with genetic algorithms. This model was created based on weights of different seven attributes: embankment extension, revetment extension, freeboard, number of drain gates, number of drainpipes, and slope covering material of 92 historical cases of irrigation-type river facility construction.

Moreover, in another paper, the authors have found that parametric cost estimation model based on multiple regression equation is an effective tool for preparing accurate strategic and conceptual cost estimates of apartment building projects in Korea. Ji et al. (2010) explain that the preprocessing of historical data adds for accurate outputs in this method. The data processing mainly includes principal component analysis and correlation analysis. Besides the above-mentioned parameters affecting cost and schedule growth. There is also a special investigation required in assessment on bid documents. The evaluation of bid documents is based on the degree of response to bid document, construction

المنارة للاستشارات
www.manaraa.com

organization design, and firm honor and competence in Chinese construction market (Lai et.al, 2004). Also, Carr (1987) writes that competitive bid analyses include resource restraints and opportunity cost for a better understanding of a firm's competitive position. Moreover, Ho and Hsu (2013) write that bid compensation during bid preparation process for project planning and schematic design motivates bidders to get best bid amount mainly in the case of heterogeneous bidders. The bid preparation process involves a lot of effort for bidders in terms of time and cost. Moreover, Yu et al. (2012) discussed that besides going only for lowest bid the owner should consider best value approach which provides more qualitative results for a project. Salem Hiyassat (2001) has proposed an alternative statistical procedure to award a bid based on an average bid price rather than the lowest bid price to control final cost growth of a project.

Kim et al. (2004) conducted a comparison of construction cost estimating models developed based on regression analysis, neural networks, and case-based reasoning. The cost estimating models were developed using input variables such as year, gross floor area, number of stories, total unit, duration, roof types, foundation types, usage of the basement, and finish grades of 490 residential buildings that were built by general contractors between 1997 and 2000 in Korea. Among the three models developed, the best model was found to be neural network model based on a mean absolute error between predicted and actual cost. Later, An et al. (2007) conducted a comparison of three case-based reasoning cost estimating model (assumed equal weights, gradient descent method, and analytical hierarchy process) on 540 residential buildings built by general contractors in Korea between 1997 and 2002. The cost prediction model consists of nine variables: gross floor area, number of stories, total unit, unit area, location, roof types, foundation types, usage of the basement, and finish grades that were selected from the interviews of engineers working in construction companies. The analytical hierarchy process was found better in comparison to other two case-based reasoning cost estimation model based on the mean absolute error

المنارة للاستشارات
www.manaraa.com

between predicted and actual cost.

(Chou, 2009) used a parametric estimating technique to develop a generalized linear model for predicting quantities of a total of 10 geometry related activities of a transportation project based on each project's characteristics. The geometry related activities were preparing right of way, excavation, embankment, flexible base, lime treatment for materials used as subgrade, lime treatment for base courses, surface treatments, planning and texturing pavement, concrete pavement, and work zone pavement markings. The project characteristics used for model development were engineering quantity, average daily traffic, designed speed, average lane width, lane number, number of trucks on the existing highway, the percentage of trucks, road length, project type, shoulder width, terrain type, the highway system, trunk system, and project location.

2.2 Schedule Performance Metrics

Construction duration is a time required to complete a project. It is also defined as the elapsed period from commencement of site works to the completion of the project. The experience with planning and calculation techniques are mainly used to forecast the duration of the project. The desired timescale is given in the initial phase of the project construction.

Based on findings of Memon et al. (2012), 92% of the respondents involving in construction projects are facing time overrun in Malaysia. The different perceptions were found among the owner, contractor, and consultant on reasons for delays in the project (Assaf & Al-Hejji, 2006). Some of the examples of these factors that were considered in the study are a shortage of labors, delay in progress payments by owner, inflexibility of a consultant, and delay in producing design documents. Table 8 shows some other factors illustrated in the past studies resulting schedule overrun in a project.

المنارة للاستشارات
www.manaraa.com

Table 8 Top Three Factors Influencing Schedule Overrun

Authors	Major Factors of Schedule Overrun	Respondents Type	Project Location
Tumi et al. (2009)	Improper planning Lack of effective communication Design errors	Literature review and questionnaire survey among contractors, clients, and consultants	Libya
Iyer and Jha (2006)	Lack of coordination among project participants Project manager's ignorance Lack of knowledge	112 industry professionals	India
Kaming et al. (1997)	Design changes Poor labor productivity Inadequate planning	31 project managers working on high rise projects	Indonesia
Nkado (1995)	Specified sequence of completion Programming construction work Form of construction	29 members of national contractors group	UK

In addition, Frimpong et al. (2003) found factors such as monthly payment difficulties from agencies, poor contractor management, and material procurement lead to both schedule and cost overruns in a project. These findings are based on questionnaire survey conducted among 72 respondents, personnel from owners, consultants, and contractors involved in groundwater projects in Ghana. Moreover, Meeampol and Ogunlan (2006) stated that the improvements in construction method, construction resource management, schedule management, human resource management, supervision and control, and communication system would help to reduce schedule overrun of a construction project. Moreover, Chua and Hossain (2012) developed a model for predicting the impact of external changes on design schedule. These external changes are requested by client's due to unforeseen events, change initiated by construction methods or field conditions, change initiated by fabricator or supplier. The unforeseen events include incorrect assumption about market conditions, future customer needs and available technology. The authors found that the project size has no effect on schedule growth.

Martin et al. (2006) determined relationships between construction time and construction cost

المنارة للاستشارات
www.manaraa.com

for various types of buildings. For swimming pools and libraries, cost growth was significantly correlated to construction time growth, whereas for factories, warehouse, and hotel or motel the construction cost has less impact on construction time. The findings were based on data from 2,700 building projects collected using questionnaire survey among construction industry clients and consultants. These building types were divided into 29 categories and were completed in the UK between 1998 and 2006. Furthermore, Chan and Kumaraswamy (1995) developed a linear model to illustrate construction duration with construction cost based on three project categories: public sector building, civil engineering works, and private sector buildings. In addition, the authors also determined the relationship of duration with total gross floor area based on three project categories. The authors found that private sector buildings behaved differently to in both models in comparison to other two types. The authors have also found site productivity levels has an impact on construction duration.

Shrestha et. al (2012) found five out of 21 input variables had a significant correlation with schedule growth. There is a positive significant individual correlation found between schedule growth and four input variables: number of interchanges, pavement type, schedule performance bonus, and partnering, whereas, negative significant correlation with construction work days per week. Moreover, there is a positive correlation, though not significant, found between schedule growth and liquidated damage rate per day. Bordat et al. (2004) found that the percentage of time overruns of a project were different based on the different state in U.S. as shown in Table 9.

المنارة للاستشارات
www.manaraa.com

Table 9 Percentage of Time Overrun in Different States of U.S.

State	Period	Time Overruns
Indiana	1996 to 2002	12%
New Mexico	2002	10%
Ohio	1994 to 2001	44% - 56%
Oregon	1998 to 2002	15% - 65%
Tennessee	1998 to 2002	14%
Texas	1998 to 2002	52% - 55%

2.3 Gaps in the Literature

The review of literature had shown that few studies had been conducted to investigate the relationship between the performance metrics and the procurement data. Some of the procurement data has been used in these studies, but none of the studies had investigated the combined effect of available procurement data on project performance in transportation projects. Some of the researchers had conducted individual correlation of project performance with project bid data but has not performed combined regression analysis. Therefore, the intent of this study is to determine the combined effect of these procurement data with the project performance data of the road projects completed by state DOTs. This will help to identify the procurement metrics that have a significant correlation with the project performance metrics.

المنارة للاستشارات
www.manaraa.com

CHAPTER 3 RESEARCH METHODOLOGY

3.1 Overview of Research Methodology

The overview of the methodology used in this study is shown in Figure 1. The objectives of the study are identified based on the gaps in the literature review. Then the project performance metrics are developed to determine the impact of contract procurement data on these metrics. Research hypotheses were created then these are converted to null hypotheses to conduct the statistical tests. Then, the multiple linear regressions models were developed for each performance metrics based on procurement data. These models were validated with the help of testing data. At the end, the conclusions of the findings and the recommendations for further studies are presented.

www.manaraa.com

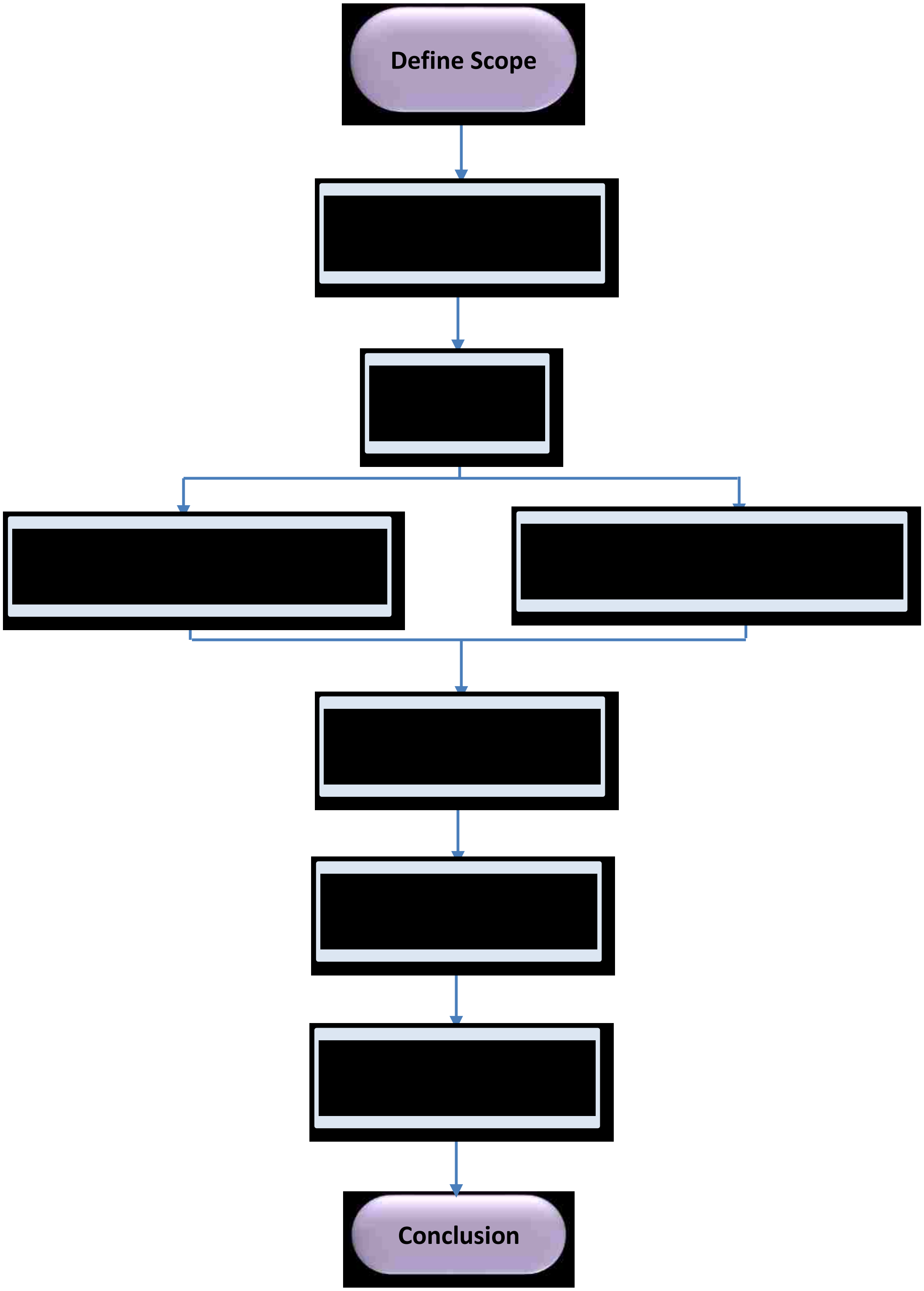


Figure 1 Model Training Procedure

www.manaraa.com

3.2 Data Types

Seven contract procurement data are used as independent variables for the multiple linear regression models and the cost and schedule data are used to develop the performance metrics. There were two performance metrics used in this study. Table 10 shows the dependent and independent variables used in this study.

Table 10 List of Independent and Dependent Variables

S. N.	Independent Variables	Dependent Variables
1	Number of Bidders	Cost Growth
2	Award-Estimate Difference	Schedule Growth
3	Liquidated Damage Rate per Day	
4	Contract Funding	
5	Contractor Type	
6	Project Location	
7	Project Size	

3.3 Independent Variables/ Contract Procurement Factors

There are seven independent variables considered for this study. They are the number of bidders, the award-estimate difference, the liquidated damage rate per day, the type of a contract funding, the type of a contractor, the project location, and the project size. Figure 2 shows the contract procurement data considered for this study. The general definition of each independent variable is described below:

المنارة للاستشارات
www.manaraa.com

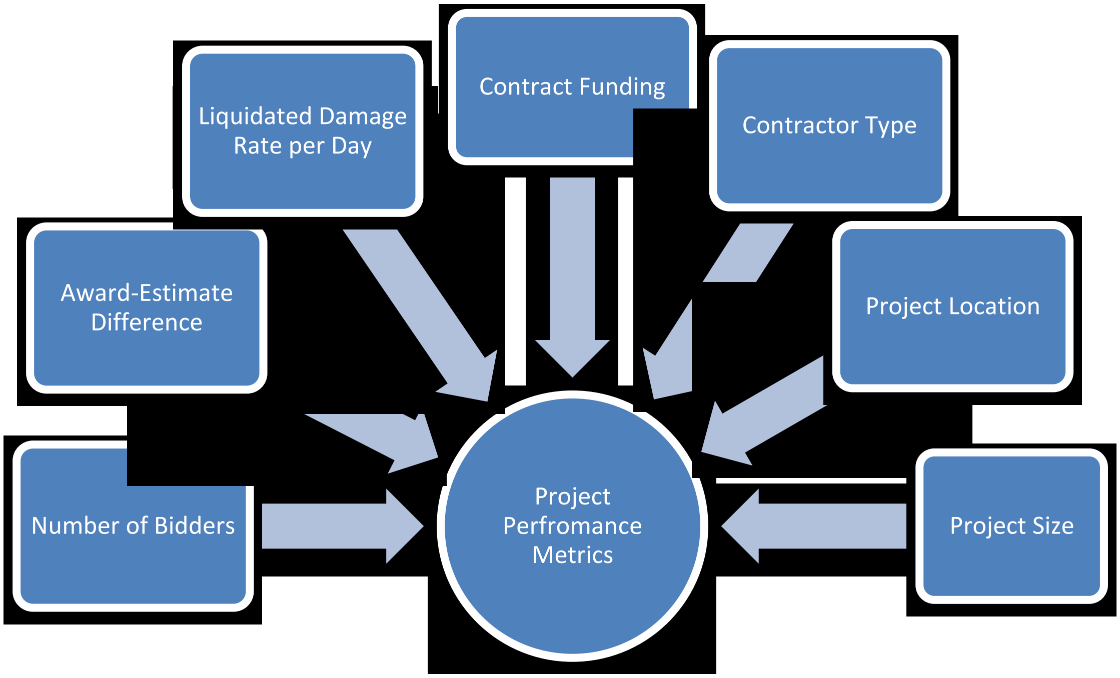


Figure 2 Relationship of Project Performance Metrics with Contract Procurement Data

3.3.1 Number of Bidders

A bidder is a person or company that estimates the cost of a proposed project, submits a tender. It is the total number of bids received on each project. The number of bidders varies from the project to project and depends on many factors that bidders take into account before bidding on the project.

3.3.2 Award-Estimate Difference

It is the percentage of the difference between contract award amount and engineer estimate is calculated. The contract award amount is the price for which the original construction contract was awarded, and engineer's estimate is the estimate of the total project cost by the project owner. The award-estimate difference is calculated using formula as shown in Equation 1.

$$\text{Award-Estimate Difference} = \frac{\text{Contract Award Amount - Engineer Estimate}}{\text{Engineer Estimate}} * 100 \qquad (1)$$

Then, the obtained percentage of difference is categorized into two types: one with a negative

للاستشارات المنارة
www.manaraa.com

difference and another with positive difference. When the contract award amount is less than the engineer's estimated amount then the difference is negative, otherwise positive.

3.3.3 Liquidated Damage Rate per Day

Liquidated Damages are specified daily charges deducted from contractor payment for each day the contractor fails to meet a milestone and/or contract completion date. In other words, it is the price the contractor must pay per day for working beyond the required completion date. It is also defined as the sum of money that is agreed between owner and contractor in a written form as the total amount of compensation an owner should get if the contractor is not able to complete the project within allocated time. Liquidated damages clauses are commonly used in construction contracts.

3.3.4 Contract Funding Type

Contract funding amount is the number of funds obligated on the contract. There are mainly two types of funding for a construction contract: Federal and State. Federal transportation funding is primarily allocated from the Federal Highway Trust Fund, which is capitalized from federal gasoline and diesel taxes; truck, bus, and trailer taxes; tire taxes; Heavy vehicle usage fees, and taxes on alternative fuels. States and local governments obtain income from a variety of sources, the breakdown changes drastically from state to state. The proportions vary based on the types of taxes and fees administered within state borders, the type of resources within the state, the number of intergovernmental transfers, and the policy priorities of state and local governments. State and local government tend to obtain the tax revenues from property taxes, sales and gross receipts taxes, personal income taxes, corporate income taxes, and other types of taxes such as motor vehicle licenses taxes. There is a fixed amount spending allocated every year for both federal and state construction projects.

3.3.5 Contractor Type

A contractor is an entity engaged in the business of construction, repair, and maintenance of

المنارة للاستشارات
www.manaraa.com

projects such as transportation. A contractor working on any project in the state must hold a license in that state. There are mainly two types of contractors, in-state contractor and out-of-state contractor. The prime contractor on each project if has his/ her home office or a permanent field office in the same local geographic area as the project job site then it is categorized under state contractor; else out-of-state contractor. Out-of-state contractors are even required to supply an extra bond in comparison to in-state contractors that equal to a certain percentage of contract depending on the state.

3.3.6 Project Location

The project location is the categorized as rural or urban. (Patrick et al., 2016) writes that if the population of the county is less than 55,000, then the county is known as rural county otherwise urban county. This criterion is based on "Rural" definition in Texas statutes and Texas administrative code as defined by TxDOT. In order to divide the projects into these categories, initially, the population of Texas and Florida were collected from (Demographics, 2016). Then, they were divided based on the population in a county.

3.3.7 Project Size

Project size is the contract award amount to complete the construction of a project.

3.4 Dependent Variables/ Performance Metrics

The performance metrics considered in this study are cost and schedule. The metrics used for cost and schedule are shown in Figure 3. These performance metrics are considered as dependent variables in this research. The prediction of each dependent variable of a project, in this study, will be conducted by determining the significant contract procurement variables. Figure 3 shows the relationship of project performance with all seven contract procurement data. The general definition of each dependent variable is described below:

www.manaraa.com

Figure 3 Division of Performance Measuring Metrics

3.4.1 Cost Growth (CG)

This metric shows the cost growth of the project from its bid phase to completion. The cost growth is defined as the percentage of difference in cost, plus or minus, between final paid amount and contract award amount. The cost growth is calculated using formula as shown in Equation 2.

$$\text{Cost Growth (\%)} = \frac{\text{Final Amount - Contract Award Amount}}{\text{Contract Award Amount}} * 100 \qquad (2)$$

3.4.2 Schedule Growth (SG)

The schedule growth is the percentage of the difference between days used and bid days. The schedule growth is calculated using formula as shown in Equation 3.

$$\text{Schedule Growth (\%)} = \frac{\text{Days Used - Bid Days}}{\text{Bid Days}} * 100 \qquad (3)$$

3.5 Questionnaire Development

Once the dependent and independent parameters are selected, the questionnaire is prepared to collect the data. It was mainly divided into three sections: General Information, Procurement Data, and

www.manaraa.com

Performance Data. The first section contains general information of a project contains contract identification number and contract description. The second section contains data that are available during procurement phase of the project. In addition, the third section contains data that are available after completion of the project.

Section 1 - General Information

- Contract #
- Contract Description

Section 2 - Procurement Data

- Number of Bidders
- Bidding Amount of all Bidders
- Liquidated Damage Rate per Day
- Type of Contract Funding
- Contractor Address
- Contractor Type
- District
- County

Section 3 - Performance Data

- State DOT Engineer's Estimated Amount
- Contract Bid Amount
- Final Paid Amount to Contractor
- Bid Days
- Days Used
- Job Letting Date
- Notice to Proceed (NTP) Date
- Final Completion Date
- Total Number of Change Orders
- Total Amount Paid for Change Orders

3.6 Data Collection

The data were collected from two state department of transportation: Texas DOT and Florida DOT. The database was created based on PDF files, Excel files, or from online database search forms.

3.6.1 Texas DOT Data Collection

The different approaches were made to collect the data from Texas DOT. At first, the defined parameters for this research were searched on TxDOT website. However, all those parameters were not available on its website, TxDOT personnel was contacted to collect all other missing parameters.

المنارة للاستشارات
www.manaraa.com

3.6.1.1 Data Availability

The data were obtained from two sources: TxDOT website and TxDOT personnel. The TxDOT website data contained parameters/variables such as liquidated damage rate per day, type of contract funding, etc. On the other hand, the TxDOT personnel data contained data such as number of bidders, bid amount of all bidders, type of contractor, number of change orders, total change order amount, etc. Both databases contained common parameters such as control ID (unique value), contract bid amount, TxDOT engineer's estimated amount, final paid amount to the contractor, bid days, days used, job letting date, notice to proceed date, completion date. By verifying these common parameters, the data from these two databases were combined to create a new database containing all the procurement and performance metrics information required for the analysis purposes.

3.6.1.2 Data Preparation

From the combined database, the project data with anomalies were eliminated before analysis. The database contained four contract type data, namely, Construction, Building, Local, and Maintenance. Out of these four types, most data had Construction as contract type, therefore; other contract type data were eliminated. A construction contract is a contract entered for the construction, reconstruction, or maintenance of a segment of the state highway system. Out of construction contract type, the new construction and reconstruction projects were only included in the study. In addition, all parameters of each project were checked for any missing values and it was found that some parameters had missing values for "Days Used" and "Paid Amount". The projects with missing values were removed and were not used in the analysis. Moreover, the project data for which the lowest bid was not the winning bid were also eliminated. After the elimination process, the combined database contained data of 1,639 projects that were completed between years 2000 and 2016 and had projects with cost range between 62 thousand and 191 million dollars.

المنارة للاستشارات
www.manaraa.com

3.6.2 Florida DOT Data Collection

Similarly, the different approaches were made to collect the data from Florida DOT. At first, the defined parameters for this research were searched on FDOT website. However, all those parameters were not available on its website, FDOT personnel was contacted to collect all other missing parameters.

3.6.2.1 Data Availability

The data related to contract procurement and performance metric were obtained from FDOT website. The construction office report contained information such as FDOT engineers' estimate, final construction cost, bid days, days used, number of change orders, and total change order amount. In addition, the bid tabulation reports contained data that showed all bidders that bid the job and what their respective bids were for that project. Both reports contained common parameters such as contract ID and contract bid amount. By verifying these common parameters, the data from these two reports were combined to create a new database. In addition, the list of prequalified contractors gave information about bidder's name and their office address as shown in Table 11

.

المنارة للاستشارات
www.manaraa.com

Table 11 Liquidated Damage Rate per Day calculated based on Contract Bid Amount in every 2 years

Original Contract Amount	Liquidated Damage Amount per Day							
	2002-2003	**2004-2005**	**2006-2007**	**2008-2009**	**2010-2011**	**2012-2013**	**2014-2015**	**2015-2016**
$50,000 and under	$674	$674	$313	$278	$278	$642	$836	$763
Over $50,000 but less than $250,000	$544	$544	$580	$388	$388	$758	$884	$958
$250,000 but less than $500,000	$634	$634	$715	$566	$566	$966	$1,074	$1,099
$500,000 but less than $2,500,000	$1,288	$1,288	$1,423	$1,148	$1,148	$1,532	$1,742	$1,584
$2,500,000 but less than $5,000,000	$2,470	$2,470	$2,121	$1,914	$1,914	$2,374	$2,876	$2,811
$5,000,000 but less than $10,000,000	$3,730	$3,730	$3,057	$2,514	$2,514	$3,226	$3,770	$3,645
$10,000,000 but less than $15,000,000	$5,240	$5,240	$3,598	$3,300	$3,300	$4,624	$4,624	$4,217
$15,000,000 but less than $20,000,000	$6,078	$6,078	$4,544	$3,782	$3,782	$4,276	$5,696	$4,698
$20,000,000 and over	$8,624+A	$8,624+A	$8,537+A	$5,684+B	$5,684+B	$7,864+B	$9,788+B	$6,323+B

Note: A= 0.00027 of any amount over 20 million, B= 0.00005 of any amount over 20 million

المنارة للاستشارات
www.manaraa.com

3.6.2.2 Data Preparation

All parameters of each project were checked for any missing values and it was found that some parameters had missing values for "Days Used" and "Paid Amount". The projects with missing values were removed and were not used in the analysis. Moreover, the project data for which the lowest bid was not the winning bid were also eliminated. After the elimination process, the combined database contained data of 1,806 projects that were completed between years 2000 and 2016 and had projects with cost range between 19 thousand and 112 million dollars.

3.7 Data Transformation

The time adjustments were done for liquidated damage rate per day and contract bid amount to establish a more direct comparison of the projects. The ENR Cost Index was used to convert these cost variables based on their job letting calendar year to 2016 equivalent cost using time adjustment formula. Table 12 below shows the ENR Cost Indices and the multiplication factors used to adjust the cost.

Table 12 ENR Cost Indexes and Multiplication Factors from Year 2000 - 2016

Year	Index
2000	6221
2001	6334
2002	6538
2003	6695
2004	7115
2005	7446
2006	7751
2007	7967
2008	8310

Year	Index
2009	8570
2010	8800
2011	9070
2012	9308
2013	9547
2014	9806
2015	10035
2016	10379

Time Adjustment Using the Historical Cost Indexes can be done using Equation 4:

$$\frac{\text{Cost in Year A}}{\text{Cost in Year B}} = \frac{\text{Index for Year A}}{\text{Index for Year B}} \qquad (4)$$

3.8 Data Measurement Type

A variable has one of the four various levels of measurements: Nominal, Ordinal, Interval, or Ratio.

للاستشارات المنارة
www.manaraa.com

Interval and Ratio levels of measurement are sometimes called continuous or scale. Nominal variables are used for labeling variables without any quantitative value. For example, gender and hair color. Ordinal scales are used when the order of the values is important and significant but the difference between each one is not really known, e.g. satisfaction levels. Interval scales are numeric scales in which the exact difference between the values matter, e.g. difference between 60 and 50 degrees. Ratio scales are like interval scales with a clear definition of zero, e.g. height and weight.

The measurement of data types that we have for independent and dependent variables are shown in Table 13. The data used for this research are of two types: scale and nominal.

Table 13 Type of Data Measurement for Independent and Dependent Variables

Parameters Type	**Parameters**	**Data Types**
Independent Variables	Number of Bidders (count)	Scale
	Award-Estimate Difference	Nominal
	Liquidated Damage Rate per Day (in thousands)	Scale
	Contract Funding	Nominal
	Contractor Type	Nominal
	Project Location	Nominal
	Project Size	Scale
Dependent Variables	Cost Growth (%)	Scale
	Schedule Growth (%)	Scale

3.9 Research Hypothesis for Multiple Linear Regression

Two research hypotheses were developed for this study. Some of these contract procurement variables had a combined effect on each of these two project performance metrics. Each of the research hypothesis was converted to null hypothesis to conduct the statistical test. The research and null hypothesis for this study are explained below.

3.9.1 Hypothesis Related to Cost Growth

The research hypothesis for cost growth states that cost growth had a correlation with a number of bidders, award-estimate difference, liquidated damage rate per day, contract funding, contractor type, project location, and project size on projects.

www.manaraa.com

The null hypothesis states that the correlation coefficients of cost growth with a number of bidders, award-estimate difference, liquidated damage rate per day, contract funding, contractor type, project location, and project size on projects are not significantly different from zero. It can be mathematically expressed as in Equation 5.

$$\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0 \qquad (5)$$

If the p-value of each of these correlation coefficients is less than 0.05, the correlation is defined as significant. If the p-value for these correlation coefficients is less than 0.01, then the correlation is defined as highly significant.

3.9.2 Hypothesis Related to Schedule Growth

The research hypothesis for schedule growth states that schedule growth had a correlation with a number of bidders, award-estimate difference, liquidated damage rate per day, contract funding, contractor type, project location, and project size on projects.

The null hypothesis states that the correlation coefficients of schedule growth with a number of bidders, award-estimate difference, liquidated damage rate per day, contract funding, contractor type, project location, and project size on projects are not significantly different from zero. It can be mathematically expressed as in Equation 6.

$$\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0 \qquad (6)$$

If the p-value of each of these correlation coefficients is less than 0.05, the correlation is defined as significant. If the p-value for these correlation coefficients is less than 0.01, then the correlation is defined as highly significant.

www.manaraa.com

3.10 Statistical Background

There are two types of statistical analyses used to interpret the results of quantitative data. They are Descriptive Statistics and Inferential Statistics. Descriptive Statistics is used to summarize dataset using main characteristics and visual methods. Inferential Statistics explains data based on formal modeling or hypothesis.

3.10.1 Descriptive Statistics

The statistics used only to describe the sample or summarize information about the sample is known as descriptive statistics. Examples are mean, mode, and median. Box plot is a convenient way of graphically depicting groups of numerical data through their quartiles. The lines extending vertically from the boxes (whiskers) indicate variability outside the upper and lower quartiles. The spacing between the various parts of the box indicated the degree of dispersion and skewness in the data, and show outliers as shown in Figure 4.

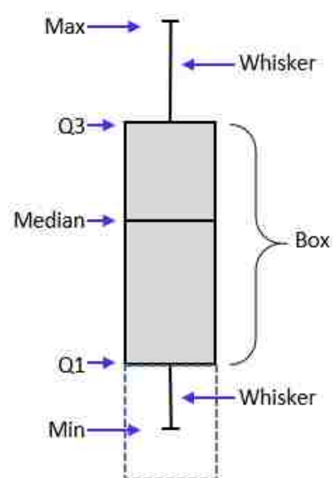


Figure 4 Interpretation of Boxplot

3.10.2 Inferential Statistics

Statistics used to make inferences or generalization about the broader population is known as

المنارة للاستشارات
www.manaraa.com

inferential statistics, e.g. correlation test, t-test, and regression analysis. There are mainly two types of data: one is quantitative, and another is qualitative. The qualitative variables are coded using dummy variables. The dummy coding compares each level of categorical variable to a fixed reference level. The independent variables: award-estimate difference, contract funding, contractor type, and project location used dummy coding during statistical analysis.

- Award-Estimate Difference: Negative = 0 and Positive =1
- Contract Funding: State = 0 and Federal = 1
- Contractor Type: Local = 0 and Out-of-State = 1
- Project Location: Rural = 0 and Urban = 1

3.10.2.1 Multiple Linear Regression Analysis

Before discussing this regression analysis procedure, it is important to discuss about the sample size requirements for this type of data analysis. Therefore, the detailed about this sample size requirement is discussed below.

3.10.2.1.1 Sample Requirement for Data Analysis

The sample is the part of the population that helps us to draw inferences about the population. The appropriate sample size is required to make the inferences about the population based on the sample. A rule of thumb for using regression analysis is at least 10 observations per independent variable. That means if we are using seven independent variables for regression analysis then the minimum sample size should be 60. However, (Green, 1991) suggest the total number of the sample should be greater than 50+m (where m is the number of independent variables) for testing the multiple correlations.

Moreover, (Knofczynski & Mundfrom, 2007) has conducted a Monte Carlo simulation to find out the minimum number of samples for using multiple linear regression for prediction. The authors have found that it mainly depends on the number of independent variables and squared multiple correlation

المنارة للاستشارات
www.manaraa.com

coefficients. As the squared multiple correlation coefficients decreased, the sample size increased at an increasing rate. The number of minimum samples required to use five and seven predictors in predicting using multiple linear regression is shown in Table 14.

Table 14 Sample Size recommendation based on R^2 and Number of Independent Variables

R^2	**Good prediction level**		**Excellent prediction level**	
	5	**7**	**5**	**7**
.10	550	700	2,200	2,800
.15	340	440	1,400	1,800
.20	260	320	950	1,300
.25	180	240	750	950
.30	150	190	600	800
.40	95	120	380	480
.50	65	85	230	320
.70	35	40	110	140
.90	11	14	35	40

Note: R^2 is squared multiple correlation coefficients

3.10.2.1.2 Multiple Linear Regression

Linear regression is an approach to modeling the relationship between a scalar dependent variable and one or more explanatory variables. The case of one explanatory variable is called simple linear regression. For more than one explanatory variable, the process is called multiple linear regression. As a predictive analysis, the multiple linear regressions are used to explain the relationship between one continuous dependent variable and two or more independent variables. An independent variable is a variable that is being manipulated in an experiment to observe the effect on a dependent variable. The independent variables can be either continuous (i.e. an interval or ratio variable) or categorical (i.e. an ordinal or nominal variable). In addition, the dependent variable is a variable that is dependent on an independent variable(s). The dependent variable should be measured on a continuous scale (i.e. either an interval or ratio variable). Regression analysis assumes a dependence or causal relationship between one or more independent and one dependent variable. The estimation technique used for multiple linear regression in R-software is an ordinary least square method. The general estimated multiple regression

المنارة للاستشارات
www.manaraa.com

relationships between dependent and independent variables is shown in Equation 7.

$$\mu = b_0 + b_1x_1 + b_2x_2 + \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots b_px_p \qquad (7)$$

where, b_0, b_1, b_2, ……. b_p is the estimates of x_0, x_1, x_2, ……. x_p

μ = predicted value of the dependent variable

The key assumptions of multiple linear regression are required to be tested (Statistics, 2017). They are described below:

- **Multivariate Normality**: It assumes that the error between observed and predicted values (the residuals of the regression) is normally distributed. It can be checked by reviewing quantile-quantile (Q-Q plot) of residual values. The Q-Q plot is an exploratory graphical device used to check the validity of a distributional assumption of a dataset. If the data follow the assumed distribution, then the points on the q-q plot will fall approximately on a straight line. In addition, normality can also be checked with a goodness of fit test (e.g. Shapiro -Wilk Normality test) on residuals.
- **Linear Relationship**: It assumes that there is a linear relationship between a) dependent variable and each of your independent variables, b) the dependent variable and independent variables collectively. The linearity assumption can best be tested with scatter plots. A plot of the standardized residuals versus the predicted y values shows whether there is a linear or curvilinear relationship.
- **Multicollinearity Test**: When independent variables are correlated with each other, there is a presence of multicollinearity. It assumes that two or more independent variables are not highly correlated with each other. Multicollinearity occurs when the independent variables are not independent of each other. Multicollinearity is checked against correlation matrix and variance inflation factor. The variance inflation factor (VIF) of the linear regression should be less than 10.

المنارة للاستشارات
www.manaraa.com

If VIF>10, there is an indication for multicollinearity to be present.

3.11 Data Modeling

The combined database contains data of transportation projects completed between years 2000 and 2016. The preprocessing of data was done based on criteria explained in section data preparation. Then, the data was divided into 80:20 ratio for training and testing data respectively. The training data is used for model development and testing data is used for its validation. The method selection allows specifying how independent variables are entered, then eliminated if not significant from the analysis. Initially, all the independent were used for the analysis. Then, the independent variable with Variance Inflation Factor (VIF) greater than 5 is eliminated one by one, considering highest VIF as the first one for elimination. Once, all variables with VIF greater than 5 are eliminated then the independent variables with p-values greater than 0.05 are eliminated individually. Only significant variables are used to create multiple linear regression models for forecasting dependent variables. Then, the model validation is conducted to determine if the predicted values are within the 20% range of the actual values. Figure 5 shows the flowchart for a step-by-step process of the data modeling and validation.

المنارة للاستشارات
www.manaraa.com

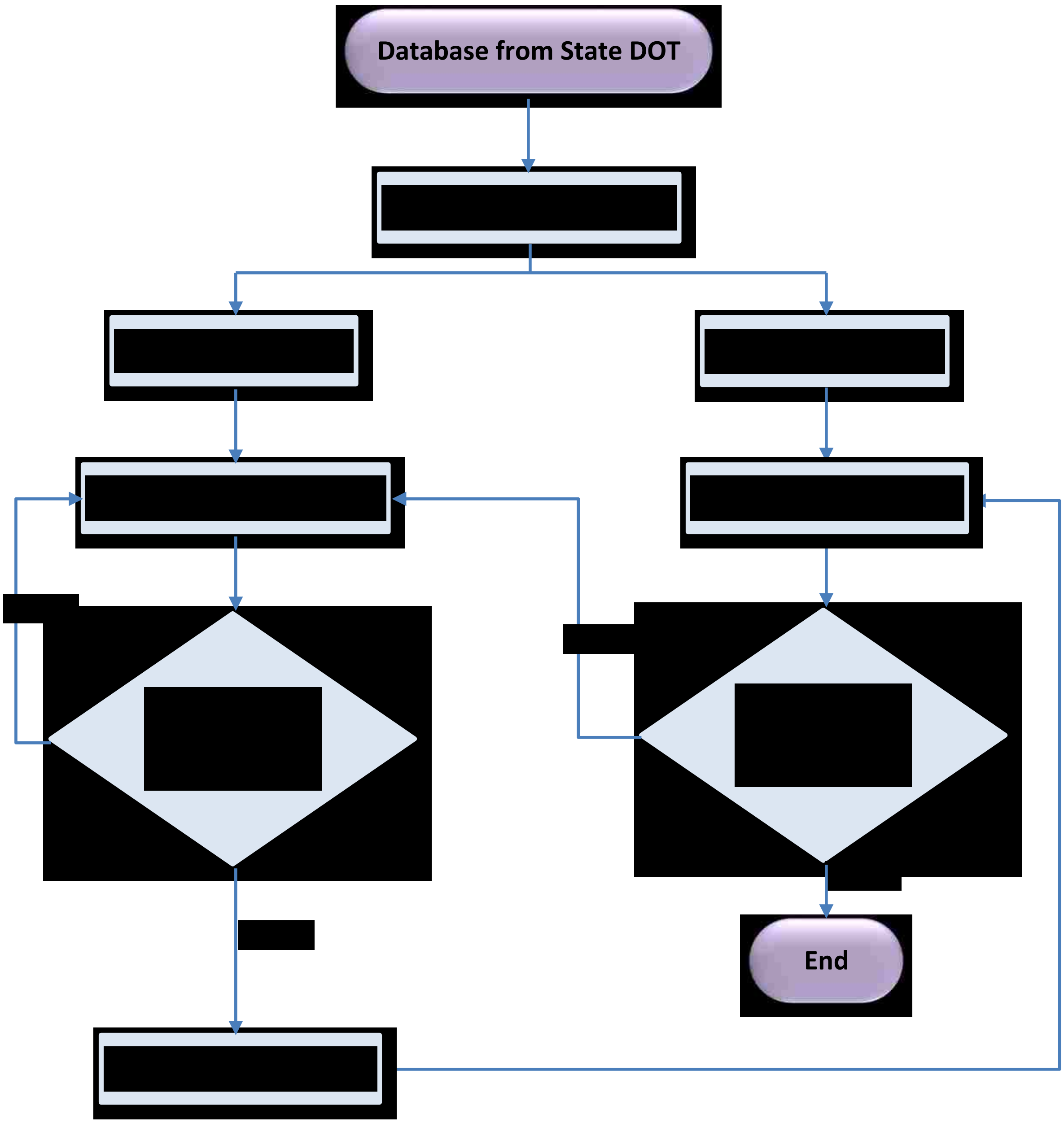


Figure 5 Flow Chart of Data Modelling

www.manaraa.com

3.12 Data Separation

The collected data were further divided at 80:20 percent ratios for training and testing of the models respectively. Training dataset is for creating a model and testing dataset for then validating the models.

3.12.1 Texas DOT Data Separation

Out of 1,639 completed projects, 1,309 projects were separated randomly as a training dataset. The remaining 330 projects were used later for testing of the final models. Figure 6 shows two data division: the training dataset for creating a predicting model for the cost growth and the scheduled growth and testing dataset for validating the obtained models.

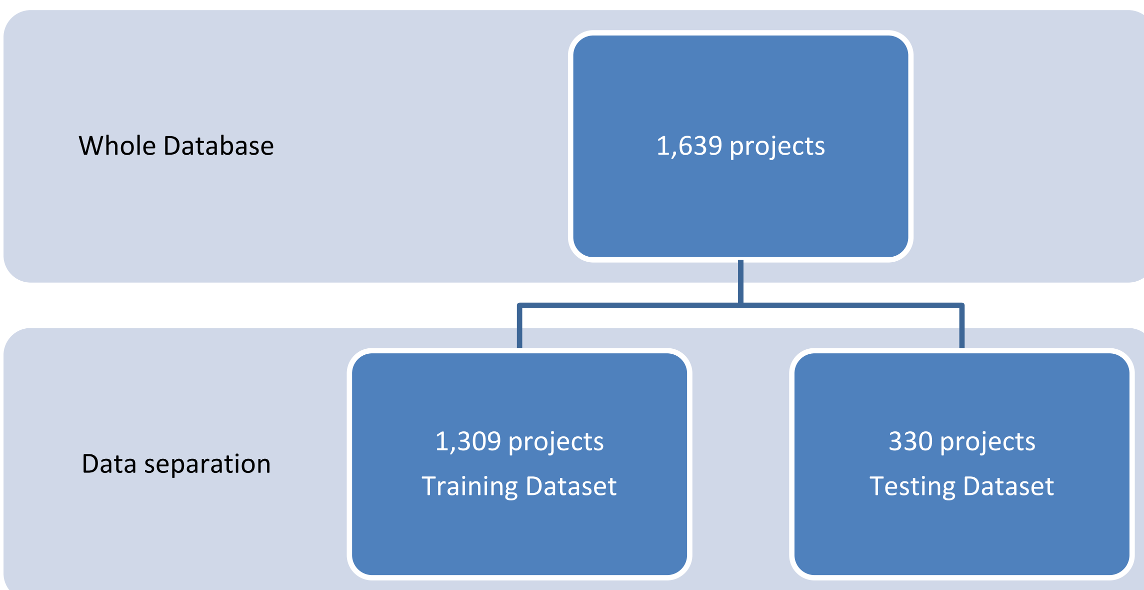


Figure 6 Division of Data for TxDOT projects

3.12.2 Florida DOT Data Separation

Out of 1,806 completed projects, 1,445 projects were separated randomly as a training dataset. The remaining 361 projects were used later for testing of the final models. Figure 7 shows two data division: the training dataset for creating a predicting model for the cost growth and the schedule growth

www.manaraa.com

and testing dataset for validating the obtained models.

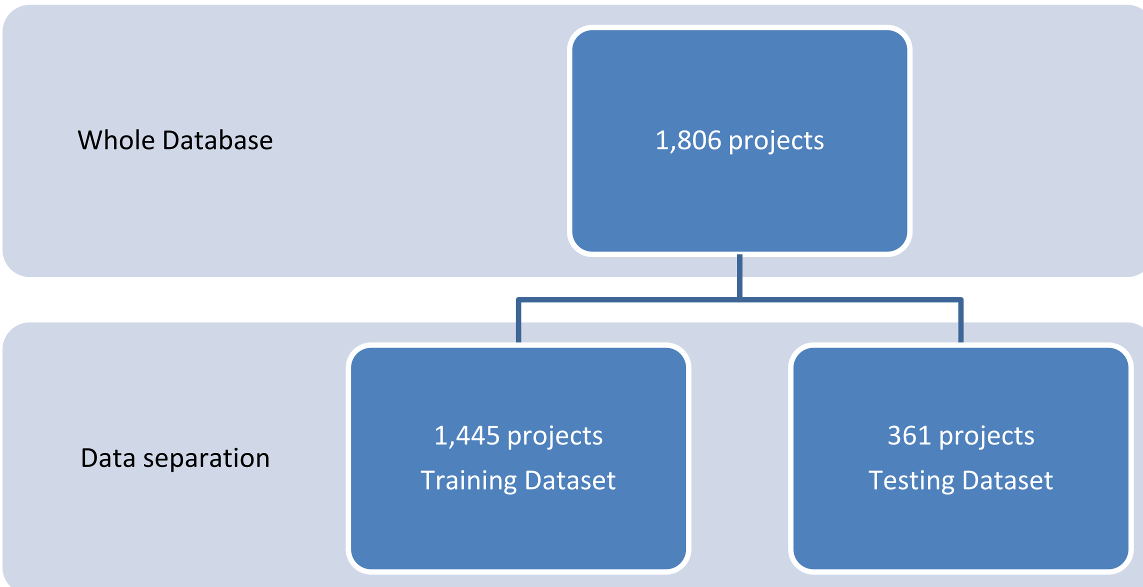


Figure 7 Division of Data for FDOT projects

3.13 Model Validation

It is the process of checking the models, whether the predicted values from the models are closed to the actual values. The validation process can involve analyzing the goodness of fit of the regression, analyzing whether the regression residuals are random, and checking whether the model's predictive performance deteriorates substantially when applied to data that were not used in the model estimation. In other words, the model validation involves assigning values of independent variables to the generated linear regression equation for each dependent variable and then measuring them against the desired response using predefined performance criterion e.g. mean absolute percentage error.

www.manaraa.com

CHAPTER 4 RESULTS

The data is initially analyzed using descriptive statistics such as boxplots, histograms, and summary tables. Then, statistical tests such as multiple linear regression were performed on data using the R-software. These statistical tests were conducted to create multiple linear regression models to understand the combined effect of independent variables on each of the two dependent variables: cost growth and schedule growth. Each model will predict each dependent variable. The results from the data analysis are described separately below for both state DOTs, starting with projects collected from TxDOT and then, followed by FDOT. Lastly, the data analysis for combined data from both state DOTs is presented.

4.1 Results of Texas DOT Data

Firstly, the descriptive statistics of all independent and dependent variables were presented for TxDOT projects. Secondly, the multiple linear regression was conducted for each performance metric. Before conducting multiple linear regression, the related assumptions of the statistical test were initially verified which allows the model development. Later, the developed models were validated using independent set of the data by checking the variation between observed and predicted value of each dependent variable

4.1.1 Descriptive Statistics of TxDOT Projects

For the 1,639 projects collected, the data is further divided into two sets: training and testing datasets. The training dataset consists of 1,309 projects and testing dataset contains 330 projects. The descriptive and statistical analysis is done for the training dataset only. The testing dataset is used for validation purposes.

There were several types of projects such as construction and reconstruction. The maximum and minimum contract duration of a project as per bidding document was found 1,103 days and 13 days

www.manaraa.com

respectively. On average, the bid duration of a project was found 179 days. However, due to several reasons such as unforeseen changes, the act of God, delays due to the owner, contractor, or consultant, there is variation in the actual completion duration of the project. The maximum and minimum actual duration found to complete a project were 1,910 days and 3 days respectively. On average, the bid duration of a project was found 201 days. Table 15 shows the statistics of the bid and actual days for training dataset.

Table 15 Summary of Bid and Actual Days of TxDOT projects (N=1,309)

Days	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
Bid	13	67	125	179	230	1103
Actual	3	75	135	201	257	1910

4.1.1.1 Descriptive Statistics of Independent Variables

The descriptive statistics of seven independent variables are presented separately in each section. There are mainly two data types of independent variables: continuous and categorical. The boxplot is plotted for each continuous variable considered for this research. The box plot is a useful graphical display for describing the behavior of the data in the middle as well as at the ends of the distribution. The continuous variables are a number of bidders, award-estimate difference, and liquidated damage rate per day. In addition, the bar chart is plotted for three categorical independent variables: contract funding type, contractor type, and project location.

4.1.1.1.1 Number of Bidders

Out of the total 1,309 projects, the maximum number of bidders observed for a project was 19 and the minimum number of bidders observed for a project was 3. On average, there were 5.67 bidders for each project as shown in Table 16.

المنارة للاستشارات
www.manaraa.com

Table 16 Descriptive Statistics of Number of Bidders of TxDOT Projects

Independent Variable	Units	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
Number of Bidders	Count	3	4	5	5.67	7	19

The boxplot for the number of bidders, shown in Fig. A - 1, indicated that there is a presence of outliers. The frequency distribution for a number of bidders is shown in Appendix B (Fig. B - 1). The histogram shows that the distribution is skewed to left indicating that there are more projects with less number of bidders than projects with a higher number of bidders in the database.

4.1.1.1.2 Award-Estimate Difference

The deviation in percentage between the awarded amount and engineer's estimate is either positive or negative. The deviation is negative when the awarded amount is less than engineer's estimated amount and vice versa. Among the projects collected, 65.01% (851) projects had awarded amount less than engineer's estimate and 34.99% (458) of the project had awarded amount greater than engineer's estimate.

4.1.1.1.3 Liquidated Damage Rate per Day

The liquidated damage rate per day ranged from almost $390 to $3,390. On average, the liquidated damage rate per day was found around $1,001 as shown in Table 17.

Table 17 Descriptive Statistics of Liquidated Damage Rate per Day of TxDOT Projects

Independent Variable	Units	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
LD per Day	$/Day	0.39	0.62	0.88	1.00	1.09	3.39

The boxplot of the liquidated damage rate per day, shown in Fig. A - 3, indicated that there is a presence of extreme outliers. The frequency distribution for liquidated damage rate per day is shown by plotting histogram (refer Appendix B, Fig. B - 3). The distribution is skewed to left indicating that there are more projects with less than $1,001 liquidated damage rate than projects with higher than $1,001 in the database.

المنارة للاستشارات
www.manaraa.com

4.1.1.1.4 Contract Funding Type

Federal and State funding were the two type of funding sources for the project. Among the projects collected, 16.81% (220) of the projects were state-funded and 83.19% (1,089) of projects were federally funded.

4.1.1.1.5 Contractor Type

The construction contractors involved in the project can be local or from out-of-state. Among collected projects, local contractors completed 95.19% (1,246) of total projects and 4.81% (63) of the projects were completed by out-of-state contractors.

4.1.1.1.6 Project Location

The project is either located in rural or urban. Among collected projects from TxDOT, 36.21% (474) of projects were in the rural area and 63.79% (835) of the projects were in the urban area.

4.1.1.1.7 Project Size

Out of 1,309 projects, the contract bid amount ranged from around 0.062 to 191 million dollars. The average project size was found to be around 5 million dollars. Table 18 shows a summary of contract bid amount.

Table 18 Descriptive Statistics of Contract Bid Amount of TxDOT projects (N= 1,309)

State	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
Texas	\$62,578	\$513,732	\$1,083,155	\$5,362,244	\$3,350,239	\$191,240,387

4.1.1.2 Descriptive Statistics of Dependent Variables

The descriptive statistics of two dependent variables are described separately in each section. The dependent variables have a continuous data type. The continuous dependent variables are cost growth and schedule growth.

4.1.1.2.1 Cost Growth

Out of 1,309 projects, the maximum cost underrun and overrun from contract award phase to

المنارة للاستشارات
www.manaraa.com

construction phase was found to be 76.11% and 186.65% respectively. On average, total cost saving was found to be around 2.39% as shown in Table 19. The mean and median values are not close indicating the data are not distributed normally.

Table 19 Descriptive Statistics of Cost Growth of TxDOT Projects

Dependent Variable	Units	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
Cost Growth	%	-76.11	-0.69	2.39	5.19	7.79	186.65

The boxplot of the cost growth, shown in Fig. A - 5, indicated that there is a presence of extreme outliers. The frequency distribution for Cost growth is shown by plotting histogram (refer Appendix B, Fig. B - 5). The distribution of data seems to have skewed to the left.

4.1.1.2.2 Schedule Growth

Out of 1,309 projects, the maximum schedule underrun and overrun during construction phase was found to be 83.33% and 760% respectively. On average, schedule growth was found to be around 11.96% as shown in Table 20. The mean and median values are not close indicating the data are not distributed normally.

Table 20 Descriptive Statistics of Schedule Growth of TxDOT Projects

Dependent Variable	Units	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
Schedule Growth	%	- 83.33	- 7.10	4.11	11.96	25.56	760.00

The boxplot of the schedule growth, shown in Fig. A - 7, indicated that there is a presence of extreme outliers. The frequency distribution for schedule growth is shown by plotting histogram (refer Appendix B, Fig. B - 7). The distribution of data seemed to skew to left.

4.1.2 Results of Multiple Linear Regression for TxDOT Projects

The total number of projects used for creating a training model is 1,309 projects. In addition, 330 projects were used for validation purpose. The results of this multiple linear regression and validation are discussed below.

المنارة للاستشارات
www.manaraa.com

4.1.2.1 Results of Assumption Tests for Multiple Linear Regression

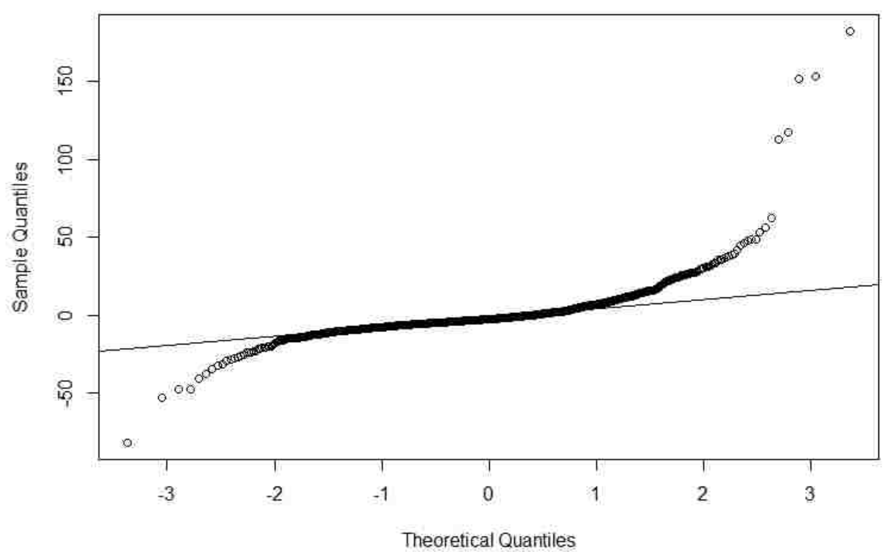
Before the multiple linear regression models are developed, it is necessary to check whether the assumptions used for the multiple linear regressions are fulfilled in the dataset. Three tests related to multivariate normality, linear relationship, and multicollinearity are conducted to verify the assumptions of multiple linear regression. The results of these tests are described below.

4.1.2.1.1 Multivariate Normality Tests of Residuals of Dependent Variables

The multiple linear regression assumes that the variables are normally distributed. Shapiro-Wilk test is conducted to determine whether each dependent variable is normally distributed. In addition, the Q-Q plot is plotted to check whether the data points lie on a straight diagonal line to be normally distributed.

4.1.2.1.1.1 Multivariate Normality Tests of Residuals of Cost Growth

Shapiro-Wilk test was used to check normality of the training dataset of CG. The p-value obtained from Shapiro-Wilk normality test was 2e-16, which is less than 0.05, so the residuals from the fitted linear model are non-normal. In addition, the Q-Q plot in Figure 8 shows the points does not lie on a straight diagonal line stating the data is not normally distributed.



www.manaraa.com

Figure 8 Normal Q-Q Plot of Cost Growth of TxDOT projects (N= 1,309)

Therefore, bootstrapping is conducted to check the significance of variables. As shown in Table 21, using bootstrapping method, zero is found outside of confidence interval for each predictor. Therefore, each remaining predictor in the model is significant. The t-values are also computed for testing whether beta is equal to 0 or not for each predictor. Since each t-value is greater than 2 and error degrees of freedom is 1305, each predictor in the model is highly significant. This shows that the non-normality data of CG will not affect the outcome of the regression model.

Table 21 Confidence Interval and T-Values from Bootstrapping for CG of TxDOT projects

Independent Variables	CIbetaL	CIbetaU	T Observed
Intercept	0.42	7.09	2.21
Liquidated Damage Rate per Day	1.22	4.16	3.59
Contract Funding	- 5.49	- 0.05	- 2.00
Project Location	0.15	3.16	2.16

4.1.2.1.1.2 Multivariate Normality Tests of Residuals of Schedule Growth

Shapiro-Wilk test was used to check the normality of training dataset of SG. The p-value obtained from Shapiro-Wilk normality test was 2e-16, which is less than 0.05, so the residuals from the fitted linear model are non-normal. In addition, the Q-Q plot in Figure 9 shows the points does not lie on a straight diagonal line stating the data is not normally distributed.

المنارة للاستشارات
www.manaraa.com

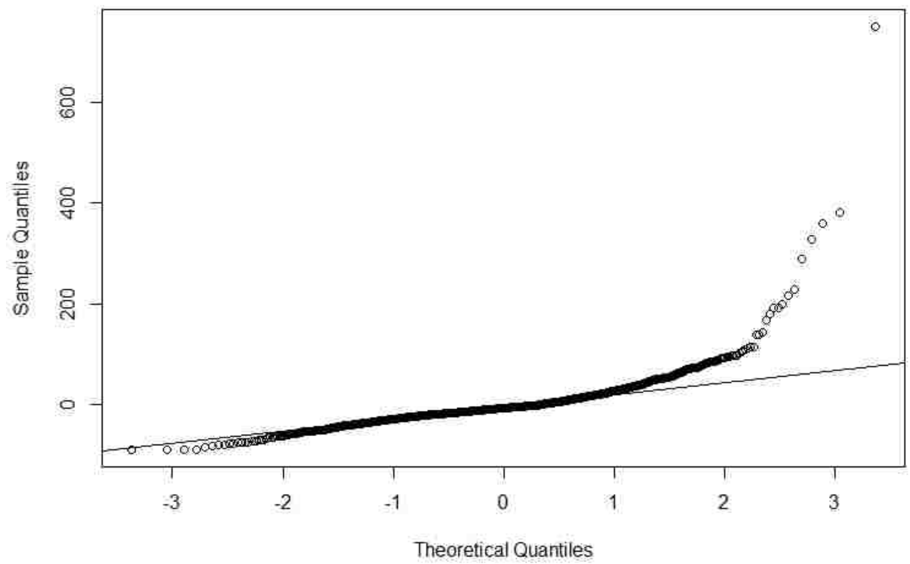


Figure 9 Normal Q-Q Plot of Schedule Growth of TxDOT projects (N= 1,309)

Therefore, bootstrapping is conducted to check the significance of variables. As shown in Table 22, using bootstrapping method, zero is found outside of confidence interval for each predictor. Therefore, each remaining predictor in the model is significant. The t-values as shown below are computed for testing whether beta is equal to 0 or not for each predictor. Since each t-value is greater than 2 and error degrees of freedom is 1304, each predictor in the model is highly significant. This shows that the non-normality data of SG will not affect the outcome of the regression model.

Table 22 Confidence Interval and T-Values from Bootstrapping for SG of TxDOT projects

Independent Variables	CIbetaL	CIbetaU	T Observed
Intercept	- 15.32	- 0.32	- 2.04
Award-Estimate Difference	0.85	11.09	2.28
Liquidated Damage Rate per Day	3.75	11.66	3.82
Contract Funding	2.28	13.46	2.76
Project Location	0.13	10.61	2.01

4.1.2.1.2 Linear Relationship Tests of Dependent Variables

A plot of the standardized residuals versus the predicted values of each outcome is created to determine if linear relationship fits the model. The results of these tests for each of the dependent variables are described below.

المنارة للاستشارات
www.manaraa.com

4.1.2.1.2.1 Linear Relationship Tests of Cost Growth

The residual plots are required to validate the model. Figure 10 shows that the observed error (residuals) of CG is consistent with random and unpredictable errors. This shows that there is linear relationship existed between the CG and all the independent variables.

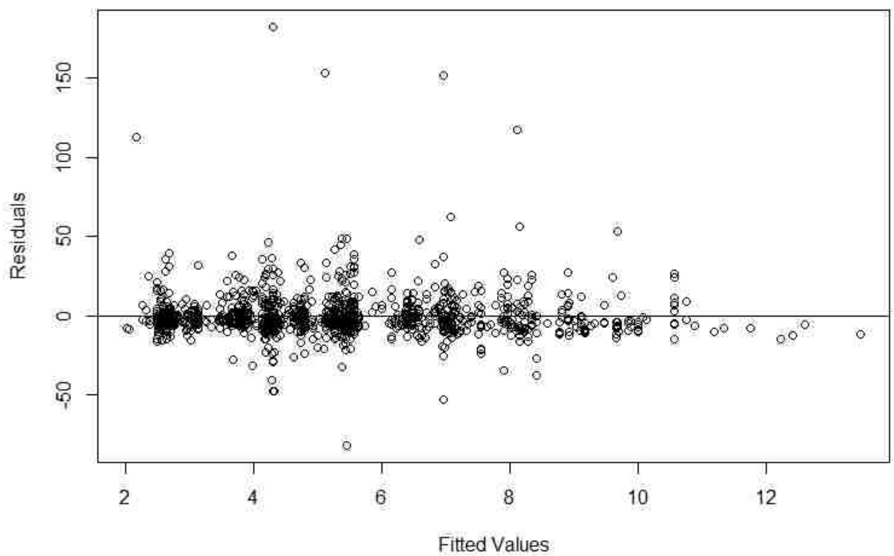


Figure 10 Residual vs Fitted Values - Cost Growth of TxDOT projects (N= 1,309)

4.1.2.1.2.2 Linear Relationship Tests of Schedule Growth

The residual plots are required to validate the model. Figure 11 shows that the observed error (residuals) of SG is consistent with random and unpredictable errors. This shows that there is linear relationship existed between the SG and all the independent variables.

www.manaraa.com

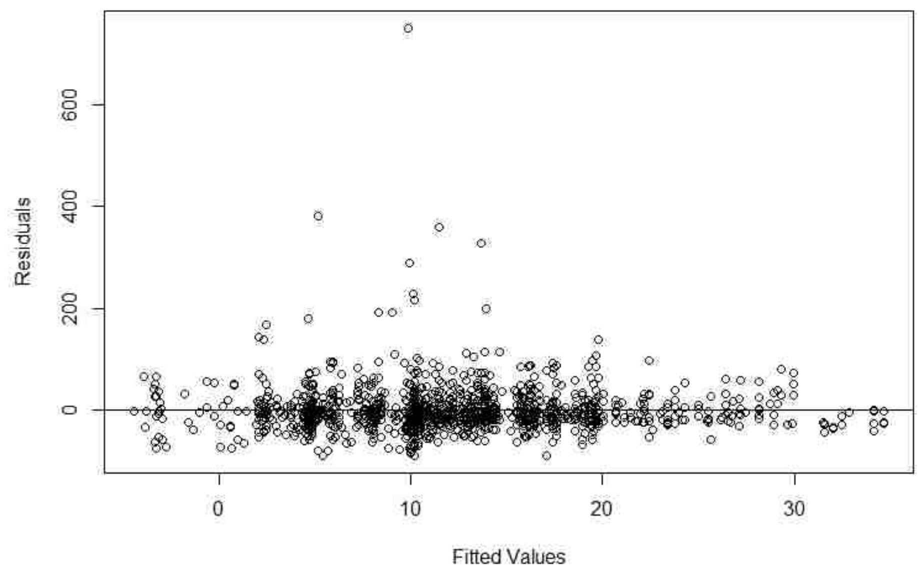


Figure 11 Residual vs Fitted Values - Schedule Growth of TxDOT projects (N= 1,309)

4.1.2.1.3 Results of Multicollinearity Tests of Dependent Variables

Another assumption to conduct multiple linear regression is there should be no multicollinearity between the independent variables. If variance inflation factor (VIF) is greater than 5, multicollinearity existed between the independent variables.

4.1.2.2 Results of Regression Models

A multiple linear regression was used to find a relationship between each dependent variable with seven independent variables i.e. Number of Bidders, Award-Estimate Difference, Liquidated Damage Rate per Day, Contract Funding, Contractor Type, Project Location, and Project Size.

4.1.2.2.1 Multiple Linear Regression of Cost Growth

A multiple linear regression was used to determine the relationship between CG and seven independent variables. Out of the seven independent variables, three of them were found to have a significant correlation with CG as shown in Table 23. The R-square for the model is found to be 2%.

المنارة للاستشارات
www.manaraa.com

Table 23 Significance Level of CG with Independent Variables in TxDOT projects

Independent Variables	Estimate	Std. Error	T-value	P-value	95% Conf. Interval		VIF
					CI upper	CI lower	
Intercept	3.76	1.25	3.00	2e-03**	1.30	6.21	
LD per Day	2.69	0.74	3.61	3e-04**	1.23	4.15	1.03
Contract Funding	- 2.77	1.05	-2.64	8e-03**	- 4.83	- 0.71	1.01
Project Location	1.65	0.82	2.01	0.04*	0.04	3.26	1.03

*significant at p-values less than 0.05 **significant at p-values less than 0.01

Number of Observations = 1,309
F (3, 1305) = 8.67
Prob. > F = 1e-05
R-squared = 0.019
Adjusted R-squared = 0.017

The multiple linear regression equation adopted for CG can be described mathematically as in

Cost Growth = 3.76 + (2.69 * Liquidated Damage Rate per Day) - (2.77 * Contract Funding) + (1.65 * Project Location)

The regression equation for CG shows that:

- The coefficient for liquidated damage rate per day is significant, the sign is positive implying it is reasonable. The cost growth decrease by 2.69% when liquidated damage rate per day decrease by one thousand dollars. The one of the probable reason for this could be the projects had excusable and compensable delays which increased the final construction cost of the project. The liquidated damage rate per day does not apply when the delay was due to weather condition, an act of god, and unforeseen conditions. Also, another reason could be a comparative low compensation each day for delaying a project TxDOT based on the size and daily spending of the project. The average LD per day was found around $1,000 for Texas project.
- The coefficient for contract funding is significant, the sign is negative indicates that it is reasonable. The cost growth decrease by 2.77% when contract funding is federal. The likely reason for this can be the federally funded projects have federal requirements, such as meeting DBE goals, environmental compliance documentation, etc. which results in high bid from contractor.

المنارة للاستشارات
www.manaraa.com

The bidder's higher bid on federal projects decreases the possibility of an increase in construction cost.

- The coefficient for project location is significant, the sign is positive that means it is reasonable. The cost growth increase by 1.65% when project location is urban. One of the probable reason is in urban areas, the competitive nature of bidding forced contractors to bid low. Therefore, they will try to make up the difference in cost by adding change orders during the construction phase of a project. Another probable reason for this is the highways built in an urban area are congested and the contractors need to more work in planning to execute the project successfully. Moreover, the contractors need to be more innovative in managing materials and choosing construction methods, so that there will be less disturbance in the movement of the people and vehicles. This can lead to higher cost overruns.
- The cost growth will be 3.76% when liquidated damage rate per day is zero; contract funding is state; and project location is rural.

4.1.2.2.2 Multiple Linear Regression of Schedule Growth

A multiple linear regression was used to determine the relationship between SG and seven independent variables. Out of the seven independent variables, four of them were found to have a significant correlation with SG as shown Table 24. The R-square for the model is found to be 2%.

Table 24 Significance Level of SG with Independent Variables in TxDOT Projects

Independent Variables	Estimate	Std. Error	T-value	P-value	95% Conf. Interval		VIF
					CI upper	CI lower	
Intercept	- 7.82	4.11	-1.91	0.05*	- 15.87	0.23	
Award-Estimate Difference	5.97	2.64	2.26	0.02*	0.78	11.15	1.01
LD per Day	7.71	2.41	3.20	1e-03**	2.98	12.43	1.04
Contract Funding	7.87	3.38	2.33	0.02*	1.24	14.50	1.02
Project Location	5.37	2.66	2.02	0.04*	0.16	10.58	1.04

*significant at p-values less than 0.05 **significant at p-values less than 0.01

Number of Observations = 1,309
F (4, 1304) = 7.28

المنارة للاستشارات
www.manaraa.com

Prob. > F = 8e-06
R-squared = 0.022
Adjusted R-squared = 0.019

The multiple linear regression equation adopted for SG can be described mathematically as in

Schedule Growth = - 7.82 + (5.97 * Award-Estimate Difference) + (7.71 * Liquidated Damage Rate per Day) + (7.87 * Contract Funding) + (5.37 * Project Location)

The regression equation for SG shows that:

- The coefficient for the difference between award and estimate is significant, the sign is positive that shows it is reasonable. The schedule growth increase by 5.97% when award amount is greater than engineer's estimated amount. There could be many reasons besides the positive deviation between the awarded amount and the estimated amount such as weather condition, geographic location, and change orders during the construction phase of a project which leads to the project delay. The study does not consider the quantitative analysis to find out by how much the positive or negative deviation between the awarded amount and the estimated amount affects the schedule growth and is recommended for future studies.
- The coefficient for liquidated damage rate per day is significant, the sign is positive that means it is reasonable. The schedule growth increase by 7.71% when liquidated damage rate per day increase by one thousand dollars. This could be because when the projects had a higher LD per day, the contractors were more driven to take the risk to complete the project on time. The resulting schedule pressure led to out-of-sequence work that delays the project. In addition, when the liquidated damage rate per day is lesser than the cost to accelerate per day during the construction of a project, the contractor prefers to delay the project. Another reason can be similar to that mentioned in cost growth findings regarding inconsistency in providing the LD per day by TxDOT.

المنارة للاستشارات
www.manaraa.com

- The coefficient for contract funding is significant, the sign is positive implying it is reasonable. The schedule growth increase by 7.87% when contract funding is federal. The reasons could be due to federal contract requirements (such as meeting DBE goals, environmental compliance documentation, etc.) which involved more time for the contractors.
- The coefficient for project location is significant, the sign is positive that indicates it is reasonable. The schedule growth will increase by 5.37% when project location is urban. This shows that the rural projects were completed earlier than urban projects. The delay could be due to the high complexity of the urban projects compared to rural projects. As the projects became more complex, there was a higher chance of experiencing a schedule delay.
- The project will be ahead of schedule by 7.82% when the award-estimate difference is negative; liquidated damage rate per day is zero; contract funding is state; and project location is rural.

4.1.2.3 Results of Data Validation

The testing data is used to validate both models (Cost Growth and Schedule Growth) developed. The results of the validation are described below.

4.1.2.3.1 Validation of Cost Growth Model

The regression equation determined was used to calculate the cost growth for a testing dataset. The testing dataset was the 330 transportation projects that were randomly selected and not used in training dataset. The difference between the actual and predicted CG values was calculated to check the accuracy of the model. The summary and box plot of the difference between predicted and observed values are shown in Table 25 and Fig. C - 1 respectively. It shows that mean difference is 0.33% and the minimum and maximum difference values were found to be - 51.18% and 87.17%.

Table 25 Summary of Diff. between Predicted and Observed CG Values of TxDOT Projects (N=330)

www.manaraa.com

Cost Growth	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
Predicted - Observed	- 51.18	- 3.72	2.60	0.33	5.62	87.17

The numbers of projects whose difference between predicted and actual CG values are within +/- 20% were calculated. It was found that 304 projects were within the defined range. Therefore, the percentage of the project with ± 20% CG Error is 92.12%. In addition, the absolute average difference in CG between predicted and observed is found to be 7.60%.

4.1.2.3.2 Validation of Schedule Growth Model

The regression equation determined was used to calculate the SG for a testing dataset. The testing dataset was the 330 completed transportation projects. The difference between the actual and predicted SG values was calculated to check the accuracy of the model. The summary and box plot of the difference between predicted and observed values are shown in Table 26 and Fig. C - 3 respectively. It shows that mean difference is 1.62% and the minimum and maximum difference values were found to be - 229.74% and 106.67%.

Table 26 Summary of Diff. between Predicted and Observed SG Values TxDOT Projects (N=330)

Schedule Growth	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
Predicted - Observed	- 229.74	- 15.71	8.33	1.62	21.93	106.67

The numbers of projects whose predicted SG values are within +/- 20% with actual values were calculated. It was found that 167 projects were within the defined range. Therefore, the percentage of the project with ± 20% SG Error is 50.61%. In addition, the absolute average difference in SG between predicted and observed is found to be 27.04%

4.2 Results of Florida DOT Data

Similarly, the descriptive statistics of all independent and dependent variables were firstly presented for FDOT projects. Secondly, the multiple linear regression was conducted for each performance metric. For the multiple linear regression analysis, the test assumptions were verified which

المنارة للاستشارات
www.manaraa.com

allows the model development. Then, the developed models were validated using the independent set of the data by checking the variation between observed and predicted values of each dependent variable.

4.2.1 Descriptive Statistics of Florida DOT Projects

For 1,445 projects collected, the data is divided into two sets: training and testing datasets. The training dataset consists of 1,445 projects and the testing dataset contains 361 projects. The descriptive and statistical analysis is done for the training dataset only. The testing dataset is used for validation purposes.

The maximum and minimum contract duration of a project as per bidding document was found 1,620 days and 19 days respectively. On average, the bid duration of a project was found 213 days. However, due to several reasons such as unforeseen changes, the act of God, delays due to the owner, contractor, or consultant, there is variation in the actual completion duration of the project. The maximum and minimum actual duration found to complete a project were 1,840 days and 3 days respectively. On average, the bid duration of a project was found 262 days. Table 27 shows the statistics of the bid and actual days for training dataset.

Table 27 Summary of Bid and Actual Days of FDOT Projects (N=1,445)

Days	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
Bid	19	100	160	213	250	1,620
Actual	3	110	197	262	318	1,840

4.2.1.1 Descriptive Statistics of Independent Variables

The descriptive statistics of seven independent variables are presented separately in each section. There are mainly two data types of independent variables: continuous and categorical. The boxplot is plotted for each continuous variable considered for this research. The box plot is a useful graphical display for describing the behavior of the data in the middle as well as at the ends of the distribution. The continuous variables are a number of bidders, award-estimate difference, and liquidated damage rate per

المنارة للاستشارات
www.manaraa.com

day.

4.2.1.1.1 Number of Bidders

Out of the total 1,445 projects, the maximum number of bidders observed for a project was 19 and the minimum number of bidders observed for a project was 3. On average, there were 5.31 bidders for each project as shown in Table 28.

Table 28 Descriptive Statistics of Number of Bidders of FDOT Projects

Independent Variable	Units	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
Number of Bidders	Count	3	4	5	5.31	6	19

The boxplot for the number of bidders, shown in Fig. A - 2, indicated that there is a presence of extreme outliers. The frequency distribution for a number of bidders is shown in Appendix B (Fig. B - 2). The histogram shows that the distribution is skewed to left indicating that there are more projects with less number of bidders than projects with a higher number of bidders in the database.

4.2.1.1.2 Award-Estimate Difference

The deviation in percentage between the awarded amount and engineer's estimate is either positive or negative. The deviation is negative when the awarded amount is less than engineer's estimated amount and vice versa. Among the projects collected, 75.99% (1,098) projects had awarded amount less than engineer's estimate and 24.01% (347) of the project had awarded amount greater than engineer's estimate.

4.2.1.1.3 Liquidated Damage Rate per Day

The liquidated damage rate per day ranged from almost $320 to $34,800. On average, the liquidated damage rate per day was found around $2,238 as shown in Table 29.

Table 29 Descriptive Statistics of Liquidated Damage Rate per Day of FDOT Projects

Independent Variable	Units	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
LD per Day	$/Day	320	1,310	1,710	2,238	2,580	34,800

المنارة للاستشارات
www.manaraa.com

The boxplot of the liquidated damage rate per day, shown in Fig. A - 4, indicated that there is a presence of extreme outliers. The frequency distribution for liquidated damage rate per day is shown by plotting histogram (refer Appendix B, Fig. B - 4). The distribution is skewed to left indicating that there are more projects with less than $2,238 liquidated damage rate than projects with higher than $2,238 in the database.

4.2.1.1.4 Contract Funding Type

Federal and State funding were the two type of funding sources of the project. Among the projects collected, 22.15% (320) of the projects were state-funded and 77.85% (1,125) of projects were federally funded.

4.2.1.1.5 Contractor Type

The construction contractors involved in the project can be local or from out-of-state. Among collected projects, local contractors completed 91.56% (1,323) of projects and out-of-state contractors completed 8.44% (122) of the projects.

4.2.1.1.6 Project Location

The project can be located either in rural or urban. Among collected projects, 13.29% (192) of projects were located in rural area and 86.71% (1,253) of the projects were located in urban area.

4.2.1.1.7 Project Size

Out of 1,445 projects, the contract bid amount ranged from around 0.019 to 112 million dollars. The average project size was found to be around 4 million. Table 30 shows a summary of contract bid amount.

Table 30 Descriptive Statistics of Contract Bid Amount of FDOT Projects (N= 1,445)

State	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
Texas	$19,003	$641,503	$1,875,883	$4,177,310	$4,263,425	$112,227,321

المنارة للاستشارات
www.manaraa.com

4.2.1.2 Descriptive Statistics of Dependent Variables

The descriptive statistics of two dependent variables are described separately in each section. The dependent variables have a continuous data type. The continuous dependent variables are cost growth and schedule growth.

4.2.1.2.1 Cost Growth

Out of 1,445 projects, the maximum cost underrun and overrun from contract award phase to construction phase was found to be 79.77% and 156.17% respectively. On average, total cost saving was found to be around 1.31% as shown in Table 31. The mean and median values are not close indicating the data are not distributed normally.

Table 31 Descriptive Statistics of Cost Growth of FDOT Projects

Dependent Variable	Units	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
Cost Growth	%	- 79.77	- 5.52	- 1.83	- 1.31	1.87	156.17

The boxplot of the cost growth, shown in Fig. A - 6, indicated that there is a presence of extreme outliers. The frequency distribution for cost growth is shown by plotting histogram (refer Appendix B, Fig. B - 6). The distribution of data seems to have skewed to the left.

4.2.1.2.2 Schedule Growth

Out of 1,445 projects, the maximum schedule underrun and overrun during construction phase was found to be 91.11% and 377.27% respectively. On average, schedule growth was found to be around 23.06% as shown in Table 32. The mean and median values are not close indicating the data are not distributed normally.

Table 32 Descriptive Statistics of Schedule Growth of FDOT Projects

Dependent Variable	Units	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
Schedule Growth	%	- 91.11	4.76	17.78	23.06	35.00	377.27

The boxplot of the schedule growth, shown in Fig. A - 8, indicated that there is a presence of

المنارة للاستشارات
www.manaraa.com

extreme outliers. The frequency distribution for schedule growth is shown by plotting histogram (refer Appendix B, Fig. B - 8). The distribution of data seemed to skew to left.

4.2.2 Results of Multiple Linear Regression for FDOT Projects

The total number of projects used for creating a training model is 1,445 projects. In addition, 361 projects were used for validation purpose. The results of this multiple linear regression and validation are discussed below.

4.2.2.1 Results of Assumption Tests for Multiple Linear Regression

Before the multiple linear regression models are developed, it is necessary to check whether the assumptions used for the multiple linear regressions are fulfilled in the dataset. Three tests related to multivariate normality, linear relationship, and multicollinearity are conducted to verify the assumptions of multiple linear regression. The results of these tests are described below.

4.2.2.1.1 Multivariate Normality Tests of Residuals of Dependent Variables

The multiple linear regression assumes that the variables are normally distributed. Shapiro-Wilk test is conducted to determine whether each dependent variable is normally distributed. In addition, the Q-Q plot is plotted to check whether the data points lie on a straight diagonal line to be normally distributed.

4.2.2.1.1.1 Multivariate Normality Tests of Residuals of Cost Growth

Shapiro-Wilk test was used to check normality of the training dataset of CG. The p-value obtained from Shapiro-Wilk normality test was 2e-16, which is less than 0.05, so the residuals from the fitted linear model are non-normal. In addition, the Q-Q plot in Figure 12 shows the points does not lie on a straight diagonal line stating the data is not normally distributed.

www.manaraa.com

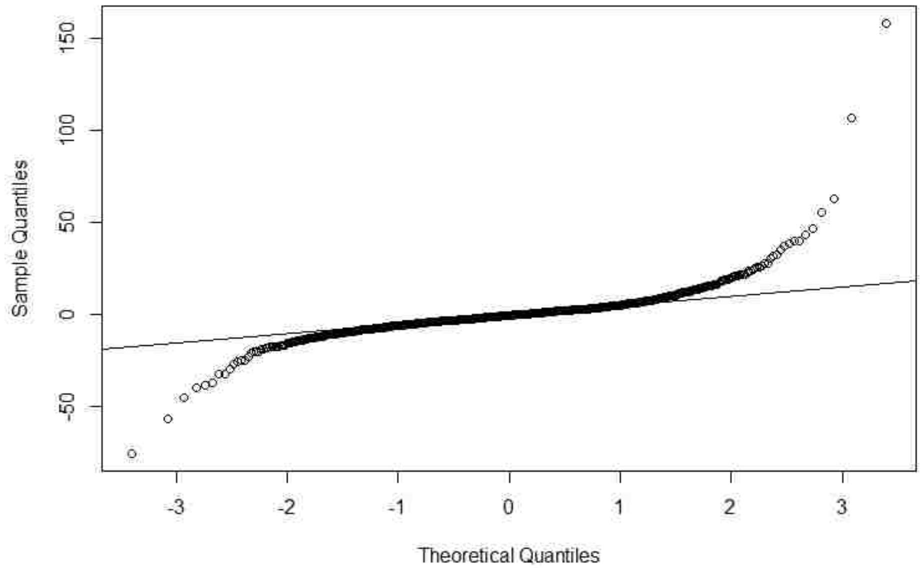


Figure 12 Normal Q-Q Plot of Cost Growth FDOT projects (N= 1,445)

Therefore, bootstrapping is conducted to check the significance of variables. As shown in Table 34, using bootstrapping method, zero is found outside of confidence interval for each predictor. Therefore, each remaining predictor in the model is significant. The t-values are also computed for testing whether beta is equal to 0 or not for each predictor. Since each t-value is greater than 2 and error degrees of freedom is 1442, each predictor in the model is highly significant. This shows that the non-normality data of CG will not affect the outcome of the regression model.

Table 33 Confidence Interval and T-Values from Bootstrapping for CG of FDOT projects

Independent Variables	CIbetaL	CIbetaU	T Observed
Intercept	- 4.02	- 2.26	- 5.93
Liquidated Damage Rate per Day	0.58	1.19	5.33
Contractor Type	- 3.67	-0.03	- 2.00

4.2.2.1.1.2 Multivariate Normality Tests of Residuals of Schedule Growth

Shapiro-Wilk test was used to check the normality of training dataset of SG. The p-value obtained from Shapiro-Wilk normality test was 2e-16, which is less than 0.05, so the residuals from the fitted linear model are non-normal. In addition, the Q-Q plot in Figure 13 shows the points does not lie on a straight

www.manaraa.com

diagonal line stating the data is not normally distributed.

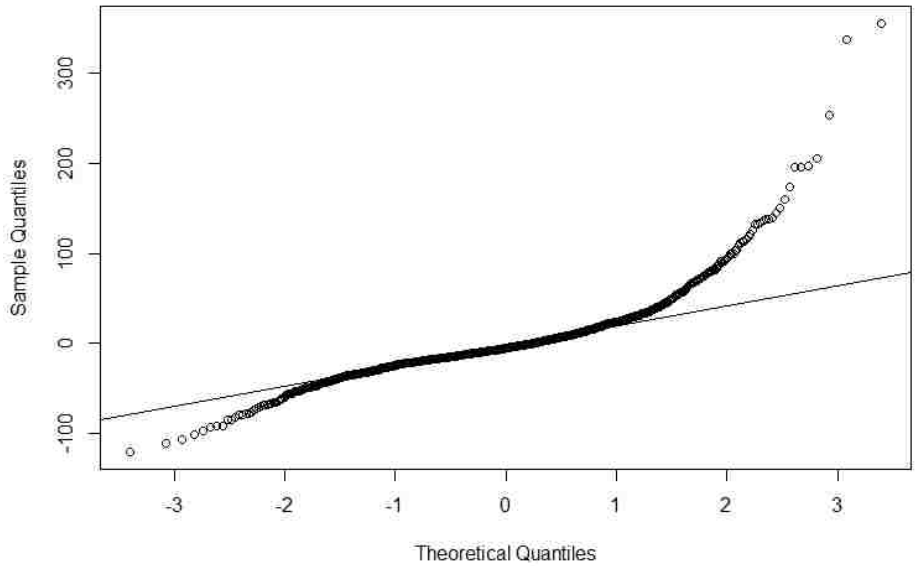


Figure 13 Normal Q-Q Plot of Schedule Growth of FDOT projects (N= 1,445)

Therefore, bootstrapping is conducted to check the significance of variables. As shown in Table 34, using bootstrapping method, zero is found outside of confidence interval for each predictor. Therefore, each remaining predictor in the model is significant. The t-values as shown below are computed for testing whether beta is equal to 0 or not for each predictor. Since each t-value is greater than 2 and error degrees of freedom is 1440, each predictor in the model is highly significant. This shows that the non-normality data of SG will not affect the outcome of the regression model.

Table 34 Confidence Interval and T-Values from Bootstrapping for SG of FDOT projects

Independent Variables	CIbetaL	CIbetaU	T Observed
Intercept	6.21	16.93	4.23
Award-Estimate Difference	3.34	13.56	3.24
Liquidated Damage Rate per Day	0.11	1.80	2.23
Contractor Type	- 15.19	- 2.73	- 2.82
Project Location	3.68	14.95	3.24

4.2.2.1.2 Linear Relationship Tests of Dependent Variables

A plot of the standardized residuals versus the predicted values of each outcome is created to

المنارة للاستشارات
www.manaraa.com

determine if linear relationship fits the model. The results of these tests for each of the dependent variables are described below.

4.2.2.1.2.1 Linear Relationship Tests of Cost Growth

The residual plots are required to validate the model. Figure 14 shows that the observed error (residuals) of CG is consistent with random and unpredictable errors. This shows that there is linear relationship existed between the CG and all the independent variables.

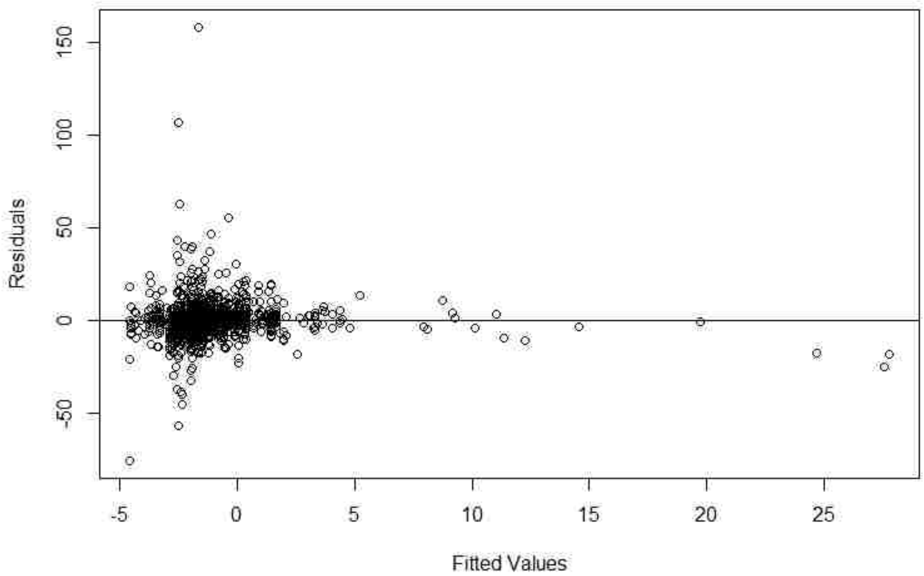


Figure 14 Residual vs Fitted Values - Cost Growth of FDOT projects (N= 1,445)

4.2.2.1.2.2 Linear Relationship Tests of Schedule Growth

The residual plots are required to validate the model. Figure 15 shows that the observed error (residuals) of SG is consistent with random and unpredictable errors. This shows that there is linear relationship existed between the SG and all the independent variables.

www.manaraa.com

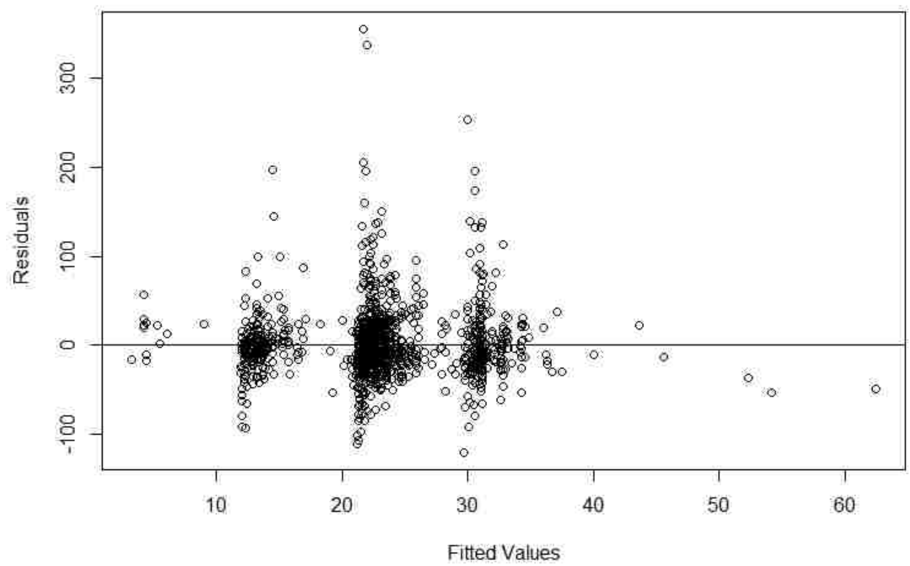


Figure 15 Residual vs Fitted Values - Schedule Growth of FDOT projects (N= 1,445)

4.2.2.1.3 Results of Multicollinearity Tests of Dependent Variables

Another assumption to conduct multiple linear regression is there should be no multicollinearity between the independent variables. The variance inflation factor (VIF) is greater than 5, multicollinearity existed between the independent variables.

4.2.2.2 Results of Regression Models

A multiple linear regression was used to find a relationship between each dependent variable with seven independent variables i.e. Number of Bidders, Award-Estimate Difference, Liquidated Damage Rate per Day, Contract Funding, Contractor Type, Project Location, and Project Size.

4.2.2.2.1 Multiple Linear Regression of Cost Growth

A multiple linear regression was used to determine the relationship between CG and seven independent variables. Out of the seven independent variables, two of them were found to have a significant correlation with CG as shown in Table 35. The R-square for the model is found to be 5%.

www.manaraa.com

Table 35 Significance Level of CG with Independent Variables in FDOT Projects

Independent Variables	Estimate	Std. Error	T-value	P-value	95% Conf. Interval		VIF
					CI upper	CI lower	
Intercept	- 3.14	0.36	- 8.67	2e-16**	- 3.86	- 2.43	
LD per Day	0.89	0.11	8.12	9e-16**	0.67	1.10	1.01
Contractor Type	- 1.79	0.94	- 1.89	0.05*	- 3.63	- 0.02	1.01

*significant at p-values less than 0.05   **significant at p-values less than 0.01

Number of Observations = 1,445
F (2, 1442) = 33.9
Prob. > F = 4e-15
R-squared = 0.045
Adjusted R-squared = 0.043

The multiple linear regression equation adopted for CG can be described mathematically as in

Cost Growth = - 3.14 + (0.89 * Liquidated Damage Rate per Day) - (1.79 * Contractor Type)

The regression equation for CG shows that:

- The coefficient for liquidated damage rate per day is significant, the sign is positive implying it is reasonable. The cost growth decrease by 0.89% when liquidated damage rate per day decrease by one thousand dollars. The one of the probable reason for this could be the projects had excusable and compensable delays which increased the final construction cost of the project. The liquidated damage rate per day does not apply when the delay was due to weather condition, an act of god, and unforeseen conditions. Also, another reason could be a comparative low compensation each day for delaying a project FDOT based on the size and daily spending of the project. The average LD per day was found around $2,300 for Florida projects. On the initial analysis of FDOT projects, it was found that FDOT charged LD per day to the contractor based upon the project size, however when the data was further examined, it was found that the LD per day charge changed randomly in the consecutive year. For example, FDOT charged $674 as LD per for project size up to $50,000 contract in 2002/ 2003 and 2004/2005. However, they changed this amount to $313 in 2006/2007 and $278 in 2008 to 2011.

المنارة للاستشارات
www.manaraa.com

- The coefficient for contractor type is significant, the sign is negative that means it is reasonable. The cost growth decrease by 1.79 % when the contractor is from out-of-state. The likely reason for this could be the out-of-state contractor would be more aggressive to complete the project within budget in comparison to in-state contractors to establish themselves in the state.
- The cost underrun will be 3.14% in a project when liquidated damage rate per day is zero; and contractor is local.

4.2.2.2.2 Multiple Linear Regression of Schedule Growth

A multiple linear regression was used to determine the relationship between SG and seven independent variables. Out of the seven independent variables, four of them were found to have a significant correlation with SG as shown Table 36. The R-square for the model is found to be 3%.

Table 36 Significance Level of SG with Independent Variables in FDOT Projects

Independent Variables	Estimate	Std. Error	T-value	P-value	95% Conf. Interval		VIF
					CI upper	CI lower	
Intercept	11.57	2.82	4.09	4e-05**	6.03	17.11	
Award-Estimate Difference	8.45	2.29	3.69	2e-04**	3.96	12.93	1.00
LD per Day	0.96	0.41	2.34	0.02*	0.15	1.75	1.01
Contractor Type	- 8.96	3.52	-2.55	0.01**	- 15.86	- 2.06	1.01
Project Location	9.32	2.88	3.24	1e-03**	3.67	14.96	1.01

*significant at p-values less than 0.05 **significant at p-values less than 0.01

Number of Observations = 1,445
F (4, 1440) = 9.42
Prob. > F = 1e-07
R-squared = 0.026
Adjusted R-squared = 0.023

The multiple linear regression equation adopted for SG can be described mathematically as in

Schedule Growth = 11.57 + (8.45 * Award-Estimate Difference) + (0.96* Liquidated Damage Rate per Day) - (8.96 * Contractor Type) + (9.32 * Project Location)

The regression equation for SG shows that:

- The coefficient for the difference between award and estimate is significant, the sign is positive

المنارة للاستشارات
www.manaraa.com

that indicates it is reasonable. The schedule growth increase by 8.45% when award amount is greater than engineer's estimated amount. There could be many reasons besides the positive deviation between the awarded amount and the estimated amount such as weather condition, geographic location, and change orders during the construction phase of a project which leads to the project delay. The study does not consider the quantitative analysis to find out by how much the positive or negative deviation between the awarded amount and the estimated amount affects the schedule growth and is recommended for future studies.

- The coefficient for liquidated damage rate per day is significant, the sign is positive implying it is reasonable. The schedule growth increase by 0.96% when liquidated damage rate per day increase by one thousand dollars. This could be because when the projects had a higher LD per day, the contractors were more driven to take the risk to complete the project on time. The resulting schedule pressure led to out-of-sequence work that delays the project. In addition, when the liquidated damage rate per day is lesser than the cost to accelerate per day during the construction of a project, the contractor prefers to delay the project. Another reason can be similar to that mentioned in cost growth findings regarding inconsistency in providing the LD per day by FDOT.
- The coefficient for contractor type is significant, the sign is negative that shows it is reasonable. The schedule growth decrease by 8.96% when the contractor is from out-of-state. The possible reasons can be due to aggressive nature of out-of-state contractors to establish themselves in the state in comparison to in-state contractors.
- The coefficient for project location is significant, the sign is positive that indicates it is reasonable. The schedule growth will increase by 9.32% when project location is urban. This shows that the rural projects were completed earlier than urban projects. The delay could be due to the high

www.manaraa.com

complexity of the urban projects compared to rural projects. As the projects became more complex, there was a higher chance of experiencing a schedule delay.

- The project will be behind the schedule by 11.57% when the award-estimate difference is negative; liquidated damage rate per day is zero; contractor is local; and project location is rural.

4.2.2.3 Results of Data Validation

The testing data is used to validate the models. The validation is done for all the seven models. The results of the validation are described below.

4.2.2.3.1 Validation of Cost Growth Model

The regression equation determined was used to calculate the CG for a testing dataset. The testing dataset was the 361 transportation projects that were randomly selected. The difference between the actual and predicted CG values was calculated to check the accuracy of the model. The summary and box plot of the difference between predicted and observed values are shown in Table 37 and Fig. C - 2 respectively. It shows that mean difference is 0.87% and the minimum and maximum difference values were found to be -28.58% and 51.11%.

Table 37 Summary of Diff. between Predicted and Observed CG Values of FDOT Projects (N=361)

Cost Growth	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
Predicted - Observed	- 28.58	- 3.10	1.03	0.87	4.30	51.11

The numbers of projects whose predicted CG values are within +/- 20% with actual values were calculated. It was found that 353 projects were within defined range. Therefore, the percentage of the project with ± 20% CG error is 97.78%. In addition, the absolute average difference in CG between predicted and observed is found to be 5.34%.

4.2.2.3.2 Validation of Schedule Growth Model

The regression equation determined was used to calculate the SG for a testing dataset. The testing dataset was the 361 completed transportation projects. The difference between the actual and predicted

المنارة للاستشارات
www.manaraa.com

SG values was calculated to check the accuracy of the model. The summary and box plot of the difference between predicted and observed values are shown in Table 38 and Fig. C - 4 respectively. It shows that mean difference is -0.77% and the minimum and maximum difference values were found to be -382.90% and 97.87%.

Table 38 Summary of Diff. between Predicted and Observed SG Values of FDOT Projects (N=361)

Schedule Growth	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
Predicted - Observed	- 382.90	- 14.28	6.32	- 0.77	19.60	97.87

The numbers of projects whose predicted SG values are within +/- 20% with actual values were calculated. It was found that 195 projects were within defined range. Therefore, the percentage of the project with ± 20% SG error is 54.02%. In addition, the absolute average difference in SG between predicted and observed is found to be 26.13%

4.3 Results of Combined State DOT Projects

The multiple linear regression was conducted for each performance metric. The developed models were validated using independent set of the data by checking the variation between observed and predicted values of each dependent variable.

4.3.1 Results of Multiple Linear Regression for Combined State DOTs Projects

The total number of projects used for creating a training model is 2,754 projects. In addition, 691 projects were used for validation purpose. The results of this multiple linear regression and validation are discussed below.

4.3.1.1 Results of Regression Models

A multiple linear regression was used to find a relationship between each dependent variable with seven independent variables i.e. Number of Bidders, Award-Estimate Difference, Liquidated Damage Rate per Day, Contract Funding, Contractor Type, Project Location, and Project Size.

المنارة للاستشارات
www.manaraa.com

4.3.1.1.1 Multiple Linear Regression of Cost Growth

A multiple linear regression was used to determine the relationship between CG and seven independent variables. Out of the seven independent variables, four of them were found to have a significant correlation with CG as shown in Table 39. The R-square for the model is found to be 2%.

Table 39 Significance Level of CG with Independent Variables in Combined DOT Projects

Independent Variables	Estimate	Std. Error	T-value	P-value	95% Conf. Interval		VIF
					CI upper	CI lower	
Intercept	3.17	0.72	4.42	1e-05**	1.76	4.57	
Project Size	0.14e-06	1e-08	7.09	1e-12**	9e-08	1e-07	1.01
Contract Funding	- 1.18	0.60	-1.95	0.05*	- 2.36	0.05	1.01
Contractor Type	- 2.60	0.96	-2.72	6e-03**	- 4.46	- 0.70	1.00
Project Location	- 1.20	0.56	-2.13	0.03*	- 2.29	- 0.09	1.01

*significant at p-values less than 0.05 **significant at p-values less than 0.01

Number of Observations = 2,754
F (4, 2749) = 15.3
Prob. > F = 2e-12
R-squared = 0.022
Adjusted R-squared = 0.020

The multiple linear regression equation adopted for CG can be described mathematically as in

Cost Growth = 3.17 + (0.14e-06 * Project Size) - (1.18 * Contract Funding) - (2.60 * Contractor Type) - (1.20 * Project Location)

The regression equation shows that CG will:

- increase by 0.14% when project size increase by 1 million dollars,
- decrease 1.18% when contract funding is federal,
- decrease by 2.60% when contractor is from out-of-state,
- decrease by 1.20% when project location is urban,
- be 3.17% cost overrun in a project when project size is zero; contract funding is state; contractor is local; and project location is rural.

المنارة للاستشارات
www.manaraa.com

4.3.1.1.2 Multiple Linear Regression of Schedule Growth

A multiple linear regression was used to determine the relationship between SG and seven independent variables. Out of the seven independent variables, five of them were found to have a significant correlation with SG as shown Table 40. The R-square for the model is found to be 3%.

Table 40 Significance Level of SG with Independent Variables in Combined DOT Projects

Independent Variables	**Estimate**	**Std. Error**	**T-value**	**P-value**	**95% Conf. Interval**		**VIF**
					CI upper	**CI lower**	
Intercept	2.29	2.50	0.91	0.36	- 2.61	7.19	
Award-Estimate Difference	6.38	1.74	3.67	2e-04**	2.96	9.79	1.00
Liquidated Damage Rate per Day	2.03	0.43	4.76	2e-06**	1.19	2.87	1.03
Contract Funding	4.09	2.00	2.04	0.04*	0.16	8.02	1.01
Contractor Type	- 7.54	3.17	-2.38	0.01**	- 13.75	- 1.32	1.01
Project Location	9.88	1.87	5.28	1e-07**	6.21	13.55	1.03

*significant at p-values less than 0.05 **significant at p-values less than 0.01

Number of Observations = 2,754
F (5, 2748) = 15.07
Prob. > F = 1e-14
R-squared = 0.027
Adjusted R-squared = 0.024

The multiple linear regression equation adopted for SG can be described mathematically as in

Schedule Growth = 2.29 + (6.38 * Award-Estimate Difference) + (2.03 * Liquidated Damage Rate per Day) + (4.09 * Contract Funding) - (7.54 * Contractor Type) + (9.88 * Project Location)

The regression equation shows that SG will:

- increase by 6.38% when award amount is greater than engineer’s estimated amount,
- increase by 2.03% when liquidated damage rate per day increase by one thousand dollars,
- increase by 4.09% when contract funding is federal,
- decrease by 7.54% when the contractor is from out-of-state,
- increase by 9.88% when project location is urban,
- be 2.29% behind of the schedule when all the independent variables, the award-estimate difference is negative; liquidated damage rate per day is zero; contract funding is state; contractor

المنارة للاستشارات
www.manaraa.com

is local; and project location is rural.

4.3.1.2 Results of Data Validation

The testing data is used to validate the models. The validation is done for all the seven models. The results of the validation are described below.

4.3.1.2.1 Validation of Cost Growth Model

The regression equation determined was used to calculate the CG for a testing dataset. The testing dataset was the 691 transportation projects that were randomly selected. The difference between the actual and predicted CG values was calculated to check the accuracy of the model. The summary of the difference between predicted and observed values are shown in Table 41. It shows that mean difference is 0.63% and the minimum and maximum difference values were found to be - 55.78% and 79.39%.

Table 41 Summary of Diff. between Predicted and Observed CG Values of Combined Projects (N=691)

Cost Growth	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
Predicted - Observed	- 55.78	- 3.34	1.59	0.63	5.73	79.39

The numbers of projects whose predicted CG values are within +/- 20% with actual values were calculated. It was found that 653 projects were within defined range. Therefore, the percentage of the project with ± 20% CG error is 94.50%. In addition, the absolute average difference in CG between predicted and observed is found to be 6.92%.

4.3.1.2.2 Validation of Schedule Growth Model

The regression equation determined was used to calculate the SG for a testing dataset. The testing dataset was the 691 completed transportation projects. The difference between the actual and predicted SG values was calculated to check the accuracy of the model. The summary of the difference between predicted and observed values are shown in Table 42. It shows that mean difference is 0.09% and the minimum and maximum difference values were found to be - 386.51% and 94.43%.

Table 42 Summary of Diff. between Predicted and Observed SG Values of Combined Projects (N=691)

المنارة للاستشارات
www.manaraa.com

Schedule Growth	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
Predicted - Observed	- 386.51	- 15.12	6.06	0.09	21.81	94.43

The numbers of projects whose predicted SG values are within +/- 20% with actual values were calculated. It was found that 357 projects were within defined range. Therefore, the percentage of the project with ± 20% SG error is 51.66%. In addition, the absolute average difference in SG between predicted and observed is found to be 26.94%.

4.4 Results of Comparison of Developed Models for Separate and Combined State DOT Projects

Chow test is mainly conducted to see whether the true coefficients in two linear regressions on different data sets are equal. In this study, the chow test was conducted for both cost and schedule performance metrics among three datasets (TxDOT, FDOT, and Combined State DOTs). The models were developed using all seven independent variables for each dataset. Firstly, the results of chow test for the cost growth model were presented. Secondly, schedule growth model outputs were described. The separate and pool equations as shown below was used to determine whether the combined or separate model for each performance model is better for future application.

Equation 1: $Y_i = \beta_1 + \beta_2 X_{2i} + \ldots\ldots + \beta_k X_{ki} + \varepsilon_I$, i= 1, …, N

Equation 2: $Y_j = \alpha_1 + \alpha_2 X_{2j} + \ldots\ldots + \alpha_k X_{kj} + \varepsilon_I$, j= 1, …, N

Pooled Equation 3: $Y_i = \beta_1 + \beta_2 X_{2i} + \ldots\ldots + \beta_k X_{ki} + \varepsilon_I$, i= 1, …, N+M

Null Hypothesis (Ho): $\alpha_1 = \beta_1$ and $\alpha_2 = \beta_2$ and…….. $\alpha_k = \beta_k$

Alternative Hypothesis (H_A): $\alpha_1 \neq \beta_1$ and $\alpha_2 \neq \beta_2$ and…….. $\alpha_k \neq \beta_k$

The F statistics was used to check whether the combined or separate model is better using formula as shown in Equation 8. If the critical value of F test is less than obtained F value from Equation 9, then null hypothesis is rejected stating that the separate models are better than combined model.

المنارة للاستشارات
www.manaraa.com

$$\frac{ESS_R - ESS_{UR} \,/\, \#\, of\ restriction}{ESS_{UR} \,/\, \#\, d\ of UR} \approx F_{\#n\#d} \qquad (8)$$

$$\frac{(ESS_3 - (ESS_1 + ESS_2))/\ k}{(ESS_1 + ESS_2)/\ (N + M - 2k)} \approx F_{k,\ N+M-2k} \qquad (9)$$

4.4.1 Chow Test for Cost Growth

The error sum of squares and degree of freedom found using all seven independent variables while predicting cost growth using Texas, Florida, and Combined State data are shown in Table 43.

Table 43 Error Sum of Squares and Degree of Freedom of CG Models using Three Datasets

	Texas	**Florida**	**Combined**
Error Sum of Squares	258482.5	141515	431738.7
Degree of Freedom	1301	1437	2746

Now, using F-Statistics formula,

$$\frac{(431738.7 - (258482.5 + 141515))/\ 8}{(258482.5 + 141515)/\ (1301 + 1437 - 2 * 8)} \approx F_{8,\ 1301+1437-2*8}$$

$$27.00 \approx F_{8,\ 2722}$$

The results showed that $F_{critical}$ (1.09) < $F_{8,2722}$(27.00) which shows that the pool test rejects the null hypothesis. Therefore, the separate model for each state is suggested in comparison to combined model for predicting cost growth in a project.

4.4.2 Chow Test for Schedule Growth

The error sum of squares and degree of freedom found using all seven independent variables while predicting schedule growth using Texas, Florida, and Combined State data are shown in Table 44.

Table 44 Error Sum of Squares and Degree of Freedom of SG Models using Three Datasets

	Texas	**Florida**	**Combined**
Error Sum of Squares	2672237	1968365	4721757
Degree of Freedom	1301	1437	2746

المنارة للاستشارات
www.manaraa.com

Now, using F-Statistics formula,

$$\frac{(4721757 - (2672237 + 1968365))/\,8}{(2672237 + 1968365)/\,(1301 + 1437 - 2 * 8)} \approx F_{8,\ 1301+1437-2*8}$$

$$5.95 \approx F_{8,\ 2722}$$

The results showed that $F_{critical}$ (1.09) < $F_{8,2722}$(5.95) which shows that the pool test rejects the null hypothesis. Therefore, the separate model for each state is suggested in comparison to combined model for schedule growth in a project.

المنارة للاستشارات
www.manaraa.com

CHAPTER 5 DISCUSSION

The statistical analysis showed that cost and schedule performance metrics had a correlation with contract procurement data. In this study, the combined effect of seven contract procurement data on project performances was determined. In addition, the study results showed that these project performances could be predicted using the contract procurement data. The findings of this study of TxDOT and FDOT data revealed there was some procurement factor that had a correlation with cost and schedule metrics and was common in both cases. The similarities and the differences between results in these two datasets are described below.

5.1 Findings of Cost Growth

In both DOT studies for Texas and Florida, procurement factors affected cost. Table 45 shows the major findings regarding the effect of procurement factors on cost growth metrics. The cost growth was positively affected by the amount of liquidated damage rate per day set by their state DOTs. The results indicated that if the liquidated damage rate increased, the cost growth also increased. The study made by Akinci and Fischer (1998) had mentioned that contract clauses such as a liquidated damages clause heighten the risks of cost overrun in a project. This study reinforced the statement that by increasing the liquidated damage rate in the contract, it is more likely to have a cost overrun in a project. In addition, Shrestha et al. (2012) found a positive correlation, though not significant, between the liquidated damage rate per day and the cost growth. For TxDOT projects, the cost overrun decreased if the project was funded by federal money. Kishore and Abraham (2009) found the cost of constructing highways with federal funds was higher than the state-funded highway projects. Additionally, this study found that the cost overrun was higher when a project was in urban areas. Rosmond (1984) supported this finding. Their study found that the higher change orders caused cost overruns in an urban project. For FDOT projects, the cost overrun decreased when the selected contractor was from out-of-state.

www.manaraa.com

When the dataset for both state DOTs projects was combined for analysis, the cost growth was found significantly affected by all the factors listed in Table 45 except liquidated damage rate per day. In addition, the cost growth was affected by the project size. The cost growth increased as the project size increased. The probable reason for this could be as the project size increases the complexity of a project also increases. The cost growth decreased when the contract is federally funded and selected contractor from out-of-state. Moreover, the cost growth increased when the project is in an urban area.

Table 45 Significant Procurement Factors affecting Cost Growth

Procurement Factors	**Significant Relationships**		
	Texas DOT	**Florida DOT**	**Combined DOT**
Project Size	None	None	Positive
LD per Day	Positive	Positive	None
Contract Funding Type	Less for federal projects	None	Less for federal projects
Contractor Type	None	Less for out-of-state	Less for out-of-state
Project Location	More for urban projects	None	Less for urban projects

5.2 Findings of Schedule Growth

Procurement factors significantly affected the schedule growth as seen in both TxDOT and FDOT studies. Table 46 illustrates the major findings regarding the effects of procurement factors on schedule growth metrics. Not only the difference between the awarded amount and the engineer's estimate but also the amount of liquidated damage rate per day positively affected schedule growth in both state DOT projects. The schedule growth is found when the awarded amount is greater than the estimated amount. The liquidated damage rate per day increased as the schedule growth increased (Table 46). This finding is similar to that of Shrestha et al.(2012). The authors found that there was a positive correlation between the schedule growth and the liquidated damage rate per day on projects. However, the correlation was not statistically significant. Another finding is the schedule growth was less when the project was in rural areas for both state DOT projects, which shows that the rural projects were completed earlier than urban projects. The state-funded projects had less schedule growth in comparison to federally funded projects

المنارة للاستشارات
www.manaraa.com

in Texas DOT. Another finding is the schedule growth reduces if the selected contractor was from out-of-state in Florida DOT projects.

When the dataset for both state DOTs projects was combined for analysis, the schedule growth was found significantly affected by all the factors listed in Table 46. The schedule growth was found when the bid amount was higher than the estimated amount, the project was federally funded, and located in urban area. In addition, as the liquidated damage rate per day of a project increased, the schedule growth also increased. Moreover, the out-of-state contractors had schedule savings in a project in comparison to in-state contractors.

Table 46 Significant Procurement Factors affecting Schedule Growth

Procurement Factors	**Significant Relationships**		
	Texas DOT	**Florida DOT**	**Combined DOT**
Award-Estimate Difference	More when bid > estimated	More when bid > estimate	More when bid > estimate
LD per Day	Positive	Positive	Positive
Contract Funding Type	More for federal projects	None	More for federal projects
Contractor Type	None	Less when out-of-state	Less when out-of-state
Project Location	More for urban projects	More for urban projects	More for urban projects

www.manaraa.com

CHAPTER 6 CONCLUSION AND RECOMMENDATIONS

6.1 Conclusions

The research focused on determining the relationships between contract procurement factors and performance metrics of a transportation project built following the design-bid-build project delivery method. A comprehensive questionnaire was developed to collect those data from state DOTs. The research hypothesis for this study assumes there is relationship between cost growth and schedule growth and the seven contract procurement factors available in the study. The study aimed to collect five state DOTs data for the analysis, however, due to a lack of available data, only TxDOT and FDOT were analyzed for this study. The multiple linear regression showed that the linear relationships among these procurement factors and cost and schedule metrics depended on the individual state's DOT based on the standard error comparison of separate and pool models.

The study proved that only four procurement factors shared a significant relationship with the cost growth of any given project based on individual state analysis. Those procurement factors include the liquidated damage rate per day, the type of a contract funding, the type of a contractor, and the project location. However, the liquidated damage rate per day was the common factor that had a similar positive effect on both state DOT cost growth values. In TxDOT, the type of a contract funding and the project location were found to have a significant relationship with the cost growth. In contrast, the type of a contractor was found to have a significant relationship with the cost growth of FDOT projects. When the data from both state DOT projects were combined, a significant relationship was found between the cost growth and all factors except the liquidated damage rate per day found on individual state analysis. In addition, the project size was found to have a significant relationship with cost growth.

The multiple linear regression analysis showed that five procurement factors had a significant relationship with the schedule growth. In this case, the award-estimate difference, liquidated damage

المنارة للاستشارات
www.manaraa.com

rate, and project location had significant positive correlation with the schedule growth. However, rest of the two procurement factors were not common in TxDOT and FDOT project data. In TxDOT projects, type of project funding was identified as a significant procurement factor that affected the schedule growth. Nevertheless, in FDOT data, type of contractor was found to be a significant factor affecting the schedule growth. When the data from both state DOT projects were combined, a significant relationship was found between the schedule growth and all factors found on individual state analysis.

The analysis of these two state DOT showed that the procurement factors that affect the cost and schedule performance varied by DOTs. The variations on the significant factors may be due to a unique way of contracting in each state DOTs. However, there were some procurement factors identified which were common in both DOTs that significantly affect the project performance. As this study is exploratory in nature, more state DOTs data analysis need to be conducted to reach the concrete conclusions.

The primary contribution of this dissertation to the body of knowledge is that this study identified the combined effect of seven procurement factors on the cost growth and the schedule growth. Previous studies had not incorporated all these procurement factors in the study and those studies determined the individual effect of some of these procurement factors on cost and schedule performances. This study also collected two state DOTs data and compared the difference in findings from these two state models. It also created a single model from the data collected from two state DOTs and identified the difference between these two models. These types of work had not been conducted in previous research. Therefore, this study will pave new research opportunities in predicting the project performance based on the project procurement factors.

6.2 Limitation / Recommendation

The major limitations of this study are related to the results of the regression model. The regression model has very low R^2 but significant within 95% confidence level. The low R-square

المنارة للاستشارات
www.manaraa.com

value is generally attributed due to a large number of sample sizes. Outliers can dramatically alter the relationship of performance metrics with contract procurement data in the regression equation. The outliers were not removed because it will reduce the number of samples. It is further recommended to conduct the similar study by collecting more state DOTs data and determine the difference between findings among these DOTs. The author would also like to recommend collecting more procurement factors and performance metrics data to conduct comprehensive analysis.

www.manaraa.com

APPENDIX A BOXPLOT OF QUANTITATIVE INDEPENDENT AND DEPENDENT VARIABLES

www.manaraa.com

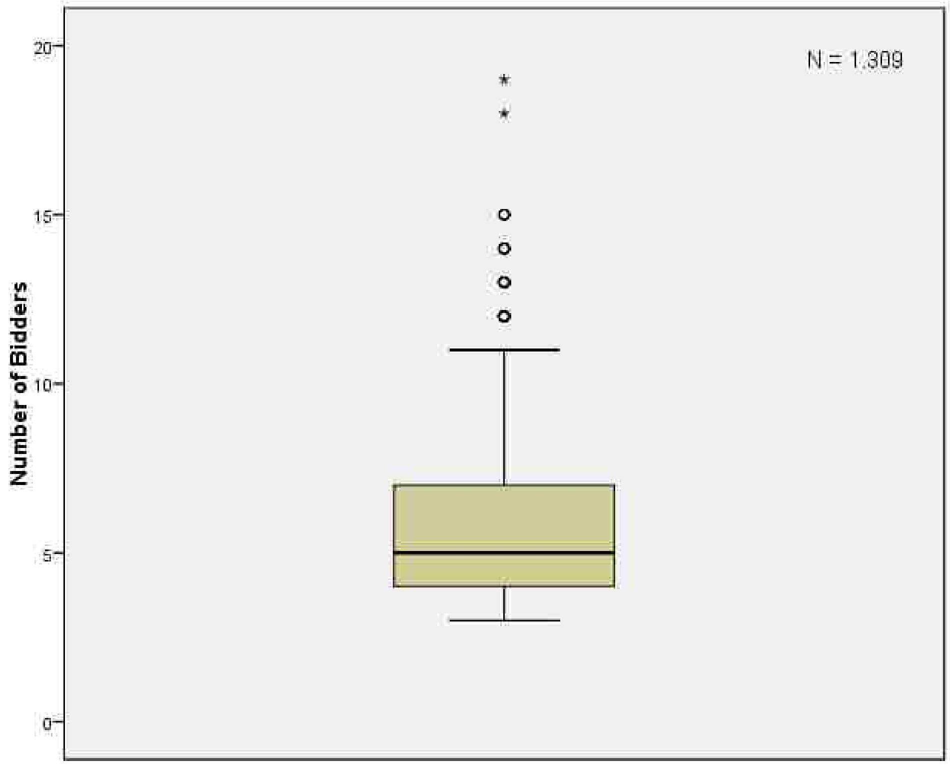


Fig. A - 1 Boxplot of Number of Bidders of TxDOT projects

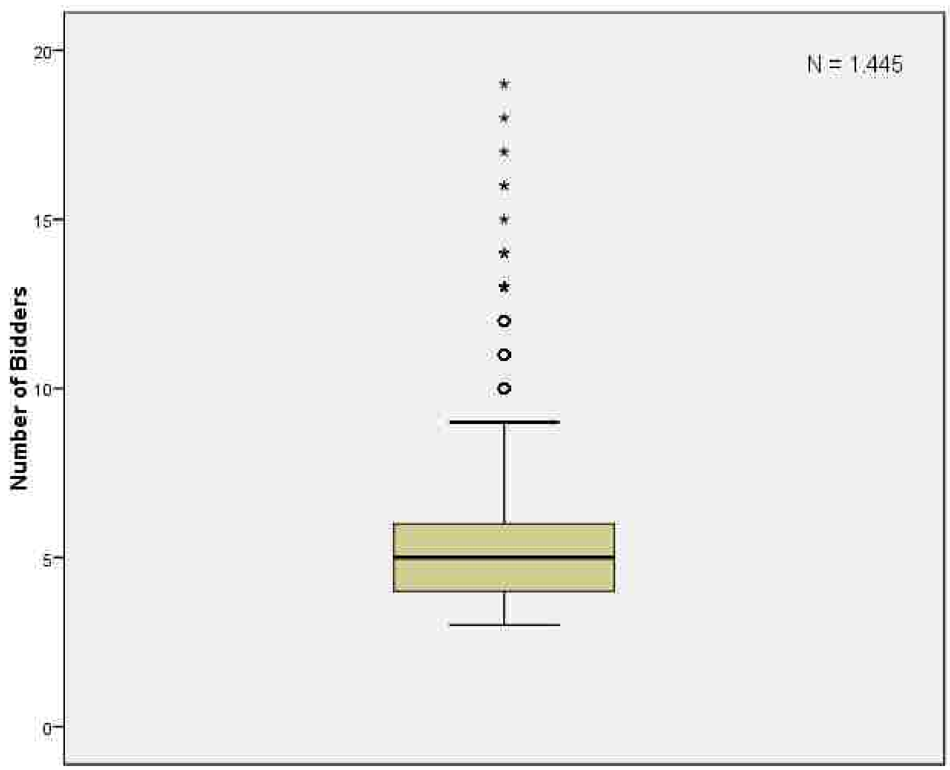


Fig. A - 2 Boxplot of Number of Bidders of FDOT Projects

www.manaraa.com

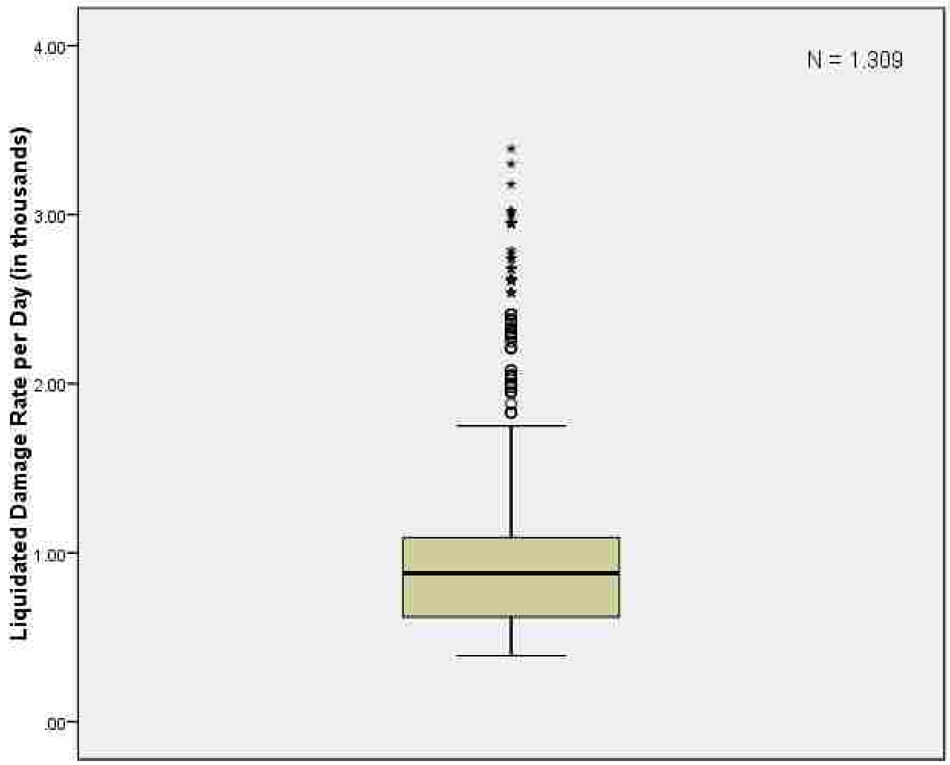


Fig. A - 3 Boxplot of Liquidated Damage Rate per Day of TxDOT Projects

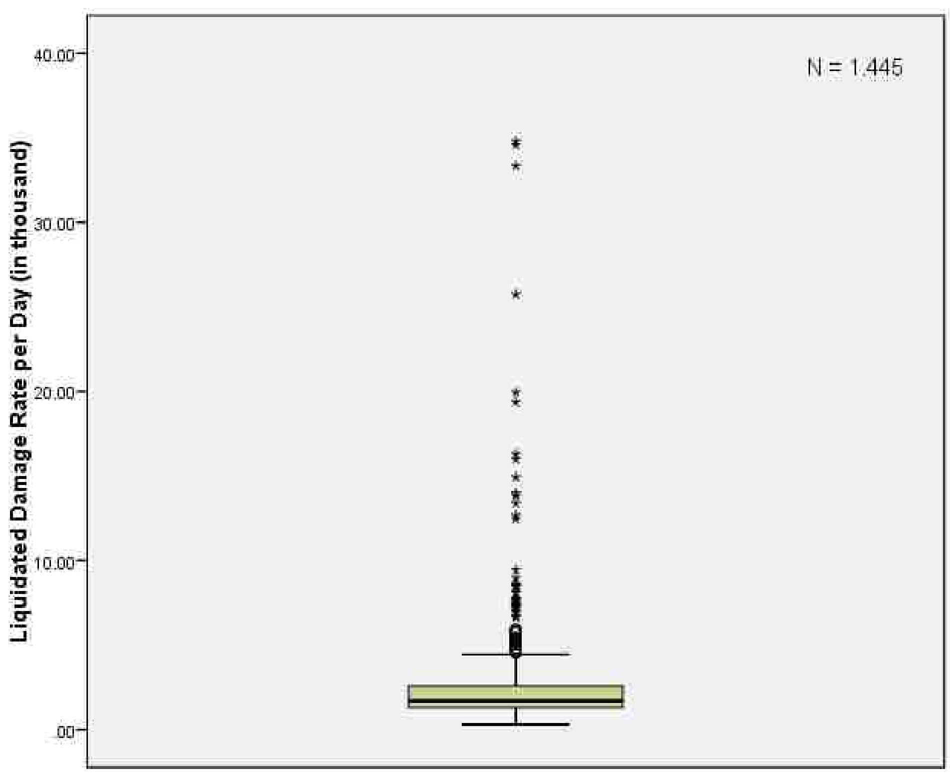


Fig. A - 4 Boxplot of Liquidated Damage Rate per Day of FDOT Projects

www.manaraa.com

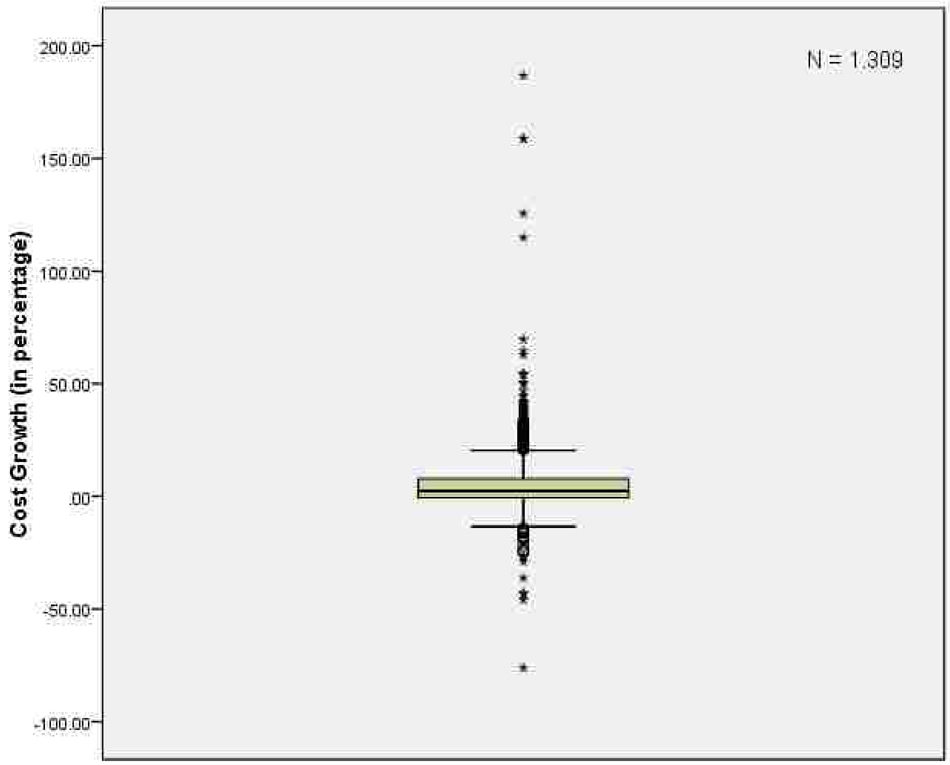


Fig. A - 5 Boxplot of Cost Growth of TxDOT Projects

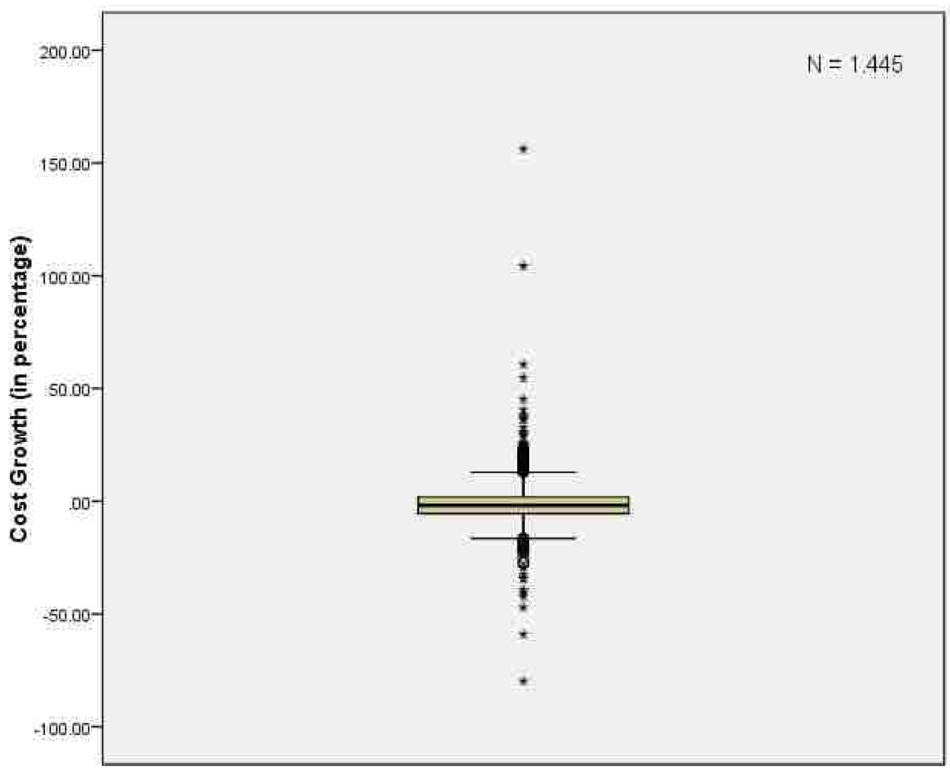


Fig. A - 6 Boxplot of Cost Growth of FDOT Projects

المنارة للاستشارات
www.manaraa.com

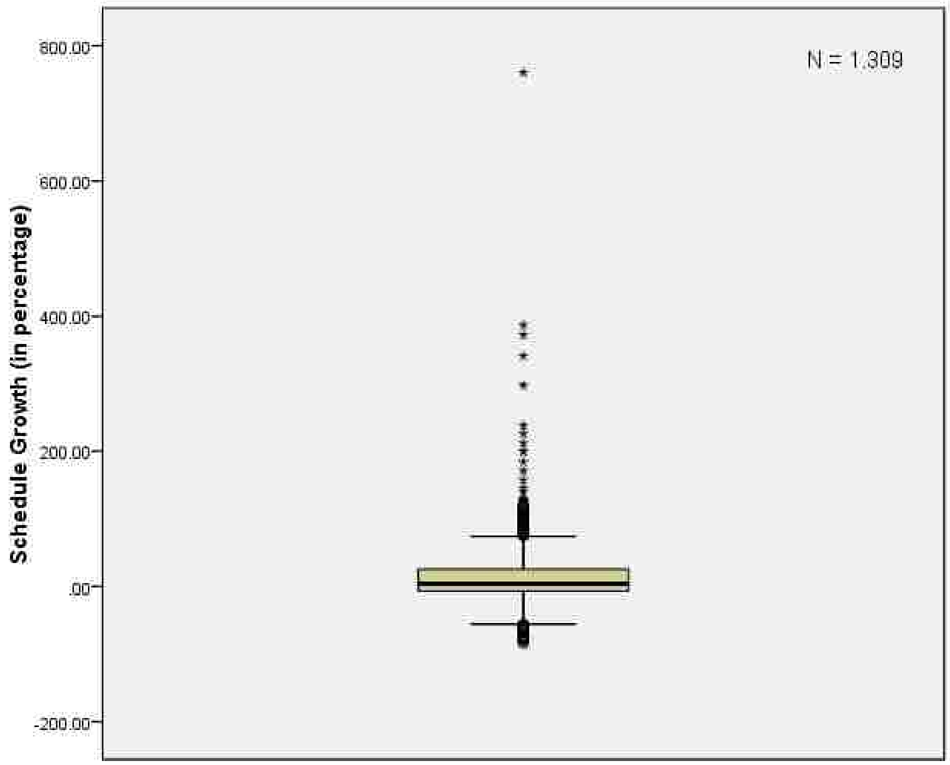


Fig. A - 7 Boxplot of Schedule Growth of TxDOT Projects

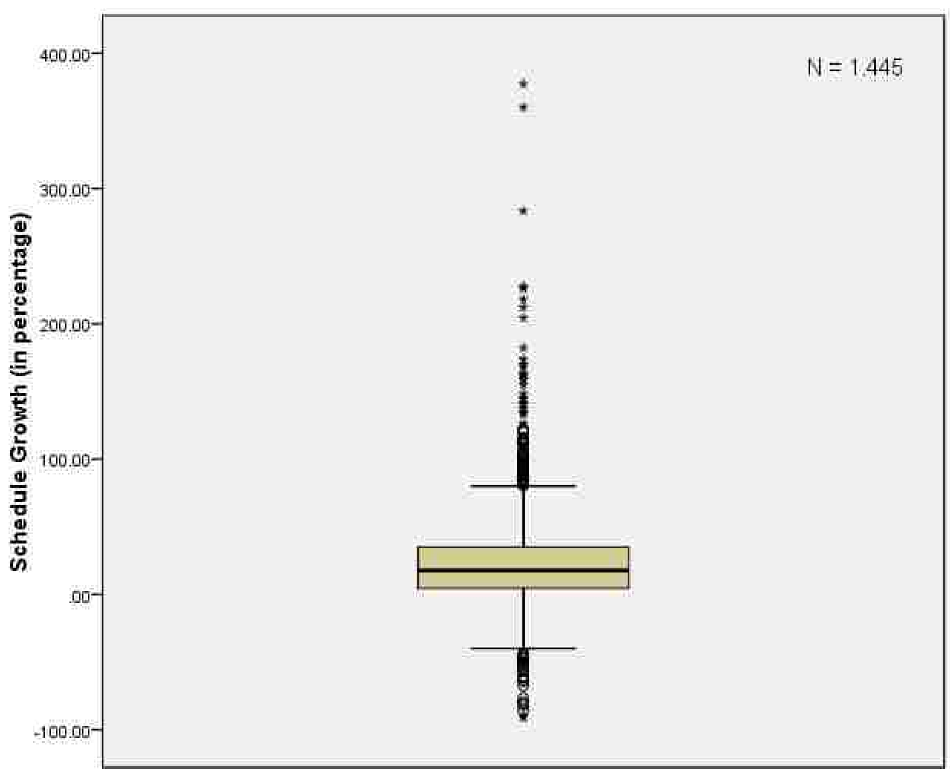


Fig. A - 8 Boxplot of Schedule Growth of FDOT Projects

المنارة للاستشارات
www.manaraa.com

APPENDIX B HISTOGRAM OF QUANTITATIVE INDEPENDENT AND DEPENDENT VARIABLES

المنارة للاستشارات

www.manaraa.com

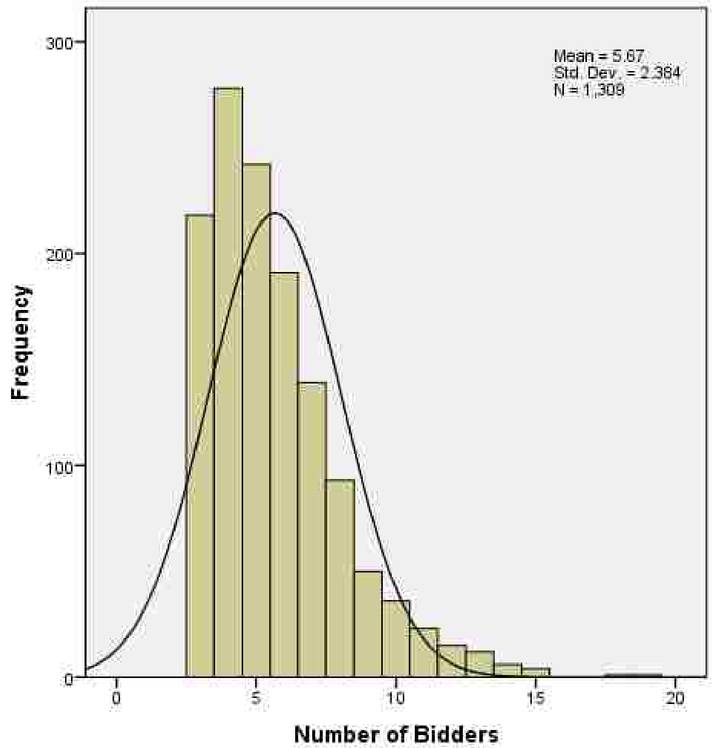


Fig. B - 1 Histogram of Number of Bidders of TxDOT projects (N = 1,309)

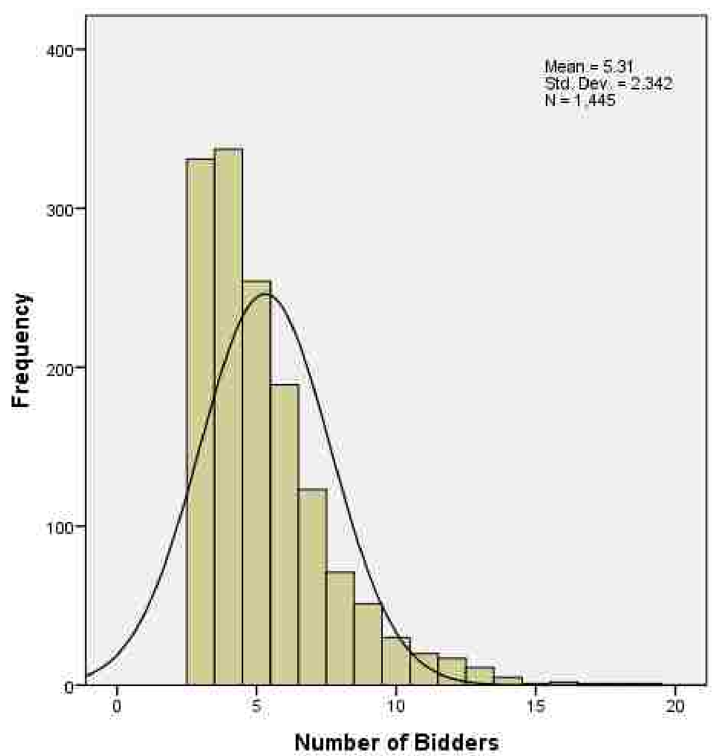


Fig. B - 2 Histogram of Number of Bidders of FDOT projects (N = 1,445)

المنارة للاستشارات
www.manaraa.com

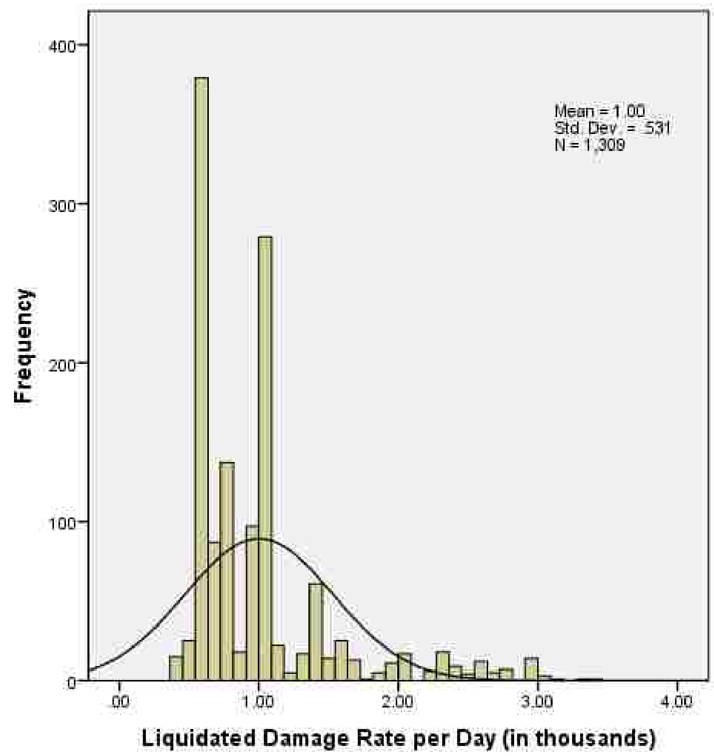


Fig. B - 3 Histogram of Liquidated Damage Rate per Day of TxDOT projects (N = 1,309)

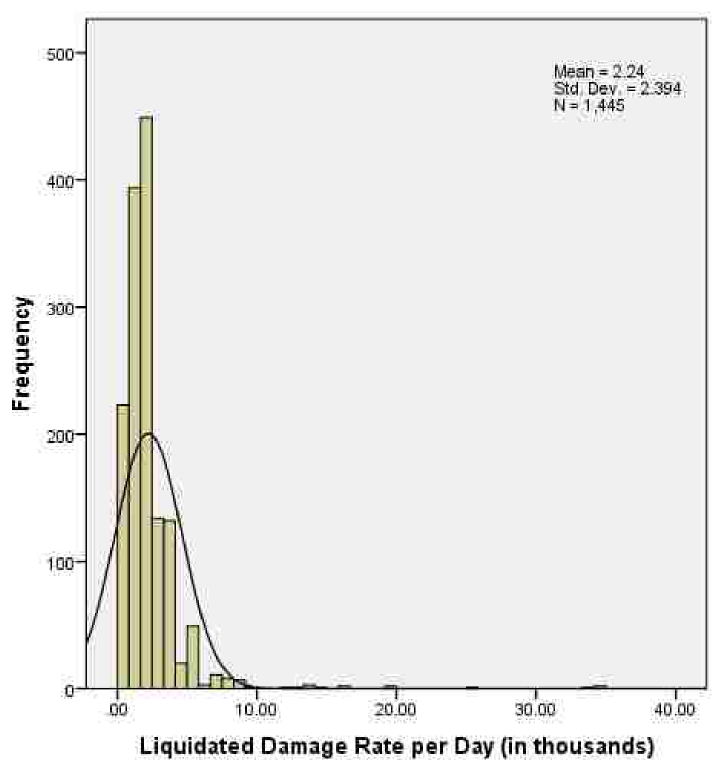


Fig. B - 4 Histogram of Liquidated Damage Rate per Day of FDOT projects (N = 1,445)

للاستشارات المنارة
www.manaraa.com

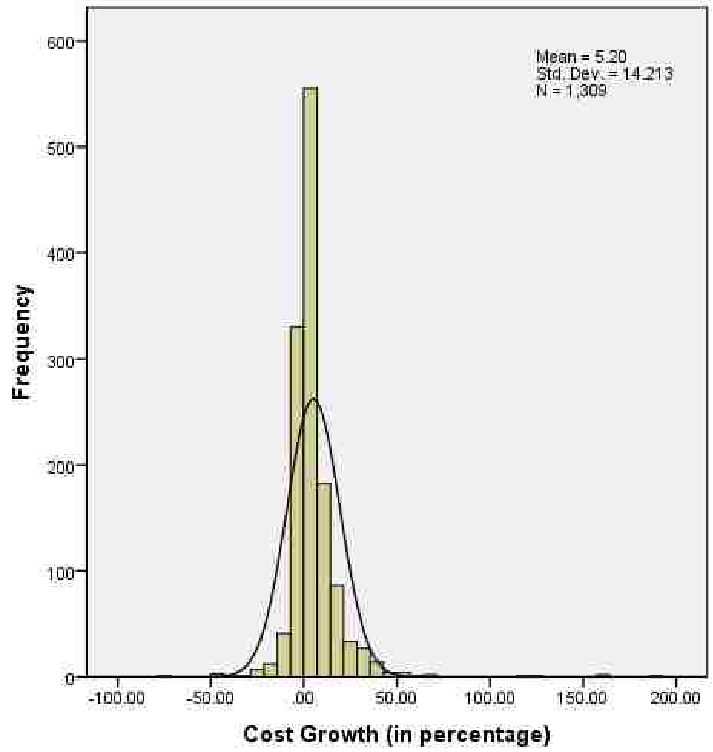


Fig. B - 5 Histogram of Cost Growth of TxDOT projects (N = 1,309)

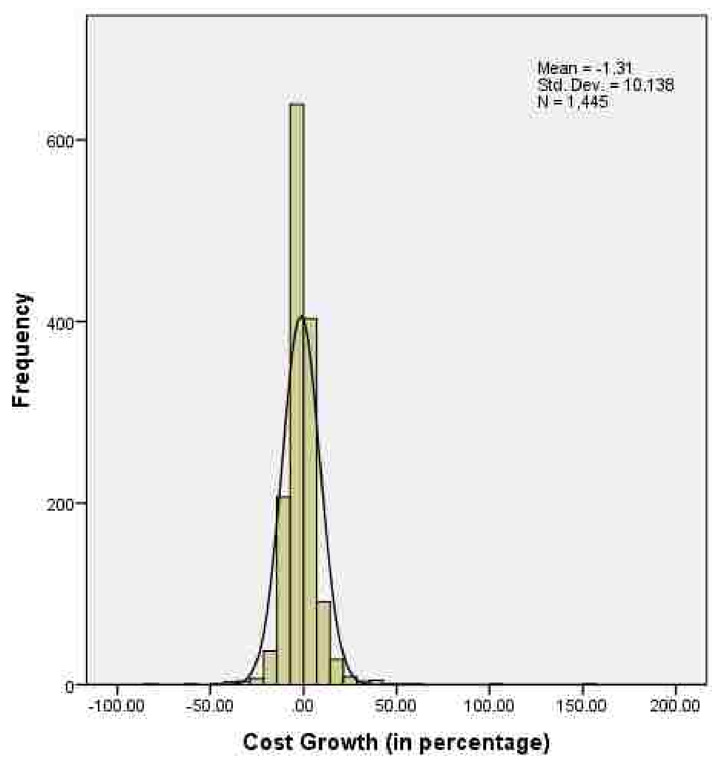


Fig. B - 6 Histogram of Cost Growth of FDOT projects (N = 1,445)

المنارة للاستشارات
www.manaraa.com

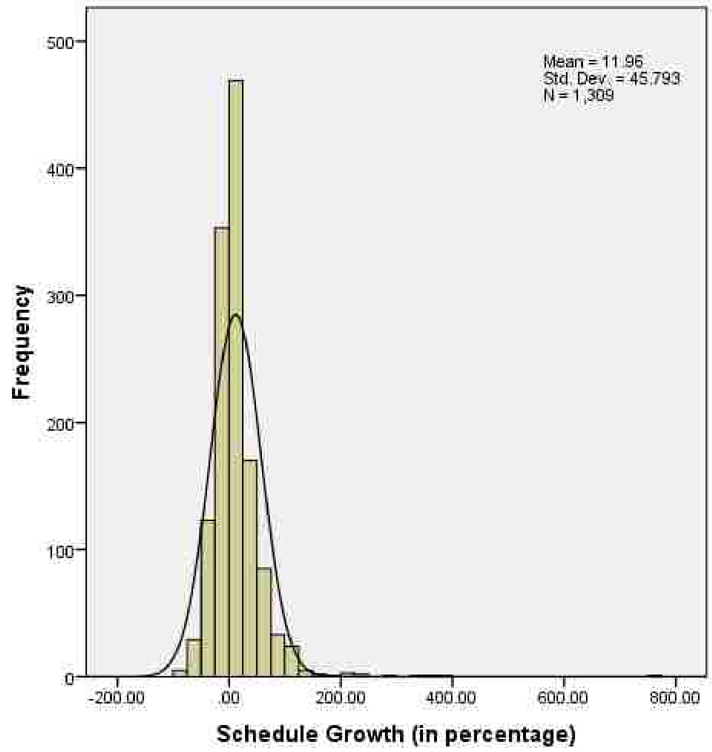


Fig. B - 7 Histogram of Schedule Growth of TxDOT projects (N = 1,309)

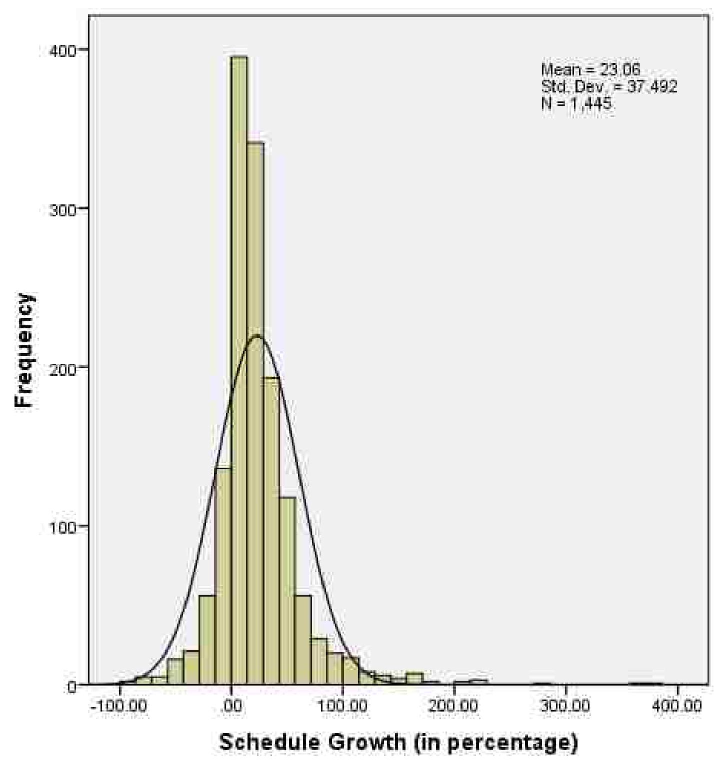


Fig. B - 8 Histogram of Schedule Growth of FDOT projects (N = 1,445)

للاستشارات المنارة
www.manaraa.com

APPENDIX C BOXPLOT OF DIFF. BET. PREDICTED AND OBSERVED DEPENDENT VALUES

www.manaraa.com

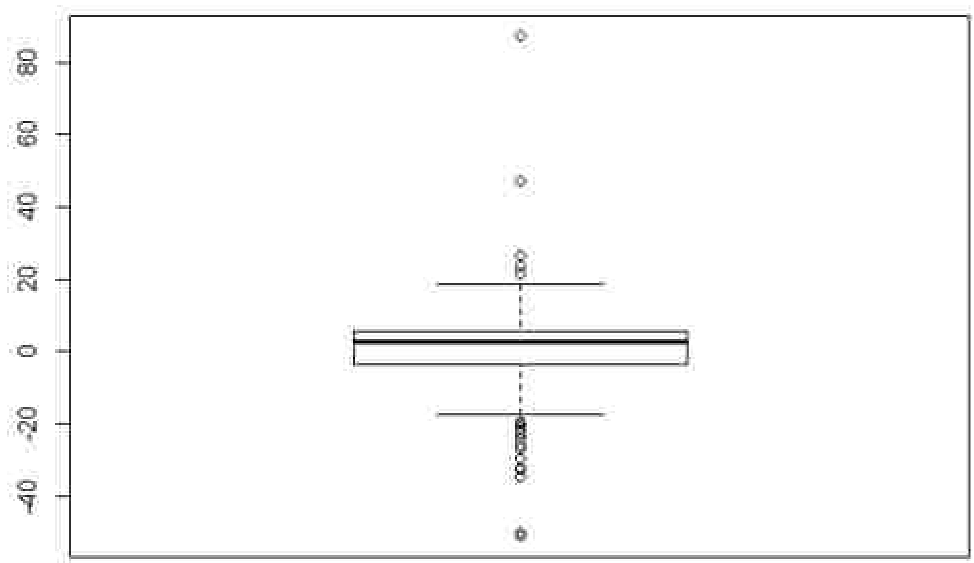


Fig. C - 1 Boxplot of Difference between Predicted and Observed CG values of TxDOT projects (N=330)

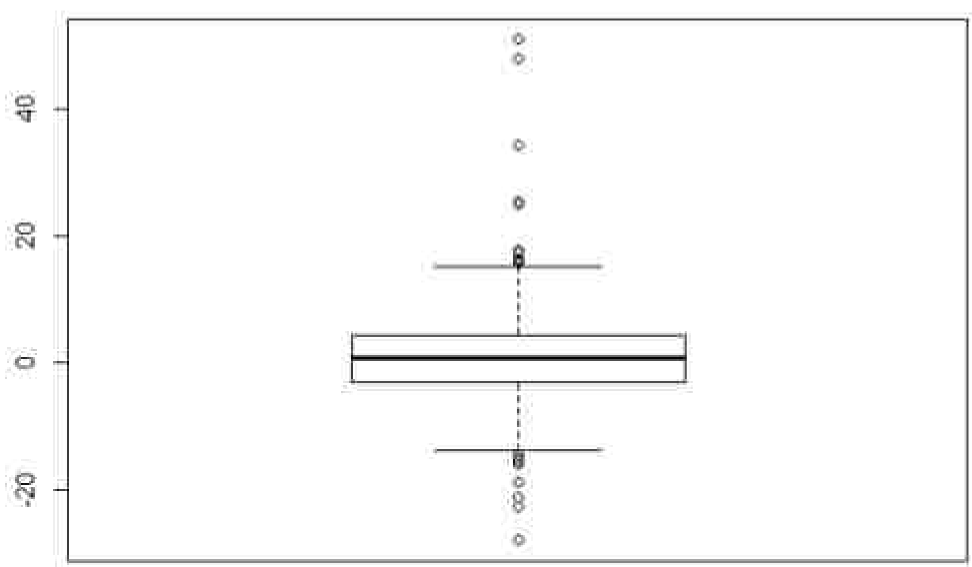


Fig. C - 2 Boxplot of Difference between Predicted and Observed CG values of FDOT projects (N=361)

المنارة للاستشارات
www.manaraa.com

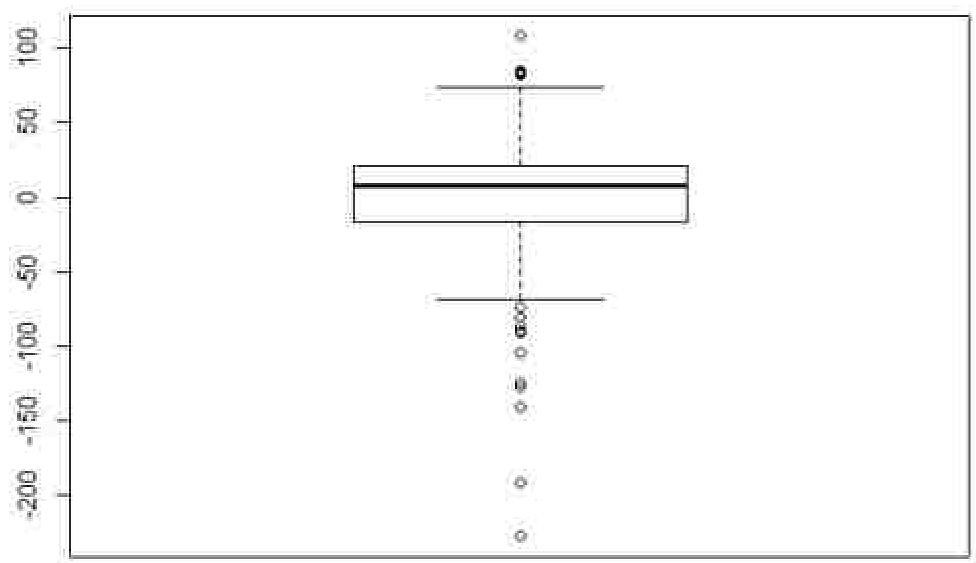


Fig. C - 3 Boxplot of Difference between Predicted and Observed SG Values of TxDOT projects (N=330)

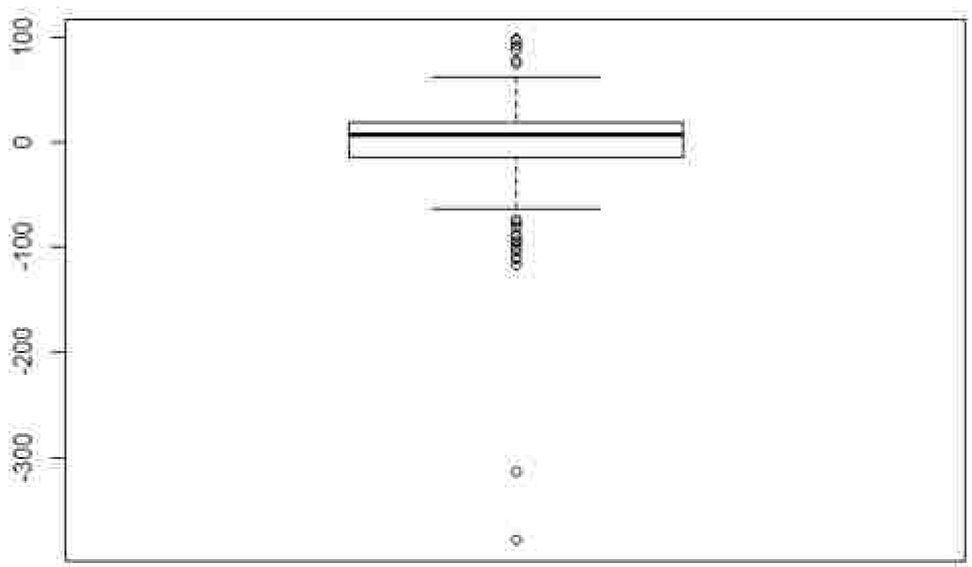


Fig. C - 4 Boxplot of Difference between Predicted and Observed SG values of FDOT projects (N=361)

المنارة للاستشارات
www.manaraa.com

REFERENCES

Ahiaga-Dagbui, D. D., & Smith, S. D. (2014). Rethinking construction cost overruns: cognition, learning, and estimation. *Journal of Financial Management of Property and Construction*, *19*(1), 38–54.

Ahmad, I., & Minkarah, I. (1988). Questionaire survey on bidding in construction. *Journal of Management in Engineering*, *4*(3), 229–243.

Akinci, B., & Fischer, M. (1998). Factors affecting contractors' risk, *14*(February), 67–76.

An, S.-H., Kim, G.-H., & Kang, K.-I. (2007). A case-based reasoning cost estimating model using experience by analytic hierarchy process. *Building and Environment*, *42*(7), 2573–2579. https://doi.org/10.1016/j.buildenv.2006.06.007

Assaf, S. A., & Al-Hejji, S. (2006). Causes of delay in large construction projects. *International Journal of Project Management*, *24*(4), 349–357. https://doi.org/10.1016/j.ijproman.2005.11.010

Bageis, A. S., & Fortune, C. (2009). Factors affecting the bid/no bid decision in the Saudi Arabian construction contractors. *Construction Managment and Economics*, *27*(1), 53–71.

Bordat, C., McCullouch, B. G., Labi, S., & Sinha, K. (2004). An analysis of cost overruns and time delays of INDOT projects. *Transportation Research*, (December), 193. https://doi.org/10.5703/1288284313134

Carr, P. G. (2005). Investigation of bid price competition measured through pre-bid project estimates, actual bid prices, and Number of bidders. *Journal of Construction Engineering and Management*, *131*(11), 1165–1172. https://doi.org/10.1061/(ASCE)0733-9364(2005)131:11(1165)

Carr, R. I. (1983). Impact of number of bidders on competition. *Journal of Construction Engineering and Management*, *109*(1), 61–73. https://doi.org/10.1061/(ASCE)0733-9364(1983)109:1(61)

Carr, R. I. (1987). Competitive bidding and opportunity costs. *Journal of Construction Engineering and Management*, *113*(1), 151–165.

المنارة للاستشارات
www.manaraa.com

Chan, A. P., & Chan, A. P. (2004). Key performance indicators for measuring construction success. *Benchmarking: An International Journal*, *11*(2), 203–221.

Chan, D. W., & Kumaraswamy, M. M. (1995). A study of the factors affecting construction durations in Hong Kong. *Construction Management and Economics*, *13*(4), 319–333.

Chou, J. S. (2009). Generalized linear model-based expert system for estimating the cost of transportation projects. *Expert Systems with Applications*, *36*, 4253–4267. https://doi.org/10.1016/j.eswa.2008.03.017

Chua, D. K. H. (2000). Key factors in bid reasoning model. *Journal of Construction Engineering and Management*, *126*(5), 349–357.

Chua, D. K., & Hossain, M. A. (2012). Predicting change propagation and impact on design schedule due to external changes. *IEEE Transactions on Engineering Management*, *59*(3), 483–493. https://doi.org/10.1109/TEM.2011.2164082

Demographics, T. and F. (2016). Texas and Florida counties by population.

Edwards, C., & Kaeding, N. (2015). Federal government cost overruns. Retrieved from https://www.downsizinggovernment.org/government-cost-overruns

Frimpong, Y., Oluwoye, J., & Crawford, L. (2003). Causes of delay and cost overruns in construction of groundwater projects in developing countries; Ghana as a case study. *International Journal of Project Management*, *21*(5), 321–326. https://doi.org/10.1016/S0263-7863(02)00055-8

Green, S. B. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research*. https://doi.org/10.1207/s15327906mbr2603

Harbuck, R. H. (2004). Competitive bidding for highway construction projects. *AACE International Transactions*.

Hinze, J., & Selestead, A. G. (1991). *Analysis of WSDOT construction cost overruns*.

المنارة للاستشارات

www.manaraa.com

Ho, S. P., & Hsu, Y. (2013). Bid compensation theory and strategies for projects with heterogeneous bidders: A game theoretic analysis. *Journal of Management in Engineering*, *30*(5), 4014022. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000212

Iyer, K. C., & Jha, K. N. (2006). Critical factors affecting schedule performance: Evidence from Indian construction projects. *Journal of Construction Engineering and Management*, *132*(8), 871–881. https://doi.org/10.1061/(ASCE)0733-9364(2006)132

Jahren, C. T., & Ashe, A. M. (1990). Predictors of cost-overrun rates. *Journal of Construction Engineering and Management*, *116*(3), 548–552.

Jennings, W. (2012). Why costs overrun: risk, optimism and uncertainty in budgeting for the London 2012 Olympic Games. *Construction Management and Economics*, *30*(6), 455–462. https://doi.org/10.1080/01446193.2012.668200

Ji, S. H., Park, M., & Lee, H. S. (2010). Data preprocessing-based parametric cost model for building projects: Case studies of Korean construction projects. *Journal of Construction Engineering and Management*, *136*(8), 844–853. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000197

Kagioglou, M., Cooper, R., & Aouad, G. (2001). Performance management in construction: a conceptual framework. *Construction Management and Economics*, *19*(1), 85–95. https://doi.org/10.1080/01446190010003425

Kaming, P. F., Olomolaiye, P. O., Holt, G. D., & Harris, F. C. (1997). Factors influencing construction time and cost overruns on high-rise projects in Indonesia. *Construction Management and Economics*, *15*(1), 83–94. https://doi.org/10.1080/014461997373132

Kim, B. S. (2011). The approximate cost estimating model for railway bridge project in the planning phase using CBR method. *KSCE Journal of Civil Engineering*, *15*(7), 1149–1159. https://doi.org/10.1007/s12205-011-1342-2

المنارة للاستشارات

www.manaraa.com

Kim, G.-H., An, S.-H., & Kang, K.-I. (2004). Comparison of construction cost estimating models based on regression analysis, neural networks, and case-based reasoning. *Building and Environment*, *39*(10), 1235–1242. https://doi.org/10.1016/j.buildenv.2004.02.013

Kim, M., Lee, S., Woo, S., & Shin, D. H. (2012). Approximate cost estimating model for river facility construction based on case-based reasoning with genetic algorithms. *KSCE Journal of Civil Engineering*, *16*(3), 283–292. https://doi.org/10.1007/s12205-012-1482-z

Kishore, V., & Abraham, D. M. (2009). Construction costs - using federal vs. local funds. *Highway*, (July). https://doi.org/10.5703/1288284314310.This

Knofczynski, G. T., & Mundfrom, D. (2007). Sample sizes when using multiple linear regression for prediction. *Educational and Psychological Measurement*, *68*(3), 431–442. https://doi.org/10.1177/0013164407310131

Lai, K. K., Liu, S. L., & Wang, S. Y. (2004). A method used for evaluating bids in the Chinese construction industry. *International Journal of Project Management*, *22*(3), 193–201. https://doi.org/10.1016/S0263-7863(03)00009-7

Leśniak, A., & Plebankiewicz, E. (2013). Modeling of the decision-making process concerning participation in the construction bidding. *Journal of Management in Engineering*, *31*(2). https://doi.org/10.1061/(ASCE)ME.1943-5479.0000237

Li, S., Foulger, J. R., & Philips, P. W. (2008). Analysis of the impacts of the number of bidders upon bid values: Implications for contractor prequalification and project timing and bundling. *Public Works Management & Policy*, *12*(3), 503–514. https://doi.org/10.1177/1087724X07312144

LTAP. (2005). *Indiana county highway departments bridge Replacement cost estimation procedures*.

Mansfield, N. R., Ugwu, O. O., & Doran, T. (1994). Causes of delay and cost overruns in Nigerian construction projects. *International Journal of Project Management*, *12*(4), 254–260.

www.manaraa.com

https://doi.org/10.1016/0263-7863(94)90050-7

Martin, J., Burrows, T. K., & Pegg, I. (2006). Predicting construction duration of building projects. *In XXIII Congreso FIG, October 2006*.

Meeampol, S., & Ogunlan, S. O. (2006). Factors affecting cost and time performance on highway construction projects: Evidence from Thailand. *Journal of Financial Management of Property and Construction*, *11*(1), 3–20. https://doi.org/10.1108/13664380680001076

Memon, A. H., Rahman, I. A., & Azis, A. A. A. (2011). Preliminary study on causative factors leading to construction cost overrun. *International Journal of Sustainable Construction Engineering & Technology*, *2*(1).

Memon, A. H., Rahman, I. A., & Azis, A. A. A. (2012). Time and cost performance in construction projects in southern and central regions of Peninsular Malaysia. *International Journal of Advances in Applied Sciences*, *1*(1), 45–52. https://doi.org/10.1109/CHUSER.2012.6504280

Ngai, S. C., Drew, D. S., Lo, H. P., & Skitmore, M. (2002). A theoretical framework for determining the minimum number of bidders in construction bidding competitions. *Construction Management and Economics*, *20*(6), 473–482. https://doi.org/10.1080/01446190210151041

Nkado, R. N. (1995). Construction time-influencing factors: the contractor's perspective. *Construction Management and Economics*, *13*(1), 81–89. https://doi.org/10.1080/01446199500000009

Patrick, D., Straus, J., & Archer, J. (2016). *Definitions of " Rural " in Texas statutes and the Texas administrative code*.

Randolph, D. A., Rajendra, K., & Campfield, J. J. (1987). Using risk management techniques to control contract cost. *Journal of Management in Engineering*, *3*(4), 314–324.

Runeson, G., & Skitmore, M. (1999). Tendering theory revisited. *Construction Management and Economics*, *17*(3), 285–296. https://doi.org/10.1080/014461999371493

المنارة للاستشارات

www.manaraa.com

Salem Hiyassat, M. A. (2001). Construction bid price evaluation. *Canadian Journal of Civil Engineering*, *28*(2), 264–270. https://doi.org/10.1139/cjce-28-2-264

Shash, A. A. (1993). Factors considered in tendering decisions by top UK contractors. *Construction Management and Economics*, *11*(2), 111–118. https://doi.org/10.1080/01446199300000004

Shrestha, P. P., O'Connor, J. T., & Gibson, G. E. (2012). Performance Comparison of Large Design-Build and Design-Bid-Build Highway Projects, *138*(January), 1–13. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000390.

Shrestha, P., & Pradhananga, N. (2010). Correlating bid price with the number of bidders and final construction cost of public street projects. *Journal of the Transportation Research Board*, (2151), 3–10. https://doi.org/10.3141/2151-01

Statistics, L. (2017). Multiple regression analysis using SPSS statistics. Retrieved from https://statistics.laerd.com/spss-tutorials/multiple-regression-using-spss-statistics.php

Tumi, S. A. H., Omran, A., & Pakir, A. H. K. (2009). Causes of Delay in Construction Industry in Libya. *The International Conference on Economics and Administration*, 265–272.

Williams, T. P. (2003). Predicting final cost for competitively bid construction projects using regression models. *International Journal of Project Management*, *21*(8), 593–599. https://doi.org/10.1016/S0263-7863(03)00004-8

WSDOT. (2002). *Highway construction cost comparison survey*.

Yu, I., Kim, K., Jung, Y., & Chin, S. (2007). Comparable performance measurement system for construction companies, *23*(3), 131–139.

Yu, W. D., Wang, K. W., & Wang, M. T. (2012). Pricing strategy for best value tender. *Journal of Construction Engineering and Management*, *139*(6), 675–684. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000635

المنارة للاستشارات
www.manaraa.com

CURRICULUM VITAE

Ruiko Maharjan

ruikomaharjan@gmail.com

Education:

Master of Science in Civil Engineering, December 2013

Bachelor degree in Civil Engineering, March 2011

Associate degree in Computer Engineering, May 2006

Journal Publications

- Shrestha, P. P., Maharjan, R., Batista, J. R., & Shakya, B. (2016). Comparison of Utility Managers' and Project Managers' Satisfaction Rating of Alternative Project Delivery Methods Used in Water and Wastewater Infrastructures. *Public Works Management & Policy*, *21*(3), 263-279. (Published)
- Shrestha, P. P., Maharjan, R., Batista, J. R. (2017). Analysis of Design-Build and Construction Manager-at-Risk in Water and Wastewater Projects. *ASCE Journal of Professional Issues in Engineering Education and Practice* (Under review)
- Maharjan, R., & Shrestha, P. P., (2017). Correlating Change Orders with Cost and Schedule Growth of Highway Projects. *ASCE Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*. (Paper under review)

Conference Publications

- Shrestha, P. P., Maharjan, R., Shakya, B., & Batista, J. (2014). Alternative project delivery methods for water and wastewater projects: the satisfaction level of owners. *Proceedings of ASCE Construction Research Congress 2014: Construction in a Global Network*, Atlanta, May 19-21, 2014 pp. 1733-1742). (Published)
- Shrestha, P. P., Batista, J., & Maharjan, R. (2016). Risks Involved in Using Alternative Project Delivery (APD) Methods in Water and Wastewater Projects. *Procedia Engineering*, Proceedings of International Conference on Sustainable Design, Engineering, and Construction, Tempe, AZ, May 18-20, 2016, 145, 219-223. (Published)
- Shrestha, P. P., Batista, J., & Maharjan, R. (2016). Impediments in Using Design-Build or

المنارة للاستشارات

www.manaraa.com

Construction Management-at-Risk Delivery Methods for Water and Wastewater Projects. *Proceedings of ASCE Construction Research Congress*, San Juan, Puerto Rico, May 31-June 2, 2016, pp. 380-387. (Published)

- Shrestha, P. P., Shakya, B., & Maharjan, R. (2016). Performance of Design-Build versus Construction Manager/ General Contractor for Highway Projects. *Proceedings of 95th Transportation Research Board*. Washington, D.C., January 10-13, 2016. (Presented)
- Maharjan, R., & Shrestha, P. P., (2018). Relationship between Project Performance and Contract Procurement Factors for Design-Bid-Build Highway Projects. *Construction Research Congress 2018: Construction in a Global Network*. Baton Rouge, LA, April 2-5, 2018. (Published)

Dissertation Title:

Effects of Contract Procurement Factors on Performance of Transportation Projects

Examination Committee:

Pramen P. Shrestha, Ph.D. (Committee Member - Chairperson)

Jin Ouk Choi, Ph.D. (Committee Member)

Mohamed Kaseko, Ph.D. (Committee Member)

Hualiang Teng, Ph.D. (Committee Member)

Ashok Singh, Ph.D. (Graduate Faculty Representative)

www.manaraa.com