

University of Bahrain

College of Engineering

Department of Mechanical Engineering



Development of Cost Estimation Model for Ministry of Youth and Sports Affairs Construction Projects

A Case study from Kingdom of Bahrain

**A Thesis submitted in partial fulfillment of the requirements for the Master degree
in Engineering Management**

Submitted by:

Fatema Saqer Yusuf Malalla

ID number: 20083399

Supervised by:

Prof. Saad M. A. Suliman

(Professor)

Co-supervisor:

Dr. Osama Al Jamal

(Assistant of Professor)

Kingdom of Bahrain

October 2021



Thesis Discussion Committee Decision

Student Name: Fatema Saqer Yusuf Malalla	Student ID: 20083399
Department: Mechanical Engineering	College: Engineering
Program: Masters in Engineering Management	
Thesis Title: Development of Cost Estimation Model for Ministry of Youth and Sports Affairs Construction Projects "A Case study from Kingdom of Bahrain"	
Supervisor Name: Professor Saad M. A. Suliman, Dr. Osama Al Jamal	Student GPA: 3.93

The Discussion Committee met on 06-Oct-21. In MS Teams. Upon discussing the student whose details are stated above in public, the discussion committee recommends the following:

The student is awarded a grade A in the thesis subject.

- ☐ Without making any changes to the thesis.
- ☒ After making the modifications attached with this decision.

Discussion Committee Members Approval:

<u>Name</u>		<u>Signature</u>
1. Prof. S. Suliman & Dr. O. Al-Jamal	(Advisor)	<i>Saad Suliman / Osama Al-Jamal</i>
2. Prof. Mukesh Kumar	(Internal Examiner)	<i>Mukesh Kumar</i>
3. Dr. Khalid Al-Gahtani	(External Examiner)	<i>Khalid Al-Gahtani</i>

Abstract

Cost estimation is one of the crucial tasks handled in the project initiation phase and has a great effect on project performance. Ministry of Youth and sports (MYS) in Kingdom of Bahrain had indicators that support the possibility of having construction cost estimation deficiency problems. The problem to be investigated in this research is the inaccuracy of cost estimate of MYS construction projects. The main objectives of this research are to assess the level of cost estimation accuracy in MYS construction projects, identify the factors affecting the current cost estimation in MYS projects and develop a cost estimation model that enhances the cost estimation process. The objectives of the study are achieved through a comprehensive literature review about the three main issues: level of cost estimation accuracy, factors affecting cost estimation accuracy and developing a model for cost estimation. The data collected to achieve these objectives are actual construction projects data executed by MYS and questionnaire survey information collected from stakeholders who worked with MYS. The cost estimation model developed using two combined techniques: partial least square structural equation model (PLS-SEM) and artificial neural network (ANN). The application of PLS-SEM on the critical factors determined by the participants provided theoretical support for the study conceptual framework. While ANN modeling identified the influential factors affecting MYS construction cost estimation. This research provides an insight into the current cost estimation level in MYS, which provides opportunities to improve the cost estimation during the projects design phase. Also, the relationships between the research variables: availability of cost estimation unit; information quality; estimator performance; project characteristics; external factors; and contractual procedure on the construction cost estimation accuracy have been confirmed. The study suggests the need to control the highest-ranked factors: information quality, estimator performance and project characteristics to enhance the cost estimation process efficiency.

Dedication

I would like to dedicate this work to my backbone, my parents and family, for their continuous love, encouragement and support during this journey.

Also, this work will not reach to an end without the contribution of special people who provide me with their knowledge and support:

My uncle Essa Malalla, Mrs. Wafa Alammadi, Eng. Khalid Alhaj, Eng. Saleh Al Othman, Mrs. Mona Ebrahim, Mrs. Anwaar Albuarki, Mrs. Fatema Al Doy, Mr. Nadeem Parker, Mrs. Zainab Al-Awadi, Dr. Thuraya Jumaa, Mr. Redha Alhammam, Mr. Ali Awad, Mr. Mahmood Al-jaber, Eng. Shaikha Al-khulasi, Eng. Sahar Falamarzi, Eng. Hissa Alkandri and Mr. Jalal Al Majid.

This is dedicated to all of you.

Thank you for being effective part of this work.

Acknowledgement

I would like to express my gratitude to my family who supported me in my post graduate journey.

I would like to show my greatest appreciation to my supervisor Prof. Saad Suliman and co-supervisor Dr. Osama Al-Jamal for their continuous guidance in achieving this research.

List of Content

Subject	Page
Accreditation by Jury	A
Abstract	B
Dedication	C
Acknowledgement	D
List of Contents	E
List of Tables	H
List of Figures	L
List of Abbreviation	N
List of Notations	P
Glossary	R
List of Appendices	T
CHAPTER 1: Introduction	
1.1 Background	2
1.2 Problem Statement	4
1.3 Problem Questions	5
1.4 Objective of the Study	5
1.5 Significance of the Research	5
1.6 Research Terminology	6
1.7 Study Limitations	6
1.8 Summary of Chapters	7

CHAPTER 2: Literature Review	
2.1 Introduction	9
2.2 Cost Estimation	9
2.3 Level of Accuracy	10
2.4 Factors Affecting Cost Estimation Accuracy	15
2.5 Cost Estimation Models	25
2.6 Chapter Summary	29
CHAPTER 3: Methodology	
3.1 Introduction	32
3.2 Evaluation of Cost Estimation Accuracy	32
3.3 Cost Estimation Model	34
3.4 Factors Affecting Construction Cost Estimation	36
3.5 Questionnaire Survey Analysis	42
3.6 Modeling Factors Affecting Construction Cost Estimation	55
3.7 Chapter Summary	61
CHAPTER 4: Data Collection	
4.1 Introduction	63
4.2 MYS Construction Project Data	63
4.3 Factors Affecting Cost Estimation Accuracy	65
4.4 Questionnaire Results	70
4.5 Chapter Summary	75

CHAPTER 5: Data Analysis	
5.1 Introduction	77
5.2 Analysis of Actual Cost Data	77
5.3 Development of Construction Cost Estimation Model	89
5.4 Analysis of Factors Affecting Cost Estimation Accuracy	95
5.5 Modeling of Cost Estimation Accuracy using PLS-SEM Technique	110
5.6 Satisfaction with Construction Cost Estimation Model	126
5.7 Chapter Summary	138
CHAPTER 6: Results and Discussion	
6.1 Introduction	140
6.2 Level of Accuracy of MYS cost estimation process	140
6.3 Factors Affecting Cost Estimation Accuracy in MYS	143
6.4 Factors affecting Cost Estimation Model	147
6.5 Chapter Summary	148
CHAPTER 7: Conclusions, Recommendations and Further Studies	
7.1 Conclusion	151
7.2 Recommendations	152
7.3 Limitation of The Study	152
7.4 Further Recommended Studies	153
List of References	154

List of Tables

Table No	Table Title	Page
Table 3-1	Descriptive analysis performed on actual construction projects data	33
Table 3-2	Regression models assessment statistics and its decision criteria	35
Table 3-3	Relative importance index	43
Table 3-4	Sample size Recommendation a in PLS-SEM for statistical power of 80%	51
Table 3-5	Quality criteria for PLS-SEM	54
Table 4-1	Actual MYS construction project data base	64
Table 4-2	List of factors effecting construction cost estimation linked to its categories	67
Table 4-3	Factors affecting MYS construction cost estimation (client and consultant perspective)	71
Table 4-3	Factors affecting MYS construction cost estimation (client and consultant perspective) (continued)	72
Table 4-4	Factors affecting MYS construction cost estimation (main contractor perspective)	73
Table 4-4	Factors affecting MYS construction cost estimation (main contractor perspective) (continued)	74
Table 5-1	Distribution of construction projects based on project function	77
Table 5-2	Distribution of construction project based on project size	78
Table 5-3	Distribution of construction projects based on year of award period	78

Table 5-4	Comparison between projects estimated costs and contract awarded amounts	79
Table 5-5	Construction projects cost deviation statistics	80
Table 5-6	Tests results for Kolmogorov-Smirnov Normality test for overall cost deviation	81
Table 5-7	Tests results for Kolmogorov-Smirnov Normality test for underestimation and overestimation cost deviation	82
Table 5-8	Classifying the cost deviation based on Permissive criterion (-10% to +15% estimation error)	82
Table 5-9	Student t-test results for overall, underestimation, and overestimation of cost deviation	83
Table 5-10	Descriptive analysis for construction projects cost deviation based on project function, awarding period, and project size	84
Table 5-11	Inferential statistical test (ANOVA) results for cost deviation based on project function, awarding period, and project size.	85
Table 5-12	Tukey HSD test results for project size variable	86
Table 5-13	Effect size for Tukey HSD test result	87
Table 5-14	Categorization of projects based on duration spent between approving the budget and awarding the contract	87
Table 5-15	Mean cost deviation for the period between (2012-2019)	88
Table 5-16	Robust linear regression- model 1 validation	94
Table 5-17	Stepwise regression- model 2 validation	94
Table 5-18	Participants demography	95
Table 5-19	Summary of the cost estimation practice in participants organization	96

Table 5-20	Participants degree of satisfaction on pretender cost estimation accuracy	98
Table 5-21	Critical factors that affect MYS construction cost estimation	99
Table 5-22	Research categories reliability values	100
Table 5-23	Participants perception on scope quality effect on cost estimation	101
Table 5-24	Participants perception on estimator performance effect on cost estimation	102
Table 5-25	Participants perception on information quality effect on cost estimation	103
Table 5-26	Participants perception on project characteristics effect on cost estimation	104
Table 5-27	Participants perception on external factors effect on cost estimation	106
Table 5-28	Participants perception on contractual procedure effect on cost estimation	107
Table 5-29	Participants perception on organization management effect on cost estimation	108
Table 5-30	Mean categories scores based on participants managerial level	109
Table 5-31	MOW engineers perception on additional factors effect on cost estimation accuracy	110
Table 5-32	Results of (client/consultant) PLS-SEM model	114
Table 5-33	Results of contractor PLS-SEM model	117
Table 5-34	Results of common categories between participants PLS-SEM model	120

Table 5-35	Comparison between PLS-SEM models	120
Table 5-36	Assessment of common categories measurement model	122
Table 5-37	Locker and Fornell criterion results for common categories model	123
Table 5-38	Outer loading results for common categories of the model	124
Table 5-39	VIF and Q results for common categories of the model	124
Table 5-40	Summary of path coefficient, R^2 , effect size for results for common categories of the model	125
Table 5-41	The results of best neural network performance using Bayesian generalization and Scaled Conjugate Gradient Algorithms	127
Table 5-42	The results of best neural network performance using Levenberg-Marquardt learning algorithms neural networks	127
Table 5-43	(5-10-1) ANN hidden input, output and bias connection weight	132
Table 5-44	Standardization of input variables to establish the ANN hidden nodes equations	133
Table 5-45	Multiplication results of hidden connection weight by the standardized input variables	134
Table 5-46	Results of sum hidden neuron values with hidden bias and connection weight to the output neuron	134
Table 5-47	Categories ranking as per normalized connection weight method	138

List of Figures

Figure No	Figure Title	Page
Figure 2-1	Attributes of factors affecting the estimation accuracy practice	17
Figure 3-1	Conceptual framework for the categories of factors affecting accuracy of construction cost estimation	37
Figure 3-2	Range of estimation activities	38
Figure 3-3	An initial conceptual model for the relation between the factors and cost estimation accuracy	42
Figure 3-4	A simple PLS-SEM path model	46
Figure 3-5	The diagram depicts the flowchart for the PLS algorithm	48
Figure 3-6	Procedures followed for estimation of weights and latent variable in PLS-SEM model	49
Figure 3-7	Construction PLS-SEM model procedures	53
Figure 3-8	A Neural Network Model Architecture	56
Figure 5-1	Histogram distribution for construction projects cost deviation	81
Figure 5-2	Pattern of cost estimation deviation between (2012-2019)	88
Figure 5-3	History list for robust linear regression- model 1	89
Figure 5-4	Response plot for robust linear regression- model 1	90
Figure 5-5	Predicted vs true response plot for robust linear regression- model 1	90
Figure 5-6	Residuals plot for robust linear regression- model 1	91
Figure 5-7	History list for stepwise regression- model 2	92
Figure 5-8	Response plot for stepwise regression- model 2	92

Figure 5-9	Predicted vs true response plot for stepwise regression- model 2	93
Figure 5-10	Residuals plot for stepwise regression- model 2	93
Figure 5-11	Proposed client/consultant PLS-SEM model	111
Figure 5-12	Results of client/consultant PLS-SEM model	113
Figure 5-13	Proposed contractor PLS-SEM model	114
Figure 5-14	Results of contractor PLS-SEM model	116
Figure 5-15	Proposed common categories between participants PLS-SEM model	117
Figure 5-16	Results of common categories between participants PLS-SEM model	119
Figure 5-17	Coefficient of correlation of LM-10N neural network model	128
Figure 5-18	Validation performance for LM-10N model	129
Figure 5-19	Error distribution for LM-10N model	130
Figure 5-20	ANN diagram for model 5-10-1	130
Figure 5-21	The architecture schematics for 5-10-1 ANN model	131
Figure 5-22	Multiplication result of the transferred sum hidden layer values and output connection nodes for 5-10-1 ANN model	135
Figure 5-23	Coefficient of correlation for LM-10N model using cross validation data	136
Figure 5-24	Error distribution for LM-10N model using cross validation data	137

List of Abbreviations

AACE	Association for the Advancement of Cost Engineering
ANFIS	Adaptive neuro-fuzzy inference system
ANN	Artificial neural network
ARIMA	Seasonal Autoregressive Integrated Moving Average
ASTM	American Society for Testing and Materials
BC	Block Chain
BIM	Building Information Management
CBR	Case-based reasoning
CCI	Construction Cost Index
CE	Cost estimation
COPRAS	Complex Proportional Assessment method
EFA	Exploratory Factor Analysis
EMR	Electronic medical records
KBDSS	Knowledge Based Decision Support System
LR	Logistic regression
MAD	Mean absolute deviation
MAPE	Mean Absolute Percentage Error
MLP-ANN	Multilayer Perceptron artificial neural network
MOW	Ministry of Works
MRA	Multiple regression analysis
MSE	Mean squared error
MYS	Ministry of Youth and Sport Affairs
nftool	Neural Net Fitting Tool application
PLS-SEM	Partial Least Square Structural Equation Model

PNN	Probabilistic neural network
RMSE	Root mean square error
S _a	Cost estimation satisfaction score
SEM	Structural equation model
SERVQUAL	Customer satisfaction
S _{nor}	Normalized cost estimation satisfaction score
SPM	Minimum service standard
SQ	Service quality
TRSQ	Troll road service quality
VECM	Vector Error Correction Model

List of Notations

Symbol	Definition
RII	Relative importance index
n_i	The number of respondents for each Likert scale category
N	The total number of respondents
Y_i	A latent variable
Z_i	A latent variable error
x_i	Indicator used to observe latent variable
e_i	Error terms related to the indicators.
V	A vector of all latent variables
B	A matrix of path coefficients
Z	The disturbance term of the dependent latent variables
W_{bk}	The contribution of the indicator x where x_k ($K=1,2,...,k_{max}$) to Y
z	Error associated with latent variable Y
Y_j	The Dependent latent variable
Y_k	The independent latent variable
Y_{jk}	The value of the covariance between (Y_j ; Y_k),
b_{jk}	The inner weight between the adjacent latent variables Y_j and Y_k
$\tilde{Y}_{jn}$	The latent variable scores from the inside approximation
$\hat{w}_{bkj}$	Updated outer weights
$x_{k j n}$	The raw data for indicator k ($k=1, \dots, K$) of latent variable j ($j=1, \dots, J$) and observation n ($n=1, \dots, N$)
z_{jn}	The error term from a bivariate regression
Y_{nor}	Normalized output variable
f_o	Transfer function of output layer

b_o	Bias at the output layer
w_{so}	Connection weight between the s^{th} hidden neuron and the single output neuron
f_h	Transfer function of hidden layer
b_s	Bias of the s^{th} hidden neuron ($s=1, S$)
w_{rs}	Connection weight between the P_r input variables ($r=1, \dots, R$) and the s^{th} hidden neuron
X_{ni}^{nor}	Normalized P input variable (Normalized in the range of [-1,1] for tan-sigmoid transfer function)
X_{ni}	Actual given input data
X_n^{max}	The minimum and maximum data over values of category n.
X_n^{min}	The maximum data over values of category n
$Tanh(x)$	Tan-sigmoid transfer function for weighted sum of the input x
R_p	Relative importance of the input p in reference to the output
s	Number of neurons in hidden layer, $s=1, \dots, S$
$F(df_1, df_2)$	The F value for One-way Anova
df_1	The numerator d egrees of f reedom
df_2	the denominator degrees of freedom
η_p^2	Partial eta squared

Glossary

Term	Definition
Average Variance Extracted (AVE)	A measure associated with a latent variable, which is used in the assessment of the discriminant validity of a measurement instrument. Less commonly, it can also be used for convergent validity assessment.
Communality Index	The sum of the squared loadings for that latent variable, each loading associated with an indicator, divided by the number of indicators.
Composite Reliability	This is a measure of reliability associated with a latent variable. Unlike the Cronbach's alpha coefficient, the compositive reliability coefficient takes indicator loadings into consideration in its calculation.
Cross Loading	A PLS-SEM assessment that examine the loadings of an indicator on its assigned latent variable should be higher than its loading on all other latent variables.
Effect Sizes	The absolute values of the individual contributions of the corresponding indicators to the R-squared coefficients of the latent variable to which each indicator is associated.
Fornell and Larcker Criterion	For each latent variable, the square root of the average variance extracted should be higher than any of the correlations involving that latent variable.

Homoscedasticity	An essential assumption in developing structural equation modelling were the variable's residual (error) exhibits consistent variance across different levels of the variable.
Internal Consistency	This is a measure of reliability associated a latent variable. It usually increases with the number of indicators used and is often slightly lower than the composite reliability coefficient.
Model Predictive Relevance	The capability of the model to predict by repeating the observed values by the model itself through blind folding procedures. Q-squared coefficients are provided only for endogenous latent variables.
Outer Loading	The correlation coefficient for the variable and factor.
R-Squared	The percentages of explained variance and predictive validity associated with each of those latent variables.
Tenenhaus Gof	The square root of the product between the average communality index and the average R-squared.
Variance Inflation Factors (VIF)	A measure of the degree of collinearity (or multicollinearity) among variables, including both indicators and latent variables.

List of Appendices

	Appendix Title
1	Summary of top and lowest ranked factors affecting construction cost estimation from recent publications
2	Quantitative measures and methodologies used to evaluate projects cost estimation accuracy
3	Interpretation of Bayes Factors as evidence for null hypothesis (H0) and alternative hypotheses (H1)
4	Mathematical expression to calculate PLS-SEM quality terms
5	Questionnaire survey used to measure factors affecting construction cost estimation
3	Descriptive data analysis for cost deviation in MYS construction projects for the period between (2012-2019)
4	Questionnaire survey results and RII calculation
5	PLS-SEM models results
6	ANN model results and mathematical equation development
7	ANN model: Categories ranking as per normalized connection weight method

CHAPTER 1: Introduction

1.1 Background

1.2 Problem Statement

1.3 Problem Questions

1.4 Objective of the Study

1.5 Significance of the Research

1.6 Research Terminology

1.7 Study Limitations

1.8 Summary of Chapters

CHAPTER 1: Introduction

1.1 Background

Global competitiveness exerts pressure on organizations to improve their portfolio management. Diverse organization objectives obligate management of interdependent programs in an effective and efficient manner. Most inconstant discipline in an organization with regard to management is construction management. The multi-projects environment of the construction program imposes tremendous challenges on managers ability to anticipate projects risks and reduce their probability of occurrence.

The need for a sustainable development growth in projects construction imposed imperative demand to study the drawbacks that occur frequently in projects life cycle and may hinder the delivery of construction projects process. Cost performance, principally efficacy of the projects cost estimation, is fundamental criteria for construction projects success.

Cost estimation is one of the crucial tasks handled in the project initiation phase. This information comprises a considerable percentage of the weighting criteria in management decision of authorizing the project or abandon it. The cost constraint in construction projects imposed pressure on managers to trace the cost development during design phases “strategic planning, conceptual planning, preliminary design, detailed design, and construction” (AbouRizk et al., 2002). The cost estimates perform during planning phases are directed to achieve more reliable construction cost determination.

Several kinds of literature reported about mega projects that suffer from wrong budget estimation which affects adversely in their progress such as Chicago Deep Tunnel Project (Ross and Staw, 1993), Shoreham nuclear power plants that started with a cost estimation of \$75 million and finished with \$5 billion in 1989 (Ross and Staw, 1993), and Brandenburg airport where the cost

escalation reached 4.3 billion Euros with an initial forecast of 2.5 billion Euros (Flyvbjerg, 2017). Therefore, an accurate project cost estimation is a necessity in today working business.

The phenomenon of inaccurate cost estimation has been observed in different countries around the world. Flyvbjerg et al. (2002) proclaimed that the percentage of project cost underestimation among 258 transportation infrastructure projects across 20 countries is 86%. Conversely, Aibinu and Pasco (2008) noted in their study of the accuracy of pre-tender building cost estimates in Australia that pre-tender building costs are more often overestimated than are underestimated. Skitmore and Picken (2000) research findings corroborated the previous viewpoint when a pre-tendered estimating performance of 217 United State of America projects against the lowest tender values analysis revealed that the projects were overestimated by an average of 5.19%.

With regard to the Kingdom of Bahrain, a study emphasized the role of estimation as a root cause for cost overrun in the highway projects (Alkooheji and Suliman, 2017). Alkooheji and Suliman (2017) highlighted that cost estimation is among the factors to be investigated in their contribution in cost overrun. The statistical analysis of the participants responds on a questionnaire ranked the factors related to estimation in the second place in its contribution in cost overrun with an important index value of 71.03%.

The consequences of inaccurate cost estimation create serious issues for project stakeholders. Pillai, M. P. (2008) illustrated the impact of deficiency in cost estimation for both client and contractor. The client will be exposed to financial risks related to not realizing the return from capital investment and predicament in funding other projects. For contractors, the effect is more difficult, as the financial loss carried due to inaccurate estimation may drive the companies to defunct.

The negative impacts encourage organizations and researchers to work together aiming to avert the outcomes of inaccurate cost estimation. Several literatures demonstrated different aspects related to inaccurate cost estimation knowledge area. For example, Hatamleh et al. (2018) rank the factors affecting the accuracy of cost estimation in Jordan. Skitmore and Picken (2000) examined the accuracy of two hundred and seventeen building projects pre-tender prices forecast in united states of America. Researches like Cho et al. (2013) used ninety-six cost data of elementary schools in Korea to formulate a model to estimate construction cost using multiple regression analysis and neural network.

Among available publications, there were lack of researches dedicated to capture the inaccuracy of cost estimation in building construction projects at the gulf countries, and emphasize organizations efforts in implementing alternate methods to improve the cost estimation process in building construction industry.

1.2 Problem Statement

Ministry of youth and sport affairs (MYS) is a governmental ministry that is responsible for constructing infrastructure facilities for youth and sports sectors. MYS has executed more than 40 sport construction projects in the previous 5 years. Although the accumulative knowledge MYS acquired in construction of such facilities is great, nevertheless the unique nature of each project makes the preliminary project cost estimate a repetitive bottleneck in projects life cycle. The problem becomes more obvious currently. It has been noticed in eight of the ten projects listed in 2018 ministry construction plan, the bidders prices are exceeding the allocated budgets, raising concerns regarding the accuracy of cost estimation process done in the planning phase.

With MYS commitments towards accomplishment of construction projects listed in the governmental action plan, underestimating projects costs threatens MYS capability to achieve the ministry annual construction targets. Such

underestimation increases the effort made in rework activities: negotiations with lowest bidder and addressing a request for additional fund, eventually will postpone the project execution and delay its handing over to end users. In brief the inaccuracy of cost estimation of MYS construction projects is a problem that should be addresses.

1.3 Problem Questions

1. What factors are affecting current MYS cost estimation practice performance?
2. What is the level of cost estimation accuracy in MYS construction projects?
3. What is the cost estimation model that can help in improving the estimation process in MYS construction projects?

1.4 Objective of the Study

This study on the inaccuracy of cost estimate of MYS construction projects seeks to attain the following objectives:

1. To identify the factors affecting the current cost estimation practice performance in MYS projects.
2. To assess the level of cost estimation accuracy in MYS construction.
3. To develop a cost estimation model for future cost prediction.

1.5 Significance of the Research

The findings of this study will enrich the cost estimation field / knowledge related to Kingdom of Bahrain. Particularly, the investigation will contribute numerous improvements in MYS construction sports infrastructure portfolio management. The study findings influence on MYS construction projects performance includes and not limited to: demonstrate level of accuracy of current cost estimation process, identify the factors affecting significantly construction

cost estimation performance, and enhance cost estimation process through employment of the technologies available in developing a cost estimation model.

1.6 Research Terminology

Common terminologies that will be used in the research are:

1. **Project:** According to *PMI Lexicon of Project Management Terms* (2017), *project* is “a temporary endeavor undertaken to create a unique product, service, or result”.
2. **Project Budget:** According to *PMI Lexicon of Project Management Terms* (2017), *project budget* is “The sum of work package cost estimates, contingency reserve, and management reserve”.
3. **Estimate:** According to *PMI A Guide to the Project Management Body of Knowledge (PMBOK® Guide) (2013:539)*, *estimate* is “A quantitative assessment of the likely amount or outcome. Usually applied to project costs, resources, effort, and duration and is usually preceded by a modifier (i.e., preliminary, conceptual, feasibility, order-of-magnitude, definitive). It should always include some sign of accuracy (e.g., $\pm x$ percent)”.
4. **Underestimation:** According to *Cambridge Dictionary* (2018), *underestimation* is “fail to guess or understand the real cost, size, difficulty, etc. of something”.
5. **Accuracy:** According to *PMI A Guide to the Project Management Body of Knowledge (PMBOK® Guide) (2013:526)* *accuracy* is “within the quality management system, accuracy is an assessment of correctness”.

1.7 Study Limitations

The study will address the cost estimation inaccuracy occurring in MYS construction projects. Therefore, it was designed to determine the level of cost accuracy in MYS construction projects and identify the factors affecting cost estimation accuracy in MYS as per stakeholder’s perceptions. The participant will

be limited to construction engineers who involved in MYS construction projects during the last six years from three main categories: client, consultancy office and main contractors. Target participants will be from different position levels: top management, seniors engineers, construction engineers, architects and quantity surveyors.

1.8 Summary of Chapters

The research consists of 7 chapters. Chapter one contains a background of cost estimation inaccuracy, description of construction cost estimation level at MYS, research problem, and research objective. Chapter two presents a literature review for the deficiency in construction cost estimation with an illustration of the factors affecting cost estimation accuracy studied by different researchers. Chapter three describes the methodology adopted to determine the level of construction cost estimation accuracy, factors affecting MYS cost estimation process and proposed statistical analysis and data modeling to be used in the research. Followed by Chapter four that explains the procedures followed to collect the construction projects data base, selection of factors affecting construction cost estimation, questionnaire design and results. Chapter five contains a comprehensive data analysis for: construction projects data base, construction cost model, research questionnaire and factors affecting construction cost estimation models. The results and discussion will be presented in Chapter six. Finally, Chapter seven includes conclusion, recommendations and limitations of the research and suggestions for future studies.

CHAPTER 2: Literature Review

2.1 Introduction

2.2 Cost Estimation

2.3 Level of Accuracy

2.4 Factors Affecting Cost Estimation Accuracy

2.5 Cost Estimation Models

2.6 Chapter Summary

CHAPTER 2: Literature Review

2.1 Introduction

The aim of this chapter is to provide general insight into the cost estimation process. Works of several major publications in construction cost estimation will be reviewed to align research direction with the existing literatures.

The chapter will start by providing general information about cost estimation. Section 2.3 gives an illustration of previous research findings on accuracy of construction cost estimation, followed by two sections reviewing the literature related to factors affecting the construction cost estimation accuracy and researchers' contributions in developing construction cost estimation models.

2.2 Cost Estimation

Mislick and Nussbaum (2015) defined the cost estimate as follows: “cost estimate is the process of collecting and analyzing historical data and applying quantitative models, techniques, tools, and databases in order to predict an estimate of the future cost of items, product, program or task. Cost estimation is the application of the art and technology of approximation of the probable worth (or cost), extent or character of something based on information available at the time”.

Cost estimation information plays a significant role in managerial decisions during the project life cycle. During the initiation phase, it is used within the evaluation criteria to authorize the project (Aibinu and Pasco, 2008). Oladokun et al. (2011) mentioned the use of cost estimate for tender comparison or negotiation in case the submitted bids are above the budget. Schottlander (2006) emphasized the role of estimates as a management tool for budgetary control. Therefore, deficit in cost estimation is a serious

problem and will expose the project into serial associated problems like failure to award a construction project due to receiving high bids or low bids design problems, and project delays (National Research Council, 1990).

Professional associations such as Project Management Institute (PMI) developed a general framework to facilitate the cost estimation process and minimize the risk of cost estimation deficiency, but still, the application of these estimation tools in the field is limited. Akintoye and Fitzgerald (2000) studied the current cost estimation practice of 84 UK construction firms. Their questionnaire survey revealed 14 different techniques for preparing cost estimates: “estimating standard procedure, comparison with past similar projects based on personal experience, comparison with similar past projects based on documented facts, established standards, intuition, arithmetic formula, estimating software, published price information, capital estimating factors, shared information with subsidiary of the firm, range estimating, guessing, shared information from other firms and complex statistical formulas”.

Although the above tools, the developed technologies, and software are available to establish projects conceptual cost, still early cost estimation process is not controlled by the profession. Abbasi and Al-Mharmah (2000) observed through their survey of 50 industrial public firms a lack of utilizing modern project management tools in project management filed in Jordan which affects adversely on sector abilities in planning, monitor and control projects costs.

2.3 Level of Accuracy

Monitor the cost estimation process performance primary start by determining the level of its accuracy. It allows the organization an insight

about current estimation accuracy level and provides evidence for possible process improvements toward achieving high-quality projects cost estimation.

McCaffer (1976) clarified the advantages of establishing cost estimation accuracy as follows: “to quantify the reliability of cost estimate process, to use it as a base to compare with the alternative estimating procedure, and to produce correction factor”. Consequently, the absence of feedback mechanisms impacts estimation process, which eventually will cause errors accumulations and organization lost their opportunity to improve their estimation process by detecting any existing errors trend and eliminate it (Flanagan and Norman, 1983).

Several studies concentrate on assessing the level of cost estimate accuracy in construction projects. The evaluation mainly determines the cost estimation accuracy through biases in estimates (cost deviation) for sample projects. Researchers contribution in reporting level of cost estimation accuracy was started in 1968. Dysert (2007) explained the need for establishing cost estimate accuracy range in order to get better cost estimation performance. Most of the previous literature used the percentage difference between the awarded contract sum and cost estimated to measure the level of cost estimation accuracy.

Flyvbjerg and Holm (2002) observed in cost escalation evaluation of 258 various types of transportation projects large biased distribution of errors in cost estimates around zero and it was interpreted by the existence of systemic underestimation. Statistical analysis (ANOVA) confirmed the effect of project types on cost escalation and the continuity of project cost underestimation practice over time. No statistical details were reported for dispersion around the mean in the research.

In addition to cost estimation accuracy analysis, Skitmore and Picken (2000) studied the intercorrelation among multiple factors (building type, project size, year and sector) and its confounding effect on the cost estimation accuracy. Their partialled out analysis confirmed the responsibility of year factor on the forecast bias and inconsistency which has been related later to the effect of annual rate of inflation.

Aibinu and Pasco (2008) studied the cost estimation accuracy of 56 projects in Australia. The analysis shows that pre-tender cost predictions tended to be overestimated than underestimated with a significant effect of project value, number of storeys and structural material on estimated bias when it has been analyzed using ANOVA.

Similar findings has been found by Oladokun et al. (2011) who examined the accuracy of pre-tender cost estimates of 81 building projects in Nigeria between 2005-2008. Categorizing the projects in terms of size, year, type, and sector supports the finding that the projects tend to be overestimated with a significant effect of project size on bias, where the estimates of smaller projects are more biased than the larger projects. The reason mentioned for this observation was the practice of preparing small project cost estimates by junior quantity surveyors.

Ling et al. (2010) reported the estimating accuracy for 42 building projects constructed in Singapore between 1992-1998. The finding has been compared with those of Skitmore and Gunner (2000), which contain a larger construction project sample size of 181. Using student test analysis, it has been concluded that no significant difference in overall estimation accuracy results found between both studies, which is an indicator that there is no evidence for accuracy improvement over time and the presence of in-built estimation bias.

Other techniques have been implemented for understanding the cost estimation process precision. For example, Newton (2012) adopted Analytics approach based on the cumulative probability distribution method mentioned by Flyvbjerg (2008) as a tool to ascertain the estimation process performance. The plot generated suggested the addition of 32.75% as variation percentage mark up to cost calculated, it will provide a 90% probability that the budget estimate will be sufficient to cover the actual costs of any project.

Azrai and Ismail (2012) ascertained cost estimation reliability, and preformed linear multiple regression analysis with 14 predictors to demonstrate the relationship between project characteristics and mean bias using three different ratio approaches (estimate/lowest bid, estimate/acceptable bid and estimate/mean bids). The multiple regression model of mean bids and the project's characteristics (project size, number of bidders, type of schools and state) result in a lowest standard error of 8.306 and highest $R^2 = 38.4\%$. The results made the authors to advocated the preference of using the mean of the bids as the most proper parameter to measure the accuracy of the cost estimate.

Past and recent studies related to the Middle East tend to report the inaccuracy of construction cost estimation under the term of cost overrun or cost escalation. Generally, the percentage cost deviation during execution phase from the original project budget has been reported without providing statistical analysis toward the data collected. The following studies have been reviewed and mentioned as extended observations in the inaccuracy of the cost estimation area.

In Kuwait, inadequate construction budget estimation has been reported by Koushki et al. (2005). All the residential houses data investigated in the study suffered from cost overrun during the execution

phase. The average cost overrun found to be (1% to 6%), and the additional cost range starts from (KD 2000 to KD 10000) from the total budget allocated. More recent research done by Al-Sabah and Refaat (2019) highlighted the percentage of public projects that faced cost overrun. Almost 33% of large projects (contract amount of 100-150 million Kuwaiti Dinar) have encountered cost overrun. No details have been reported about the percentage of cost overrun.

A recent study by Rivera et al. (2017) investigated the global performance of the construction industry concerning rework, cost overrun, project delay, and customer satisfaction. The statistics confirmed that studied regions suffer from cost overrun and schedule delay problems. The Middle East had the third place of project over budget (65%), second place of project delayed (79%) with minimum percentage over budget amount of (15%) and percentage delay amount of (30%).

Additionally, Senouci et al. (2016) assessed the cost overrun in Qatari public construction projects. The cost overrun percentages corroborated that building projects suffered from huge cost overrun with a value of 70.3% when it compared with other types of projects cost overruns (road: 18.2%, drainage: 10.8%). The inferential analysis has been concerned with cost overruns percentages and project type, categories, size, project duration and year of completion. Analysis of variance failed to reject the presence of relationships between the cost overrun percentages and the mentioned criteria, except for the year of completion factor. The results found reject the time influence hypothesis on project cost overrun as the construction project cost overruns for the project completed between 2007 and 2013 statistically different and lower than those completed between 2000 and 2007.

In contradiction, the effect of project size has been observed in a study performed in Iran, as the project performance indicated by cost overrun and project delay was noticed to be minimum in small-budget and short-term municipal projects (Heravi and Mohammadian, 2017).

The consistency of cost performance evaluated by Kadiri (2014) confirmed the influence of project type on the prediction of construction cost estimation. It has been found that low rise buildings were underestimated by 8.87%, and high-rise buildings were overestimated by 23.43%.

Interestingly, Saidu and Shakantu (2017) observed an unusual cost overrun along with building construction projects in Nigeria. With an average of 44.46% cost overrun during project execution, the overall cost overrun range is unpredictable, where minimum of 5.56% for 90% project completion, to a maximum value of 216% for 5% project completion was observed. This finding implies the continuity of initial budget escalation in case of poor management and extension of time factors exist during project execution.

The recent results obtained from Ghana by Coffie et al. (2019) supports the inadequacy in construction cost estimation as in average 23.6% of public sector projects had cost overrun.

2.4 Factors Affecting Cost Estimation Accuracy

Cost estimation process evolution from traditional practice to database simulation was started in the early 1970s. Research works were dedicated to developing reliable cost models in cost planning and prediction. Ogunlana (1989) has detailed the pioneer efforts of: Beeston (1973 - 1983), McCaffer (1976 - 1982) and Bennet (1981 - 1982) in this field.

Given that cost estimation is a complex process containing estimating two types of factors: high visible factors such as labor, material, equipment, and low visible factors, which affecting the accuracy of the cost estimation (Al-khalidi, 1990). Hence, identifying the cost determinant variables and quantify their degree of influence will enhance the effectiveness of the estimation cost model and increase its efficiency (Elhag et al., 2005).

Precedent work in this field concentrated in finding evidence between the accuracy of design cost estimates with certain project characteristics. For example, McCaffer (1976) investigated the estimating accuracy of design cost against the number of bidders and contract value of 132 building projects. He found a strong negative relation between the ratio ($\text{lowest bid} / \text{design cost estimate} \times 100$) and the number of bidders, but an inconsistent correlation pattern was found between previously mentioned ratios against the contract value. The above findings aided the author in establishing correction factor to adjust the accuracy of the design estimate.

Morrison and Steven (1981) had raised the attention regarding the expected effect of qualitative factors on estimating project cost in their research. They demonstrated with further extension the possible effect of design team integration, project characteristics, time available for estimation and availability of data on the accuracy of the cost estimation process.

Similar to the above focus, Skitmore and Ashworth (1983) referred to other researchers findings to express the effect of availability of design information, market condition and personal characteristics on the accuracy of the design cost estimate. The authors concluded by affirming the impact

of qualities such as experience and intuition in forecasting accurate construction price which was proved by Skitmore (1985) experiment.

Starting from the late 1980s, dissertations like those of Ogunlana (1989) and Bley (1990) constituted a comprehensive framework for the inherent factors that may affect the cost estimation process. Their theoretical framework facilitated the work of many researchers, who had the same attention in determining the factors affecting cost estimation accuracy and utilize this knowledge in improving the cost estimation process.

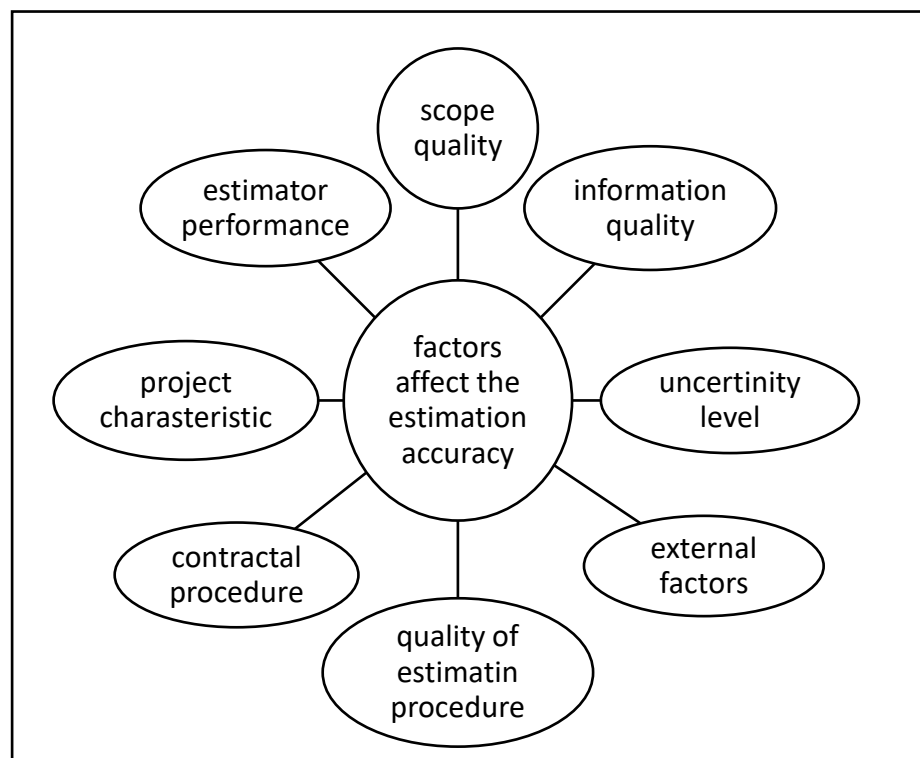


Figure 2-1 Attributes of factors affecting the estimation accuracy practice

A growing body of literature related to this problem has been reviewed. From the previous stated publications and another 24 publications listed in Appendix 1, authors have divided the factors affecting the project cost accuracy into 8 main attributes shown in Figure 2-1. The attributes are scope quality, information quality, uncertainty level,

external factors, quality of estimation procedures, project characteristics, estimator performance, and contractual procedure.

Some preliminary work was carried out in the early 1990s by Al-khaldi (1990) who studied the effect of 29 factors on the cost estimation in Saudi Arabia public building construction. The survey analysis revealed the finance group factors had the highest effect on the accuracy of cost estimation with a total main important factor of 76%. The participants agreed on considering the previous experience of contract is the severest factor in cost estimation accuracy followed by progress payment.

A More recent study done by Al bubshait and ALjuwairah (2002) studied the contribution of 42 factors in the construction cost in Saudi Arabia which will eventually affect the accuracy of the projects cost estimate. The factors found to have a severe effect on construction cost are: material cost, incorrect planning, previous experience of the contract, contract management and poor financial control on site.

In a major advance in 2002, Flyvbjerg and Holm (2002) proposed four main explanations to understand the phenomena of inaccurate cost estimates: technical, physiological, economical, and political. Having a large project sample aided in performing statistical analysis; hence, eliminating the first two explanations and promoting clear statistical evidence that economics and politics have a rational justification for the inaccuracy in cost estimation.

Further research done by Trost et al. (2003) used different approaches which were a combination of factor analysis and multivariate regression analysis to evaluate the early cost estimate performance. By applying the aforementioned analysis on respondents ranking of 45 drivers of cost estimate accuracy, it has been found that the studied factors are

responsible for 70% of variance in the data. In addition, basic process design with team experience and cost information are the greatest significant factor groups in the developed model.

Koushki et al. (2005) highlighted several circumstances that lead to time/cost increase in Kuwait, such as incomplete design drawings, contractor related problems followed by material-related problems, and owner's financial constraints.

In another major study, Elhag et al. (2005) grouped 67 cost determinants in six categories and conducted a survey to identify the influence of factors on cost estimation. They found that 52 out of 67 factors has a severity index between 65% and 95%. The category containing consultant and design parameters was ranked top, followed by client characteristics, with four out of the 10 top factors being related to this category. In their investigation into factors affecting the quality of building projects in Hong Kong, Chan and Tam (2006) used factor analysis and stepwise multiple regression analysis to identify significant factors affecting the quality performance. The factors found as follows: “project management actions, client's emphasis on time, client's emphasis on quality, and effectiveness of construction team leader”. More recent evidence, Akintoye (2010) performed factor analysis to detect the underlying relationship between the factors. The output of the analysis revealed that the most important factors are project complexity followed by technological requirements, project information, project team requirement, contract requirement, project duration, and market requirement.

Since 2013, many attempts have been made to identify the factors affecting cost estimate and construction cost in Asian and African countries. The following section will outline the main findings related to

the dissertation variables of each region. Further details of publication significant factors extracted from the mentioned literatures are tabulated in Appendix 1.

2.4.1 Factors affecting cost estimation accuracy in Africa

Nigeria is from the countries who have the lead publications that emphasized the cost estimation factors of construction projects. Alumbugu et al. (2016) evaluated the influence of 54 factors on construction cost accuracy in Nigeria. Using structured questionnaire, they found that the top ten factors affecting the accuracy of pre-tender cost estimate are: experience and skill level of the consultants, project teams experience on the construction type, clear and detail drawings and specification, completeness of cost information, accuracy and reliability of cost information, availability of all fields of specialization in a project team, quality of information and flow requirement, clear scope definition for the client, financial capability of the client, and completeness of project documents.

Oke (2017) found that the project characteristics have a great influence on contractor tender price estimation during the tendering stage followed by consultant and design characteristics. Additionally, the participants have recognized the effect of market conditions on the tendering price too.

A more recent research by Amusan et al. (2018) highlighted the critical determinate causing cost overrun in Nigerian projects during execution phase, which found to be: the contractors inexperience, inadequate planning, inflation, incessant variation order, and change in project design, complexity, shortening of project period and fraudulent practices.

The consequence effect of qualitative factors upon construction projects success criteria (cost, time) has been investigated by Famiyeh et al. (2017). The conducted study highlighted the critical factors causing cost and time escalation problem in Ghana construction projects. Among the important factors found and tie in line with current study: unrealistic contract durations imposed by clients, poorly defined project scope, under-estimation of project cost by consultants, delays in the issuance of permits by government agencies, variations in designs, lack of communications plans, poor feasibility and project analysis, and material price fluctuations.

In Ghana, Offei-Nyako et al. (2016) identified 26 out of 40 factors that are influencing the deviation between contract sum and final accounts for capitals. Using mean important indices values, the top five ranking factors are: “price fluctuations, late material delivery, changes in the scope of work, fluctuations in the market demand, and changes in design”.

2.4.2 Factors affecting cost estimation accuracy in Asia and Middle East

Enshassi et al. (2013) studied the effect of 64 factors on the accuracy of pre-tender cost estimating in Gaza strip. The results of 70 questionnaire survey distributed to clients and consultancy offices revealed that the factors related to consultants/ design parameters and information and external factors and market conditions had the most significant effect on the pre-tender cost estimation.

Cheng (2014) succeeded in identifying 16 significant factors out of 90 factors influencing the construction cost in Taiwan. By using the Modified Delphi Method (MDM), Kawakita Jiro method (KJ) and severity index computation, industry professional ranked scope of contract group as the most significant group and environmental and circumstantial had the lowest influence in cost. Worth to mention, for individual factor ranking,

the top three factors found to be: clearly defined scope of project in the contract and cost control, contract dispute, and price fluctuation.

Another contribution by Mulla and Waghmare (2015) who had the same concern about studying the factors influencing the cost estimate in Indian construction industry. The respondents in their study justified the cost overrun by: “an inaccurate estimate of cost, faulty designs, land acquisition problem, poor bidding, irregular flow of finance, delay in payment of work done”.

The effect of contract type was highlighted by Lu et al. (2017). Through applying logistic regression analysis to 14 factors identified through experts interview as critical for cost overrun and statistical analysis of 144 questionnaire survey, the authors were capable to diagnose the owner ability (such as owner team ability, owner experience, and owner finance) as the critical constructs affecting the cost performance in design-bid-build construction contracts executed in China.

In contrast with previous publications, Mahamid et al. (2014) and Badawy (2016) have reported the high ranking of uncommon factors that were found previously having low impact on cost determination such as: “level of competitors in the tendering, material price changes, communications with suppliers, communications with client, estimating method used, amount of specialist work, scale and scope of construction, complexity of design, quantities of expected variation of project, and quality of design and specification”. Also, Arif et al. (2015) reported the significance of some factors that had low effect in previous studies. Using principle component analysis, their research findings acknowledge that involvement of contractors in the estimation procedure and labor rate have a significant effect on project cost bias in Pakistan. In another research, Ali et al. (2018) investigated the uncertainty factors affecting the construction transaction cost in Pakistan. The participants highlighted a low effect of

political and environmental factors on the project uncertainty level, while the technical, commercial and project site are more influential in escalating the transaction cost.

Abd El-Karim et al. (2017) quantified the factors that influence the budget and time contingency in Egypt. Data analysis using heat map tool viewed that the factors having a large probability of occurrence and greater impact on cost are construction area, change in rules and regulations, defect work, type of fund, and team and management experience.

Abusafiya and Suliman (2017) affirmed the critical effect of design and documentation related factors on construction cost overrun in Bahrain, as this category got the highest combined importance index value among the remaining research categories.

Mahamid (2018) found an agreement between client and consultants on the fact that client-related factors dominates the cost influencing factors at the pre-tender stage of public construction projects in Palestine. The owners' ranked the balance groups as follows: contract and tendering factors followed by cost estimating factors. As for the consultants, the remaining group were ranked as follows: cost estimating factors followed by factors related to market conditions.

A similar study was conducted recently in Jordan by Hatamleh et al. (2018) examining the effects of three groups' factors (consultants/contractor/ design parameters and information, project characteristics, and market conditions) with a total of 27 factors on the cost estimation performance. The overall results indicated an agreement between consultants and contractors top five factors ranking, which found to be: a clear scope definition, the project complexity, financial capabilities of the client, site constraints (access, storage, services) and impact of project schedules. Another supporting evidence found by Bakr (2019) work, which

studied the factors affecting the tendering cost of public projects in Jordan. The participants confirmed the influence of client-related factors, especially his financial capability followed by consultants related factors on the accuracy of cost estimation.

In more recent publications, the influence of external, economic and uncertainty scales have been recognized and signified by research participants. For example, Akhund et al. (2019) reported that the exchange rate/fluctuation of prices, material and labor wage escalation (inflation) and financial instability in markets are within the top factors causing cost overrun in Pakistan. Further, the medium effect of projects and country risks discussed by Al-Sabah and Refaat (2019) will not cancel their contribution on the risks affecting public projects performance (cost and time) in Kuwait. An exception has been found by Bakr (2019), where both resources and market conditions, and external and tendering requirements related factors got the lowest ranking in the survey.

More recent research performed in India by Annamalaisami and Kuppuswamy (2019) reported the respondents' high believes on the influence of design-related factors: design/specification changes and delay in regulatory approvals. Other factors that have an influence on project cost overrun indicated as follow: poor selection of contractors, inflation/price fluctuations, and unpredicted weather conditions.

2.4.3 Factors affecting cost estimation accuracy in Europe

The perspective of the project complexity dimension has been introduced and elaborated by Sridarran et al. (2017). Through examining 23 complexity factors, it has been found that project and product size, and time frame and organization capability were among the highest-ranked complexity dimensions that impacts cost estimation. In other hand, factors

such as the political influence, external environment constraint, and environmental and safety impact ranked with low RII.

Likewise, Adafin et al. (2017) indicated risk factors responsible for the cost disparity in fifteen construction projects implemented in New Zealand. The factors listed in the study were as follows: “market conditions, clients change, design enhancements/variations, site investigation information, coordination errors, documentation errors/omissions, incomplete design information, incomplete documentation for cost plan and inadequate tender documentation”.

2.5 Cost Estimation Models

2.5.1 Cost estimation modeling in construction field

Through reviewing the existing literature, it has been observed the domination of establishing construction cost estimation (CE) models with an acceptable accuracy level for different types of infrastructure construction projects using quantitative project characteristics data inputs.

Mainly, Three CE models are repetitively used to capture the cost of building construction projects: multiple regression analysis (MRA) (Gunduz and Sahin, 2015; Coffie et al., 2019), case-based reasoning (CBR) (Tatiya et al., 2018; Leśniak and Zima, 2018), and artificial neural network (ANN) (Juszczyk et al., 2018; Chandanshive and Kambekar, 2019 ; Elmousalami, 2019).

In addition, Construction Cost Index (CCI) has been introduced as a practical method to enhance the construction cost prediction. CCI is established by calculating the weighted average price of construction activities in 20 United States of America cities. Ashuri et al. (2012) highlighted that the current methodology is subjected to unpredictable variations such as money supply and crude oil price, which leads the

researchers to consider time series analysis to predict the CCI more precisely.

Ashuri and Lu (2010) affirmed the accuracy of Seasonal Autoregressive Integrated Moving Average (ARIMA) and Holt-Winters exponential-smoothing models in the forecasting of CCI. However, performing time series analysis was questionable by other researchers. Moon and Shin (2018) argued about the reliability of performing time series analysis on cost data, due to its influence on the country's economic and political conditions. The intervention of these factors was considered in their study. The results showed that the interrupted time-series (interrupted ARIMA) provided better cost forecasts (MAPE =0.3457%) compared to ARIMA (MAPE= 0.4346%) and Holt-Winters exponential-smoothing models (MAPE =0.6231%).

Moreover, in another research, Moon and Shin (2018) found that CCI forecast is more accurate using Vector Error Correction Model (VECM) with selected search queries collected from google trends. This hybrid technique overcomes the need to analyze up-to-date data at the time of CCI forecast limitation. Among 16 VECM models, the best results were found using the query 'project manager salary' with MAPE = 0.3250%.

Looking into recent publications, new technologies are introduced to facilitate the process of construction cost estimation. Reza (2020) advocated beneficial aspects related to Block Chain (BC) technology in construction. The application of BC will aid in monitoring the procurement process, subcontractor performance, construction deliverables, and tracking the quality which eventually will improve the project execution and reduce the risk of project overrun.

These approaches are influential in the field because of data accessibility and authors' objective to automate construction cost

estimation procedures and get maximum benefits from previous experience. Unexpectedly, less attention has been noticed in exploring the intangible factors effect on the cost estimation process.

2.5.2 Modeling influential factors in various fields

Evaluation of performance of an operation is an important step in the service quality cycle. It provides opportunities to improve it. Practically, researchers used the output of questionnaire surveys to find the attributes that influence service performance and respondents overall satisfaction. This approach is adopted widely in marketing and IT research and more attention is being made in the banking and Transportation services field.

Through reviewing the previous literature, it has been observed researchers preference to analyze the customer satisfaction factors using machine learning techniques such as the neural network. The main justification to interpret their choice is the behavior of customer satisfaction, which has been found to have a highly nonlinear relation with the influencing factors that makes the nonlinear techniques more effective in detecting the relationships. For example, Zeinalizadeh et al. (2015) used Multilayer Perceptron artificial neural network (MLP-ANN) and multiple linear regression (MLR) to model factors contributing in bank customer satisfaction. The results admitted that the developed ANN has superior performance in predicting customer satisfaction (73% more than MLR). Also, the developed model used in prioritizing factors based on their weight loading in the ANN input layer.

Sharma et al. (2017) predicted quality determinants of e-learning management systems using two-step: structural equation model (SEM) to identify significant determinants, which were used as an input in the

artificial neural network (ANN) model. The mentioned method overcomes the linearity problem in SEM and enhances the predictions. The method was used by Khan and Ali (2018) in the telecom sector where the researchers investigated the factors affecting the retailer's adoption of mobile payment systems.

In the transportation field, Islam et al. (2016) developed a bus service quality (SQ) prediction model and compared the output of two methods: the probabilistic neural network (PNN) and adaptive neuro-fuzzy inference system (ANFIS). The results showed that ANFIS predicted customer preferences more accurately and better than PNN. The research was further extended to include the ranking of the SQ attributes according to their impact on the overall result from the developed model using the stepwise approach and participant opinion. Another contribution to the bus service quality studies area was done by de Oña et al. (2017). The research assessed the importance of twelve attributes related to bus service quality using decision trees and neural networks. The stability of ANN algorithms was observed against the decision tree model with a high accuracy rate (95% vs. 49.7%). Both models algorithms established the same factors important ranking affecting metropolitan bus public service.

Considering the road design field research, Zuna et al. (2016) used the ANN methodology to model the satisfaction of toll road users using three different models: Customer Satisfaction (SERVQUAL), Minimum Service Standards (SPM), and Toll Road Service Quality (TRSQ) model which was modified by two other standards. The analysis proved that to Toll Road Service Quality (TRSQ) which has been modified by the authors had the lowest values of mean absolute deviation (MAD) and root mean square error (RMSE), and highest coefficient of determination.

Similar work has been published in the health field. Alharthi (2019) employed two modeling techniques: artificial neural network and logistic regression modeling to identify critical determinants influencing physician satisfaction with their institution's electronic medical records (EMR) system. The results proved the capability of ANN to consider the significant factors that truly affecting the satisfaction measurement compared to LR, which ignored 3 factors and confirmed the nonlinearity characteristics of satisfaction variable.

2.6 Chapter Summary

Construction cost estimation is an early decision that has a great contribution to project success or failure. An overview of the past works of literature provided evidence of the construction cost estimation deficiency in the construction field. However, it has been observed lack of researches that benchmark the level of construction cost estimation accuracy in middle east countries. Absent of scientific investigation that evaluates the pattern of construction cost estimation, the effect of project characteristics, and time influence hinder the opportunities of process improvement to the concerned organizations. Diagnosing the level of construction cost estimation accuracy will provide fundamental information to align the current practice with the construction cost estimation level required to be achieved.

Another aspect introduced in the literature is the effect of factors over the accuracy of construction cost estimation. Most of the researchers were oriented toward identifying the critical factors affecting construction cost estimation using relative importance index. Although the attraction RII method got and its wide adoption in the research field, this methodical approach has limited conclusion represented by prioritizing the factors in descending order and test the degree of participant agreement on the

obtained ranking. It is argued that the methodology used in previous literature does not provide a profound understanding of the cost estimation inaccuracy phenomena and cannot be considered as conclusive to answer the question related to the relationship between the factors causing construction cost inaccuracy.

This study addresses the need for exploration research aimed to provide a conceptual framework with the ability to examine the relationship between the factors causing inaccuracy of cost estimation. To our knowledge, no study has yielded a model that answers the outlined objective above. The availability of advance software algorithms is considered as a substantial alternative to analyze the complex relationship between the factors affecting construction cost estimation.

Another promising line of research will be assessing the level of construction cost estimation in MYS in Kingdom of Bahrain and develop a cost estimation model that explore the relationship among the between multiple independent variables affecting the construction cost. The study will enrich Middle East scientific literature associated with the problem and will explore essential information regarding the level of cost estimation accuracy in similar organizations.

CHAPTER 3: Methodology

3.1 Introduction

3.2 Evaluation of Cost Estimation Accuracy

3.3 Cost Estimation Model

3.4 Factors Affecting Construction Cost Estimation

3.5 Questionnaire Survey Analysis

3.6 Modeling Factors Affecting Construction Cost Estimation

3.7 Chapter Summary

CHAPTER 3: Methodology

3.1 Introduction

This chapter contains illustrations for the methods used to evaluate MYS construction cost estimation accuracy, the proposed model to develop a cost estimation model, the factors affecting MYS construction cost estimation, and the proposed model to quantify the factors' influence on construction cost estimation.

3.2 Evaluation of Cost Estimation Accuracy

Early studies used numerous quantitative measures and methodologies to evaluate projects cost estimation accuracy. Skitmore et al. (1991) summarized previous quantitative measures and methodologies used to evaluate projects cost estimation accuracy as given in Appendix 2.

Many researchers used to report the basic evaluation of level of cost estimation accuracy, the validity and conclusions of such reports were argued by Gunner and Skitmore (1999). Gunner and Skitmore (1999) have indicated that meeting statistical tests assumptions and testing the significance of the results are mandatory to accept or reject the hypothesis related to cost deviation and consistency of cost estimation performance. Moreover, they have emphasized the need to define consistent measures for evaluating the quality of forecast performance meeting basics statistical test assumptions before running the inferential tests.

3.2.1 Statistical analysis

Following the tradition in this research topic, similar statistical concepts reported by Skitmore (1991) will be used as measures of forecast quality, the measure is as following:

Bias: which has been defined in ASTM International (Section A.21

Form and Style for ASTM Standards) as: “a systematic error that contributes to the difference between the mean of a large number of test results and an accepted reference value”.

MYS cost estimation process performance will be assessed preliminary through descriptive statistics represented by mean cost deviation and other statistics as per Table 3-1.

Table 3-1 Descriptive analysis performed on Actual construction projects data

Term	Formula
Cost deviation	$\frac{\text{estimated cost} - \text{contract sum}}{\text{contract sum}} * 100\%$
Standard deviation	$\sqrt{\frac{(x - \bar{x})^2}{n - 1}}$
Kolmogorov-Smirnov	Normality test for the cost deviation
Independent Variables	Project size, year of award, building type

Primarily, cost deviation will be calculated for each project and then the parameter means will be used to evaluate MYS project cost estimation accuracy. Positive cost deviation will be interpreted as overestimation, while a negative cost deviation will mean underestimation. Two criteria are used to classify the estimation error for each project. The strict criterion assumes that a project can only be regarded as optimally estimated when the estimated cost completely reconciles with the final awarded cost, thus the estimation error equals to 0%. The more permissive criterion regards a project as optimally estimated with the cost estimation error lies between -10 and +15%, which is an international accepted percentage as per Association for the Advancement of Cost Engineering (AACE)

International Recommended Practice (No. 17R-97) (AACE, 2003).

Adopting Skitmore and Gunner (2000) methodology, inferential statistical tests analysis (student test, analysis of variance) will be applied on the above-mentioned independent parameters to test the significance of the results findings. Statistical significance was set at 5% level. Shapiro-Wilk normality test and Levene's unequal samples size homogeneity test will be checked to ensure meeting ANOVA assumptions prior to implementing statistical tests. Once the results found to be significant, post-hoc tests Tukey's HSD (Honestly Significant Difference) and effect size (partial eta squared) results will be discussed.

3.3 Cost Estimation Model

Regression Analysis is a predictive modelling technique used to investigate the relationship between multiple variables. Analyzing the complex influence of independent variables (predictors) on each other, and their effect on the target (dependent) variable provide better understanding to casual effect relationship between the variables and allows predicting the future performance of dependent variable.

There are different types of regression techniques used to solve prediction problems. Model selection is determined based on the study objective and data characteristics. Most popular types of regression model are linear regression, nonlinear regression, logistic regression, and polynomials regression (Weisberg, 2005). The general expression for multiple linear regression as stated in (Montgomery et al., 2015):

$$Y_0 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon_i \quad \text{Equation 3-1}$$

Where Y_0 is the dependent variable, β_i ($i=0, \dots, n$) are the regression coefficients, x_i ($i=1, \dots, n$) are the independent variables, and ε_i is the random error.

In this study, two regression models will be evaluated to predict project estimated cost. First model independent variables are: project type, awarding year group, and project award amount. The second model will include the percentage cost deviation as an addition independent variable to those of the previous model. A total of 40 actual construction projects will be used to derive the regression model and another four random construction projects will be used to validate the regression model.

The Matlab software will be used to fit the data into a regression model. Matlab has a powerful in-built regression learner application. It trains and validates the regression models using different algorithms' options and machine learning techniques. Regression models' algorithms included in the application are not limited to linear regression models, regression trees, Gaussian process regression models, support vector machines and ensembles of regression trees. Automated regression model training will be used to train multiple models simultaneously.

Table 3-2 Regression models assessment statistics and its decision criteria

Statistic	Title	Description
R-Squared	Coefficient of determination	R-squared compares the trained model with the model where the response is constant and equals the mean of the training response.
MSE	Mean squared error	The MSE is the square of the RMSE.
MAE	Mean absolute error	MAE is quite similar to the RMSE, but less sensitive to outliers.

The regression model's assessment is basically measured by comparing models error Root Mean Square Error (RMSE), where the

model with lowest error will be the best model to predict the required output. Other model statistics are presented in Table 3-2.

The visual model performance will be inspected using the following plots: response plot, predicted versus actual plot, and residual plot. The response plot displays the predicted response versus actual response. The predicted versus actual plot shows the regression model ability to predict outcome using different response values. The residuals plot displays the difference between the predicted and true responses. The developed model will be further assessed and validated using data which was not included in the training set.

Due to the difficulty in extracting the linear regression equation developed in regression learner in Matlab software, alternative software maybe used to develop the mathematical expression of the regression model obtained from Matlab. Examples of software that can be used are Microsoft Excel and Statistical Package for the Social Sciences (IBM SPSS statistics V.25).

3.4 Factors Affecting Construction Cost Estimation

With the intention to achieve the second research objective, numerous factors affecting the construction cost estimation have been pooled from different published literature and dissertations. Next section will clarify the theoretical design of the second research objective instrument.

3.4.1 Literature search on theoretical framework

Former studies have similarity in their factor's selections. Researchers tended to investigate the effect of groups mentioned in Chapter 2 Figure 2-1 on the construction cost. Except for academic dissertations, it has been noticed the absence of logical explanation for factors' influence on construction cost estimation.

To summarize, much of the published literature in this area, the accuracy of construction cost estimation is mainly affected by:

- 1- Quality of activities/ information handled before contract awarding include: scope quality, estimator performance, information quality, and quality of estimation procedure.
- 2- Indirect factors influencing the construction cost include: organization management, uncertainty level, external factors, project characteristics, and contractual procedure.

Research model can be diagrammed as shown in Figure 3-1.

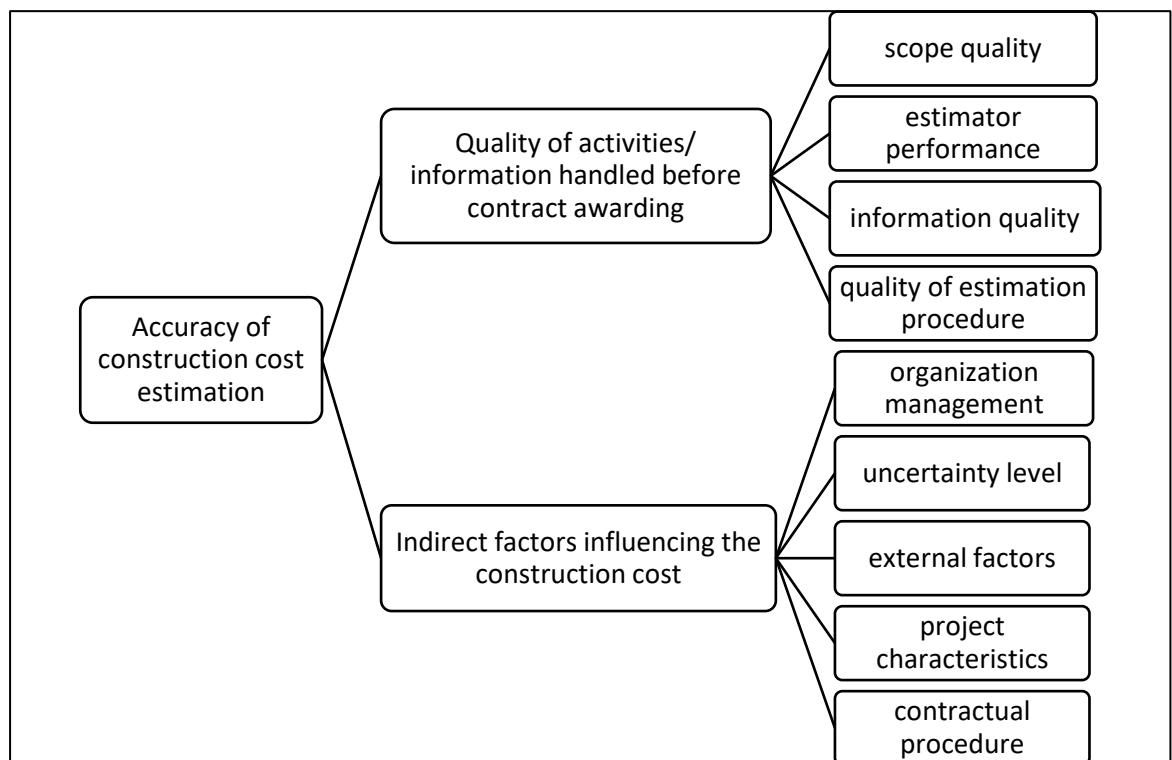


Figure 3-1 Conceptual framework for the categories of factors affecting accuracy of construction cost estimation

The conceptual illustrations for the attributes influence on construction cost estimation are discussed in the following paragraphs.

3.4.1.1 Scope quality

A clear project definition is an essential component for project

resource planning and estimation. Therefore, mismanagement of scope definition will lead to pitfall in project cost appraisal stage and serious project delivery problems. Research by Xia et al. (2016) aimed to examine the relationship between project definition and final project performance through a structural equation model, recognized that project definition has a significant direct impact on project performance. That asserts the need to understood scope by all the parties who have to make decisions throughout the project.

3.4.1.2 Estimator performance

Estimation activity consist of multiple dimensions to be considered as shown in Figure 3-2. Thus, project cost determination will depend on degree of design team integration and their ability to reconcile their experiences to forecast the work cost accurately. Such responsibility required a combination of technical and personnel skills to achieve a reliable estimation. Therefore, inadequate estimation performance may affect the precision of cost estimation.

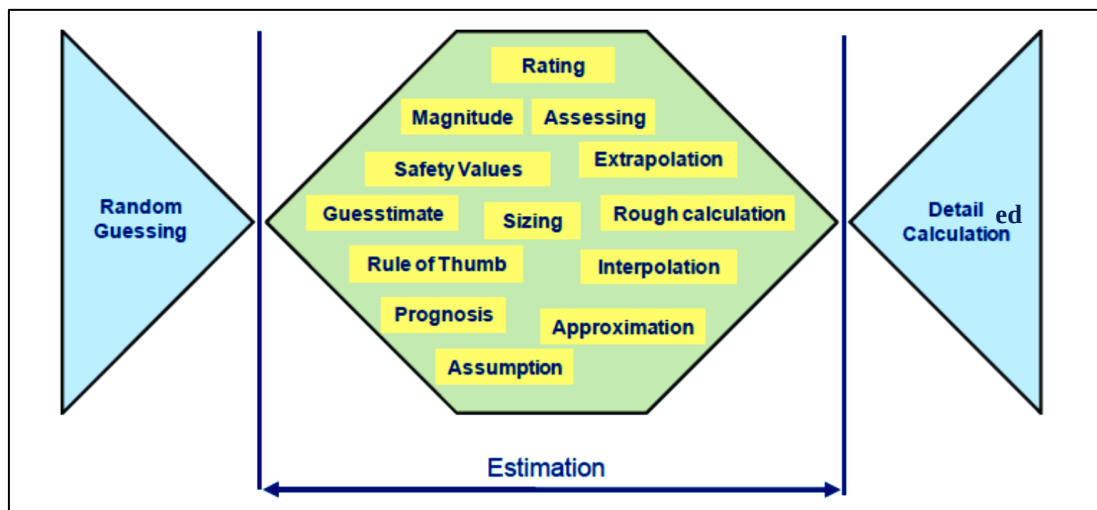


Figure 3-2 Range of estimation activities (Adolphy et al., 2009)

3.4.1.3 Information quality

Throughout the project progressive design phase, the amount of information collected, and design decisions made refines project cost

estimation. Therefore, it is very recommended to maintain a high level of cost data validity and reliability which, per contra, will assist in achieving credible cost estimation.

3.4.1.4 Quality of estimation procedure

Previous research have demonstrated influence of estimation procedure on cost estimation accuracy. Issues such as degree of consultancy office occupancy, amount of drawings and specifications details provided, and time period allowed to prepare cost estimation will have an influence on cost estimation credibility.

3.4.1.5 Organization management

The perception related to the effect of organization management on determining construction cost is not elaborated in detail in previous studies. Generally, management actions and decisions regarding bid prices will be affected by the organization financial performance. For illustration, with a good financial capability, management preferred working in profitable contracts. This is achieved through precisely determining the actual construction cost expenditures and then maximizing their profit margin during bid estimation phase. However, this is not the condition all the time. According to Kerzner (2017), one of the crucial problems in cost control is management tendency to reduce budget or bids to be competitive and increase the chances of winning the bid. Such practices may aid the main contractors to counterbalance their cash flow problems and trying to cover their financial deficiency in the short term with a larger risk to fail in delivering the project.

3.4.1.6 Uncertainty level

The construction industry is highly affected by the uncertain business environment and its influence becomes more significant in

developing countries. Since a considerable portion of construction projects depend on imported materials, economic factors such as inflation rate, currency exchange rate, price fluctuation, and taxation will influence the final construction cost. Therefore, incorporating risk and uncertainties during cost estimation is highly advisable as per AACE Recommended practice 10S-90 (AACE, 2019).

3.4.1.7 External factors

External circumstances (market condition, environment, government) can influence the construction cost through different aspects. Market condition factor has a high contribution to the construction cost. Stating the fact that the combination of direct and indirect construction costs (material, labor, overhead) account for 90% of the construction cost. Thus, inadequate approximation of activities' cost and improper anticipation of human resource allocation during preparing the construction cost estimates will result in a faulty final cost figure.

Other factors that may have an indirect effect on construction cost are environment and government factors. Due to extreme summer weather conditions, GCC countries have regulated the working hours during the summer season. In Kingdom of Bahrain, summer outdoor work ban is activated from 1st of July to 31st of August (12 pm to 4 pm). Subsequently, additional financial expenditures will be accounted on the construction project cost such as labor transportation and double shift wages.

Further, governments services will affect the construction cost when it comes to works way-leaves clearance and resolving unforeseen site condition. Such mandatory coordination may inhibit the approvals process, which ultimately will cause an undesirable change in construction works' cost over time.

3.4.1.8 Project characteristics

Range of construction cost is highly determined by the project characteristics decided during design phase Cho et al. (2009). Roughly, parameters such as project size, project type, location, type of structure, type of external finishing, and quality of finishing will indicate an accurate construction cost. In addition, client decision regarding project priority may constrain the execution time. Hence, construction cost will be influenced by construction fast tracking schedule.

3.4.1.9 Contractual procedure

Procurement system and contract terms and conditions effect on construction cost will underlie the contractual procedure factor. The common procurement practices adopted for government sector projects in the Kingdom of Bahrain is the conventional traditional procurement under the assessment of both Tender Board, which is the regulator of tendering and government procurement, and Ministry of Finance. The procedures bureaucracy involved during bidding to awarding stages causes waste of considerable time frame. Throughout this time, and with a dynamic market condition, the validity of work activities estimated rates are questionable especially if the awarding period extended more than three months.

As well, contractual clauses like performance bond, third party insurance and arbitration and legislation impose unneglectable financial implication on the construction cost (Akanni et al., 2015).

The interaction between the categories is complex, which may influence the level of cost estimation accuracy in different scenarios. Figure 3-3 illustrates an initial conceptual model for the relation between the factors and cost estimation accuracy.

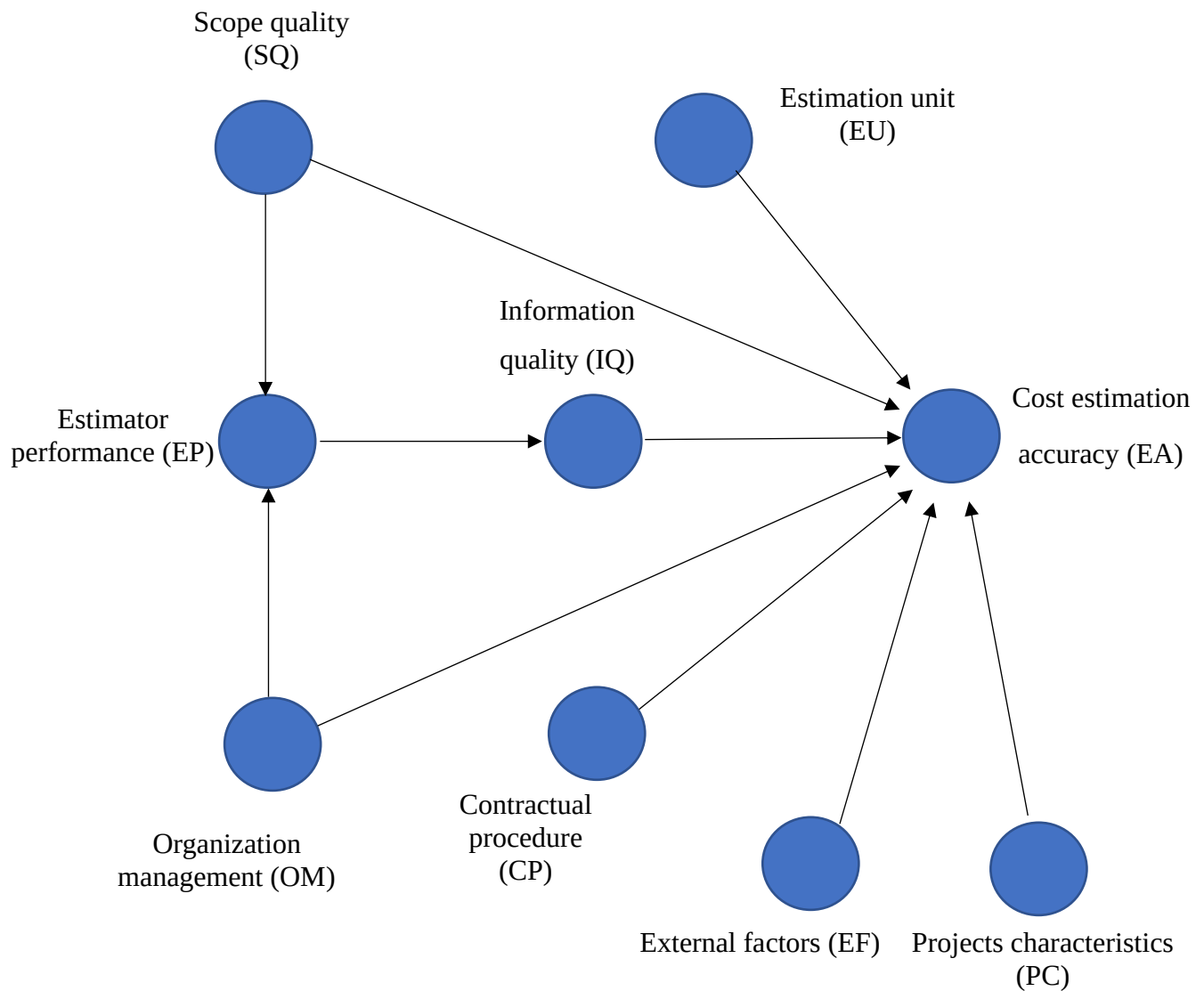


Figure 3-3 An initial conceptual model for the relation between the factors and cost estimation accuracy

3.5 Questionnaire Survey Analysis

3.5.1 Questionnaire design

The tool used to gather information about the critical factors affecting MYS construction cost estimation accuracy was the questionnaire. The target population were the professions working in MYS construction projects between the period of 2012 to 2019. The participants were representing three different entities: client, consultancy office, and main contractor. A detailed explanation for the factor's selection process,

research sample methodology and questionnaire development process will be introduced in chapter 4.

3.5.2 Statistical analysis

Data management started after completing data collection, cleaning and coding the variables in IBM SPSS V.25. The analysis framework has been differentiated through three main stages:

1- Calculate relative importance index: The relative important index (RII) is a mathematical expression used to calculate the score of indicators. RII analysis allows to prioritize the factors based on respondents believes, and it has been used frequently in previous research. RII has been computed to the scale's factors through Microsoft Excel and IBM SPSS V.25 using the following formula:

$$RII = \frac{1n_1+2n_2+3n_3+4n_4+5n_5}{5N} \quad \text{Equation 3-2}$$

Where: n_i ($i= 1,2, \dots, 5$) is the number of respondents for each Likert scale category and N is the total number of respondents. In this work, the criteria decision-making methods used to identify the vital factors affecting MYS construction cost estimation process is adopted from (Akadiri, 2011) importance level as shown in Table 3-3.

Table 3-3 Relative importance index (Akadiri, 2011)

Num	RII Values	Importance level	
1	$0.8 \leq RII \leq 1$	High	H
2	$0.6 \leq RII < 0.8$	High medium	H-M
3	$0.4 \leq RII < 0.6$	Medium	M
4	$0.2 \leq RII < 0.4$	Medium low	M-L
5	$0 \leq RII < 0.2$	Low	L

Aiming to capture the critical factors, it has been decided to consider the factors with an RII above 75% as critical factor in the represented scale and will be engaged in further analysis. The percentage proposed as a cutoff represents the upper quantile of high medium interval, which assumed theoretically will contain factors that highly affecting the cost estimation process. The factors with RII below 74.99% has been eliminated from further analysis.

2-Descriptive and inferential construct stage: The descriptive data analyses will be carried out in the IBM SPSS V.25 for the demographic, work-related, cost estimation practices, participants' perceptions of factors affecting project's cost estimation, and the actual cost discrepancies data. Categorical variables like education level, presence of estimation unit, etc. were summarized using frequency and percentage distribution for each category. Continuous variables like perceptions scales, projects' cost, etc. were analyzed using the central measures tendencies and measures of dispersions (mean, standard deviation, range, etc.).

The unidimensionalities and the reliabilities of the nine scales measuring the participants' perception of factors affecting project's cost estimation were assessed with Cronbach's Alpha measure of internal consistency. To consider a scale with a satisfactory reliability, DeVellis (2016) criteria were used and 0.7 was considered as minimum.

The categorical inferential bivariate correlational analysis was done using the Chi Square test of independence aided by Monte Carlo simulations for better estimation of tests' parameters considering the study's sample size. The continuous inferential bivariate correlational analyses were conducted with the Bayesian version of the non-parametric test when possible, like Bayesian Mann-Whitney test for two samples and

the Bayesian analysis of variance (ANOVA) (i.e., Kruskal Wallis test) between three or more samples all with Monte Carlo simulations. The repeated random sampling has been done for a sample size of 1000 at level of confidence equal to 95%. The criteria of Bayes factor interpretation adopted from Jeffreys (1961) are provided in Appendix 3.

3- Structural equation model: Two popular structural equation models are used widely in research field: conventional SEM and Partial Least Square Structural Equation Model (PLS-SEM). The preference of appropriate model selection depends on research objective and research sample size. Conventional SEM aimed to test an existing theory (confirmatory research) and it is effective in solving complex hierarchical model with large sample size, while PLS-SEM is preferred for developing new theory (exploratory research) and it is appropriate for small sample size. The study data did not fit the conventional SEM methods (with IBM AMOS) and Exploratory Factor Analysis (EFA) due to the limited sample size, the lack of the multivariate distributional normality, and the lack of homoscedasticity. This rendered the data to fit best for the Partial Least Square (PLS-SEM). For the multivariate correlational analysis and structural equation modeling (SEM), WarpPLS 6.0 software was used to specify and estimate the model's relationships' paths.

3.5.3 Partial Least Square Structural Equation Model (PLS-SEM)

Partial least square (PLS-SEM) is a widely accepted method that has been advocated to resolve structural equation modeling in many research disciplines. Unlike the covariance based structural equation modeling (CB-SEM) that deals with theories confirmation, PLS-SEM originally preformed for theories exploration. The method becomes prevalent in the research field attributable to its nonessential of multivariate data

assumptions and its capability to accommodate analysis in small sample size experiments (Willaby et al., 2015). For instance, Kock (2015) got unbiased results by implementation of PLS-SEM analysis for an experiment with a sample size of 50.

Generally, PLS-SEM is a variance-based structural equation modeling technique that create linear combinations of the indicators and subsequently estimate the model parameters using the ordinary least square algorithm (Latan and Noonan, 2017). Like other structural equation modeling technique, PLS-SEM model consist of two elements, structural model which is a diagram that connect the theory constructs and the connecting relationship will be hypothesized, and measurements model that display the relationship between the construct and its indicator variables as illustrated in Figure 3-4.

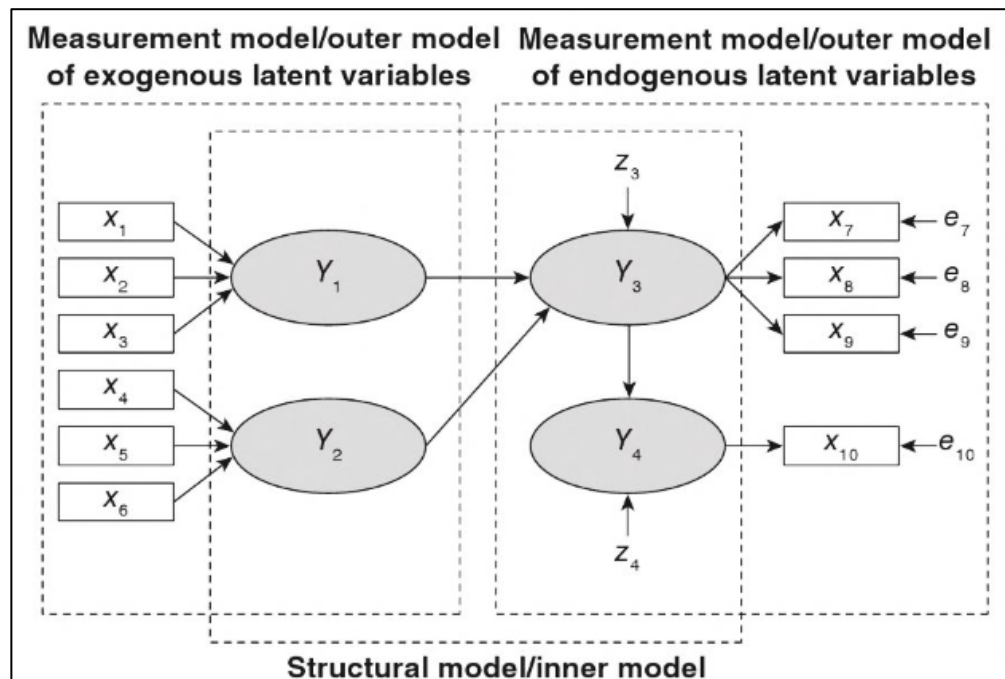


Figure 3-4 A simple PLS-SEM path model (Hair et al., 2013)

The PLS-SEM path model consist of four components displayed in Figure 3-4, where y_i ($i=1, 2, \dots$) represent the model latent variables, Z_i

($i=1, 2, \dots$) are the latent variable errors, x_i ($i=1, 2, \dots$) are the indicators which used to observe latent variables and e_i ($i=1, 2, \dots$) are the error terms related to the indicators. The structural model is defined as:

$$\mathbf{V} = \mathbf{B} \boldsymbol{\gamma} + Z \quad \text{Equation 3-3}$$

Where $\mathbf{V}$ is a vector of all latent variables, $\mathbf{B}$ is a matrix of path coefficients, and Z is the disturbance term error of the dependent latent variables (Sarstedt and Cheah, 2019).

The measurement model can be demonstrated with formative or reflective indicators. The current study implements a reflective measurement model, which is a path model that relate the construct directly with the observed indicators based on the following equation:

$$Y = \sum_{k=1}^{K_{max}} w_{bk} * x_k + z \quad \text{Equation 3-4}$$

Where Y is a latent variable, w_{bk} indicate the contribution of the indicator x_k , where x_k ($k=1, 2, \dots, K_{max}$) to Y , and z is an error associated with Y (Sarstedt and Cheah, 2019).

The algorithm used is aimed to maximize the explained variance of the dependent latent variable (R^2 value) in the path diagram through 5 steps explained by the flow chart of Figure 3-5 adopted from Monecke and Leisch (2012).

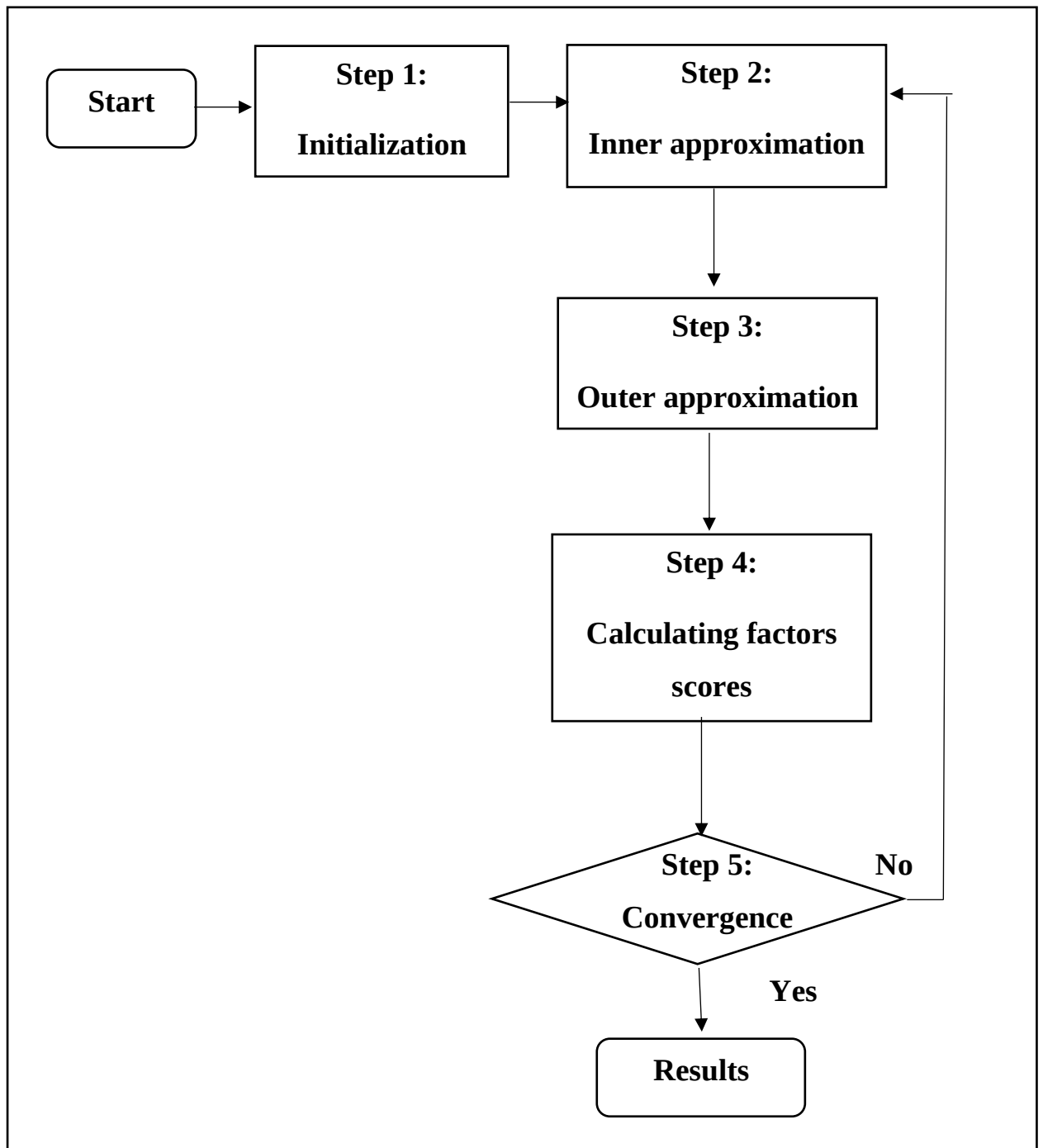


Figure 3-5 The diagram depicts the flowchart for the PLS algorithm (Monecke and Leisch, 2012)

The current study will deal with univariant PLS regression algorithm, which focus to explain a single variable with multiple exploratory variables' using the following algorithm for the extraction which has been reported by Sarstedt et al. (2019) as shown in Figure 3-6.

Initialization

#Stage 1: Iterative estimation of weights and latent variable scores

Starting at step #4, repeat steps #1 to #4 until convergence is obtained.

#1 Inner weights (here obtained by using the factor weighting scheme)

$$V_{jk} = \begin{cases} COV(Y_j; Y_k) & \text{if } Y_j \text{ and } Y_k \text{ are adjacent} \\ 0 & \text{otherwise} \end{cases}$$

Where Y_j is the dependent latent variable, Y_k is the independent variable, V_{jk} is the value of the covariance between $(Y_j; Y_k)$.

#2 Inside approximation

$$\tilde{Y}_j = \sum_i b_{jk} Y_k$$

Where b_{jk} is the inner weight between the adjacent latent variables Y_j and Y_k

#3 Outer weights; solve for

$$\tilde{Y}_{jn} = \sum_{kj} \hat{w}_{bkj} x_{k,jn} + z_{jn}$$

Where $\tilde{Y}_{jn}$ are the latent variable scores from the inside approximation in Step #2, $\hat{w}_{bkj}$ are updated outer weights from Step #3, $x_{k,jn}$ is the raw data for indicator k ($k=1, \dots, K$) of latent variable j ($j=1, \dots, J$) and observation n ($n=1, \dots, N$), z_{jn} is the error term from a bivariate regression.

#4 Outside approximation

$$Y_{jn} = \sum_{kj} \hat{w}_{bkj} x_{k,jn}$$

Stage 2: Estimation of outer weights, outer loadings, and path coefficients

Stage 3: Estimation of location parameters

Figure 3-6 Procedures followed for estimation of weights and latent variable in PLS-SEM model (Sarstedt et al., 2019)

The employment of PLS-SEM with small sample size experiments may confront practical issues with its performance adequacy, which hinders parameters estimation process. Therefore, to overcome the instability occurred, it is recommended to perform nonparametric bootstrapping step aiming to statistically substantiate the results of path coefficients obtained. Sharma and Kim (2013), Efron and Tibshirani (1986), and Davison and Hinkley (1997) explain bootstrapping in more detail.

The context of hypothesis testing in PLS-SEM is evaluated through two main criteria that can be used interchangeably: the p value of the estimated path coefficient and the confidence interval for each path coefficient. The criteria has been explained in details in a prior research (Kock, 2016). With respect to the PLS-SEM model performance, the criteria and the accepted threshold will be explained in detail in the following sub-sections.

3.5.4 Design consideration

Despite the amount of flexibility PLS-SEM has in its assumptions and procedure, many researchers insisted on having considerable sample size. Wong (2019) mentioned the following common practices to determine the sample size: “significance level (5%), statistical power (80%), the minimum coefficient of determination (at least $R^2=0.25$), the maximum number of arrows pointing to the latent variable in the model, and minimum sample size as suggested by Marcoulides and Saunders (2006)”.

WrapPLS software default setting for outer model analysis uses PLS regression algorithm, while for the inner model analysis uses Warp3 algorithm and Stable3 as resampling method. The Warp3 algorithm can detect the relationships among latent variables defined by functions whose first derivatives are U-curves (Kock, 2010). These types of relationships

follow a pattern that is more like an S-curve. Examples of S-curve functions are the sigmoid, hyperbolic sine and hyperbolic tangent (Kock, 2017).

A point to be noted that Stable3 method does not generate resamples, its rather an efficient method that outperform other resampling method available in the software in estimating paths coefficients standard errors.

Other considerations to be investigated in the data set before PLS SEM employment are missing data and number of indicators per scale. Missing data has been classified as a source of bias in PLS-SEM technique (Kock, 2014). The rule of thumb, in case the amount of missing data is less than 15%, it is possible to use the data imputation treatment, otherwise, it is recommended to discharge the items from the analysis to minimize the bias in the model. In addition, PLS-SEM predictive validity is affected considerably when single items used against multi-item scales (Sarstedt et al., 2016). Therefore, a category contains at least one factor will be eliminated from PLS-SEM analysis. Lastly, due to the expected sample size constraint in the study and Hair et al. (2013) recommendation shown in Table 3-4, it is recommended to include maximum three arrows pointing to the model's dependent variable if possible. Limiting the number of arrows will stabilize the SEM performance.

Table 3-4 Sample size recommendation in PLS-SEM for statistical power of 80% (Hair et al., 2013)

Maximum number of arrows pointing to a construct	5% Significance level, Minimum R ²			
	0.1	0.25	0.5	0.75
2	110	52	33	26
3	124	59	38	30
4	137	65	42	33
5	147	70	45	36

3.5.5 Illustrative (conceptual) model

PLS-SEM structural models will be created to test the research hypothesis which will contain the following research latent variables: availability of an estimation unit, scope quality (limited to client and consultant model), organization management (limited to contractor model), estimator performance, information quality, quality of estimation, project characteristics, external factors, contractual procedure, and one endogenous variable called as construction cost estimation accuracy.

The construction cost estimation accuracy computed for each survey participants by using the following expression:

$$CE = \frac{\text{Average Projects budgets}}{\text{Average awarding projects amounts}} \quad \text{Equation 3-5}$$

Equation 3-5 is a simple mathematical expression proposed to measure the precision between participant's projects budgets estimation to the actual awarded projects amounts. Close attention has been made during relating the respondents and the projects they were involved in their construction cost estimation. Almost 70% of the respondents were linked to at least three construction cost projects. The measurement model of the constructs consists of critical factors indicated by participants from the questionnaire survey in a formative mode.

The goal of the model is to explain the effect of independent variables mentioned above on construction cost estimation accuracy. Through model development iterations, some of the scales have been used as mediators for other scales. A mediation variable is a variable that causes mediation in the dependent and independent variables.

Partial regression model estimation in the PLS-SEM algorithm is an iterative procedure. The paths coefficient p value reported in the software is one tailed p value which can be used to calculate two tailed p values

when its required. The p-value shall be lower than or equal 0.05 to be retained in the model. The procedure of construction the PLS-SEM model followed as per Figure 3-7 flowchart.

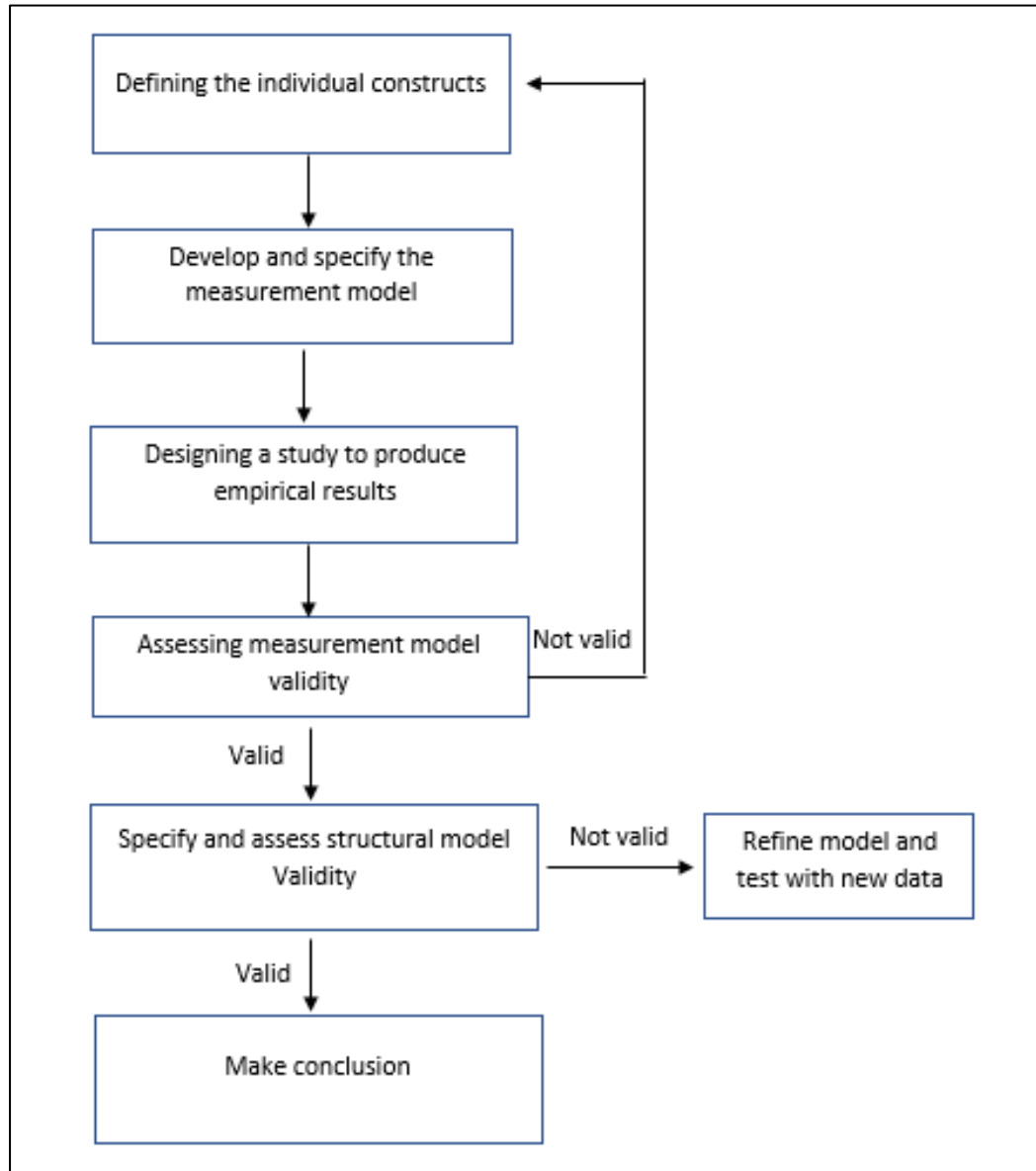


Figure 3-7 Construction PLS-SEM model procedure

The accepted model will be systemically evaluated to examine its predictive capability. The systemic evaluation is mandatory to ensure meeting the quality criteria required for the PLS-SEM as per the recommendation of common practice. Table 3-5 summarizes the quality criteria for each test, the quality criteria has been adopted from Wrap PLS-

SEM Manual (Kock, 2017; Hair et al., 2019). The mathematical expression to calculate the quality terms are provided in Appendix 4 and the definitions of the criteria are mentioned in the glossary page.

Table 3-5 Quality criteria for PLS-SEM

Model		Quality criteria
Measurement model (Reflective)	Internal consistency: composite reliability	≥ 0.6
	Convergent validity	Outer loading > 0.708 Average Variance Extracted (AVE) >0.5
	Discriminant validity	Fornell and Larcker criterion Cross loading
The structural model	Collinearity assessment	VIF < 5
	Path coefficient	$P \leq 0.05$
	Coefficient of determination (only for endogenous latent variables)	$R^2 \leq 0.25$ weak effect $0.25 < R^2 \leq 0.5$ moderate effect $R^2 \geq 0.75$ large effect
	Effect size	$f^2 < 0.02$ no effect $0.02 \leq f^2 < 0.15$ small effect $0.15 \leq f^2 < 0.35$ medium effect $f^2 \geq 0.35$ large effect
	Model predictive relevance (Stone-Geisser) (only for endogenous latent variables)	$Q^2 > 0$

3.6 Modeling Factors Affecting Construction Cost Estimation

3.6.1 Rationale of model development

Data analysis using PLS-SEM technique provides an assessment to the expected relationship between the input and desired output. In order to improve the performance of a process, data modeling is required to address the exact influence of factors affecting the desired output.

The solution proposed to address the factors influencing the cost estimation participant satisfaction is through modeling the key drivers of satisfaction level using artificial neural network. The method has been used previously with great prediction capability. The main objective of the model is to construct neural network that predict the desired satisfaction level with sufficient predictive power. The outperforming model will be used to develop the mathematical equation that expresses the satisfaction level and priorities relevant satisfaction attributes accordingly.

3.6.2 Artificial neural network

Along with the continued improvement of technology, machine learning especially artificial neural network (ANN) was a breakthrough due to its wide application in solving prediction problems in many fields. ANN is a mathematical model that approximates the output through developing a learning algorithm from the arbitrary function (Loy, 2019).

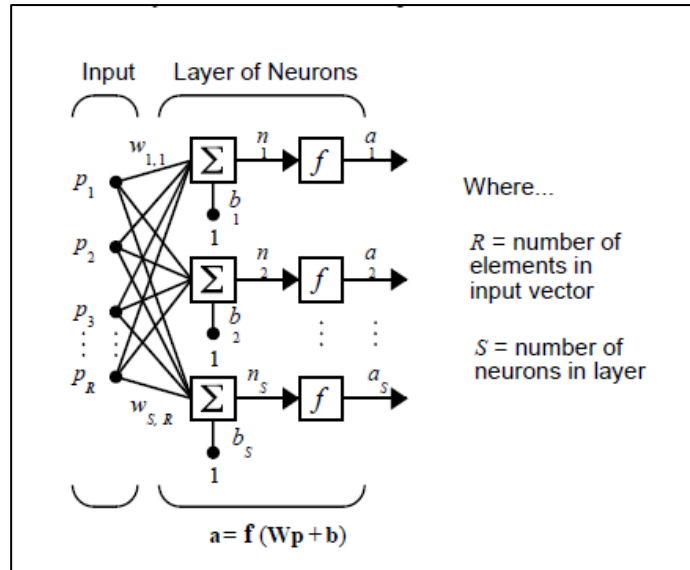


Figure 3-8 A Neural Network Model Architecture (Beale et al., 2010)

Figure 3-8 expresses the ANN model Architecture for one layer network and multi neurons. ANN workflow starts in input node layer and moving in one direction. In a typical neural network, each neuron in the input layer is connected to a neuron in the hidden layer. The signal transition occurs as follow:

1. Develop input product value (strength by weight) and adding the bias to the summing junction.
2. The value in the summing junction acts as argument of the transfer function f , which produces the scalar output a .
3. The scalar output will be transmitted to another hidden layer to predict the output and compare with the target value.

Activation the network will initiate the nodes training using the specified inputs and target values using distinct learning algorithms. The programming software's provides different options to the transfer function and training algorithms which being selected during the model development. The input weights start the training with random initial weight, then they get adjusted continuously in a complex computation and forwarded to the succeeding layer under logical argument till it reaches to

the output layer. The model configuration reached into an end when the significance inputs is defined and the model performance yield to a result similar to the output with minimum error.

ANN has a potent over the traditional approaches through its capability to recognize complex nonlinear relationships and dealing with the uncompleted database. Goh et al. (2005) introduced the mathematical equation that describe the relationship between inputs and outputs in ANN network as follow:

$$Y_{\text{nor}} = f_o \left\{ b_o + \sum_{s=1}^S [w_{so} * f_h (b_s + \sum_{pr=1}^R w_{rs} X_{\text{ni}}^{\text{nor}})] \right\} \quad \text{Equation 3-6}$$

Where,

Y_{nor} = Normalized output variable.

f_o = Transfer function of output layer.

b_o = Bias at the output layer.

w_{so} = Connection weight between the sth hidden neuron and the single output neuron.

f_h = Transfer function of hidden layer.

b_s = Bias of the sth hidden neuron ($s= 1, \dots, S$).

w_{rs} = Connection weight between the P_r input variables ($r=1, \dots, R$) and the sth hidden neuron.

$X_{\text{ni}}^{\text{nor}}$ = Normalized P input variable for category (n) (Normalized in the range of [-1,1] for tan-sigmoid transfer function).

An important step is to standardize the input values between (-1,1) before the mathematical computation of the nodes values using the following formula:

$$X_{ni}^{nor} = \frac{2(X_{ni} - X_n^{min})}{X_n^{max} - X_n^{min}} - 1 \quad \text{Equation 3-7}$$

Where, X_{ni}^{nor} is normalized value of X_{ni} , X_{ni} is actual given input data, X_n^{max} and X_n^{min} are the minimum and maximum data over values of category n. The output can be denormalized using the following expression adopted from Yu and Shih (2015):

$$Y = 0.5(Y_{nor} + 1)(Y^{max} - Y^{min}) + Y^{min} \quad \text{Equation 3-8}$$

Where Y_{nor} is a single value obtained from the neural network model.

Also, symmetric sigmoid transfer function (Tan sigmoid) that will be used in the model is defined as follows:

$$\text{Tanh}(x) = \frac{2}{1 + e^{-2x}} - 1 \quad \text{Equation 3-9}$$

Where x = weighted sum (lumped input) of the input. Tan Sigmoid activation function translates the input in the range (-1; 1).

3.6.3 Model development

In this research, the neural network will be constructed using the model that fulfilled PLS-SEM design considerations. The data set up for the neural network will prepared be as follow: the target variable matrix will contain degree of participant satisfaction on construction cost estimate, where the input variables matrix consists of the factors affecting cost estimation process. The factors will be composed in categories, averaged and used in the analysis.

Matlab 2018a software will be used to create neural network. The nature of the problem prompts using neural net fitting tool application (nftool) to construct the ANN. The application will fit the inputs to predict the target output with arbitrary accuracy. (nftool) uses a standard network

consisting of two-layer feed-forward network (one hidden layer), with a sigmoid transfer function in the hidden layer and a linear transfer function in the output layer. The user must train the network with varying neuron numbers till reaching to acceptable training performance. Optimizing the number of hidden neurons will be done by increasing its number gradually till reaching the optimum model size. Hecht-Nielsen (1989) suggested using the formula $(2i+1)$ as indication to the upper limit of neurons number in hidden layers.

The ANN model design solution will be performed using 70% of the collected data. The remaining data set (30%) will be equally split and use to validate and test the design solution. The training stage solution will continue retraining till validation set performance reached to optimum solution. Three type of training functions will be used in the hidden layers to predict the output: Levenberg-Marquardt, Bayesian generalization and Scaled Conjugate Gradient Algorithms.

Levenberg-Marquardt algorithm typically requires more memory but less time. Training automatically stops when generalization stops improving, as indicated by an increase in the mean square error of the validation samples. Bayesian generalization algorithm typically requires more time, but can result in good generalization for difficult, small or noisy datasets. Training stops according to adaptive weight minimization (regularization). Scaled Conjugate Gradient Algorithms requires less memory. Training automatically stops when generalization stops improving, as indicated by an increase in the mean square error of the validation samples.

The last stage of the process is defining the network accuracy by testing the network. The ANN models performance will be measured in terms of:

- Mean squared error is the average squared difference between outputs and targets. Lower values are better. Zero means no error. The mathematical expression of it is:

$$\text{MSE} = \left(\frac{\sum (Y_m - Y_p)}{N} \right)^2 \quad \text{Equation 3-10}$$

Where Y_m ($m=1, 2, \dots, 50$) is the desired neural network output (satisfaction level), Y_p ($p = 1, 2, \dots, 50$) is the neural network output, and N is total sample size.

- Regression R Values measure the correlation between outputs and targets. An R value of 1 means a close relationship, 0 a random relationship.

$$\text{Coefficient of correlation} = \sqrt{\frac{\sum Y_m^2 - \sum (Y_m - Y_p)^2}{\sum Y_m^2}} \quad \text{Equation 3-11}$$

Also, the best performing model observed will be further assessed using Matlab visual graphs tools such as:

1. Regression plots shows the actual network outputs plotted in terms of the associated target values.
2. Error histogram shows the error sizes distribution. Typically, most errors are near zero, with very few errors far from that.

Finally, the group's connection weights will be extracted from the best performing model and used to develop mathematical equations that facilitate the prediction of participants satisfaction level on cost estimation accuracy using the general expressions of the neural network model using Equation 3-6. Also, the degree of the input of importance will be quantified through employing connection weight method according to the following formula (Gzik et al., 2017):

$$R_p = \sum_{s=1}^S w_{rs} \cdot w_{so} \quad \text{Equation 3-12}$$

Where, R_p is the Relative importance of the input P in reference to the output, s is the number of neurons in hidden layer ($s=1, \dots, S$), W_{rs} is the Connection weight between the input neuron P_r and the hidden neuron s and W_{so} is the Connection weight between hidden neuron s and the output neuron.

Olden et al. (2004) highlighted outperformance of connection weight method among other existing methodologies and its ability to quantify the variables' importance ranking. The final step will be normalizing the importance scores. Pires dos Santos et al. (2019) addressed that this procedure will determine how much the input variables differ in proportion to the input variable that has the highest relative importance. Accordingly, it will allow prioritizing the effect of categories on participants' satisfaction level with construction cost estimation.

3.7 Chapter Summary

In this chapter, the methodologies adopted to determine MYS cost estimation accuracy and determination of factors affecting the construction cost estimation practice were discussed. The traditional approach will be followed in analyzing the cost deviations that occurred in MYS construction projects then regression analysis will be performed to develop cost estimation models. To investigate the factors affecting MYS construction cost estimation and their contribution to construction cost estimation, both partial least square structural equation modeling (PLS-SEM) and Artificial Neural Network (ANN) techniques will be performed to develop a model that examines the relationship between the factors affecting MYS construction cost and construction cost estimation accuracy firstly, then their relationship with participant's satisfaction with estimated construction cost.

CHAPTER 4: Data Collection

4.1 Introduction

4.2 MYS Construction Project Data

4.3 Factors Affecting Cost Estimation Accuracy

4.4 Questionnaire Results

4.5 Chapter Summary

CHAPTER 4: Data Collection

4.1 Introduction

This chapter summarizes the procedures followed in collection of data related to research objectives. The questionnaire design, research sample, and questionnaire results are discussed too.

4.2 MYS Construction Project Data

Prior to collecting the information needed for the study, ethical clearance was sought from MYS. The cost data of MYS construction projects awarded between years 2013 and 2019 were made available, in purpose to investigate the level of MYS construction cost estimation practice accuracy.

The archival documents used are budget reservation letters, the preliminary design reports, the bids evaluation reports, and projects awarding letters. The data extracted from electronic copies of the original documents during the period of July 2019 and October 2019. To ensure the quality and reliability of the data collected, the information sheet has been reviewed and checked by other employees from the contract unit and finance department.

The data consist of: date of forwarding budget reservation letter, amount of budget to be reserved, project type, contract awarding date and contract awarded amount. Another precaution taken is assigning a unique ID to each construction project. This procedure will ensure data anonymity.

There was difficulty in finding a record for the date of budget reservation letters relevant to old construction projects, it has been decided to keep them empty. The percentage of incomplete data concerning budget reservation date is 15% (7 projects out of 45), which is expected and can be considered a low percentage in scientific research. Table 4-1 displays the construction projects relevant data collected.

One information to be mentioned, the construction projects with a cost estimation amount below BD25,000 have been excluded from the study, as it is a small-scale construction works and normally the submitted bids cost variability is very close to the estimated budget.

Table 4-1 Actual MYS construction project data

ID	Project type	Date of contract award	Reserved Budget (BD)	Contract awarded amount (BD)
1	Sporthall	2012	400,000	318894.900
2	Sporthall	2014	450,000	465,000
3	Sporthall	2013	400,000	374,636.273
4	Sporthall	2014	450,000	412,285.500
5	Sporthall	2015	750,000	788,700
6	Sporthall	2017	400,015	456,021.195
8	Playground	2013	150,000	162,801.870
9	Playground	2014	249,700	192,898.156
10	Playground	2014	182,879.400	142,627.012
11	Playground	2014	51,395.300	37,666.464
12	Playground	2014	83,611.00	79,750.000
13	Playground	2015	53,000.00	52,800
14	Playground	2014	81,000.00	81,950
15	Playground	2015	53,000.00	51,981.490
16	Playground	2017	247,800	296,875.700
17	Playground	2016	100,357.400	83,000
18	Playground	2018	236,314	285,835
19	Playground	2018	136,722	168,769.986
20	Playground	2018	241,321	194,075.761
21	Playground	2019	222,248.250	208,357.900
22	Admin	2013	600,000	538,368.380
23	Admin	2013	90,000	61,711.400
24	Admin	2013	160,000	166,954
25	Admin	2013	150,000	139,700
26	Admin	2014	340,000	251,490.146
27	Admin	2015	150,000	240,000
28	Admin	2014	121,000.000	140,800
29	Admin	2014	800,000	711,975

Table 4-1 Actual MYS construction project data (continued)

ID	Project type	Date of contract award	Reserved Budget (BD)	Contract awarded amount (BD)
30	Admin	2016	300,000	406,594.131
31	Admin	2017	457,000	750,000
32	Admin	2018	830,000	1,339,739.445
33	Admin	2018	711,168.150	738,921.750
34	Admin	2016	2,000,000	1,963,347
35	Admin	2018	250,000	286,592.756
36	Admin	2015	2,000,000	1,978,900.223
37	Admin	2015	1,260,000	1,343,700
38	Maintenance	2014	38,500	25,277.900
39	Maintenance	2013	44,000	37,503.250
40	Maintenance	2014	67,000	61,594
41	Maintenance	2017	144,436.600	123,759.471
42	Maintenance	2014	57,000	83,385.145
43	Maintenance	2015	41,000	39,965.145
44	Maintenance	2017	69,400	51,228.800
45	Maintenance	2017	31,790.000	30,116

4.3 Factors Affecting Cost Estimation Accuracy

4.3.1 Instrument design

For the purpose of identifying the factors affecting the cost estimation practices in MYS, the instrument employed to collect the data is the survey. In general, this is an attractive approach because it's a practical and cost-efficient tool that provides actionable information.

The survey structure has been designed to provide information in three main aspects: demographic information and main cost estimation practice, factors affecting construction cost estimation accuracy, and factors affecting the construction cost.

4.3.2 Categories and factors selection

The procedures adopted in selecting factors contain two stages: review of previous literature and validation of factors. Initially, four researches published before 2010 were selected as a reference for factors selection. Collection of recent

literature published between 2010-2019 have been inspected and factors compared with the earlier works. The total number of reviewed papers is 23. It has been concluded through the literature content that all the 23 research shall be retained due to its direct relation to the study objective.

Primarily, the main factors affecting the construction cost estimation have been tabulated. It has been noticed the use of factors with different categories notations. The categories have been formulated with a consistent notation in such a way it reflects the measured construct effectively. The notations have been refined and used as per the Figure 2-1 presented in Chapter 2 (p.17 Section 2.4), to continuance with terms and facilitate results comparison with previous literature.

Through the thematic review, it has been observed many factors have synonyms notation in different literature. Therefore, the similar factors have been abbreviated, refined and linked with the related category. Using multiple factors per category in research is recommended practice, as the research constructs can be considered complex in the determination by one factor. And practically, multiple factors allow measuring the constructs in a more reliable and valid manner (Diamantopoulos et al., 2012). A total of nine categories with 54 factors were crosschecked with an expert for their applicability and validity. Two factors have been eliminated due to not applicability and sensitivity. The final number of factors used is 52 factors. Table 4-2 contains a tabulation of the research categories and the factors used to measure it.

Table 4-2 List of factors affecting construction cost estimation linked to their categories

Category		Factor	
No	description	No	Description
1	Scope Quality (Client)	1	Previous client / consultant experience of contract
		2	Clear scope definition for the client
		3	Level of team integration
		4	Buildability of design
		5	Level of involvement of client in the estimation procedure
		6	Client's priority level of the project
		7	Number of other projects with client
2	Estimator Performance	8	Experience of pricing construction
		9	Impact of team integration and alignment
		10	Estimating method used
3	Information Quality	11	Completeness of cost information
		12	Accuracy and reliability of cost information
4	Quality of Procedure	13	Time allowed for preparing cost estimate is not sufficient
		14	Clear and detail drawings and specification
5	Project Characteristics	15	Type/function (playground, administration building, model youth center, sport hall)
		16	Size/gross floor area
		17	Type of structures (steel, concrete, masonry)
		18	Location (regions; urban)
		19	Quality of finishing
		20	Priority of project
		21	Duration of the project period
		22	Clients financial situation and budget
		23	Inadequate production of raw material for country
6	External Factors	24	Cost of materials
		25	Cost of labors
		26	Cost of machinery
		27	Overhead costs
		28	Governments requirements
		29	Lack of information and coordination between government agencies
		30	Weather

Table 4-2 List of factors affecting construction cost estimation linked to their categories (continued)

Category		Factor	
No	Description	No	Description
7	Uncertainty Level	31	Economic insatiability
		32	Inflationary pressure
		33	Currency exchange
		34	Taxation in imported material
		35	Monopoly
		36	Price fluctuation
8	Contractual Procedure	37	Clients' evaluation and awarding policy
		38	Contract conditions
		39	Time between project announcement, bid submission and award of contract
		40	Availability of other projects for tendering
		41	Performance Bond/warranty
		42	Number of bidders on competitive projects
		43	Contingencies allowed
		44	Tendering duration
9	Organization Management (Contractor)	45	Management team (suitability, experience, performance)
		46	Financial capability
		47	Past relationship with client
		48	Payment record of client
		49	Relationship with subcontractors and suppliers
		50	Experience on similar projects
		51	Past loss/profit in similar projects
		52	Current workload

The influence of scope quality on construction cost estimation is apparent mostly during the design stage. Therefore, it has been decided to limit scope quality factors ranking to client and consultant participants. Similarly, organization management influence on construction cost estimation is apparent mostly during the tendering stage. Therefore, it has been decided to limit organization management factors ranking to participants from main contractors.

The Responding system about the participant perception about the factors affecting estimating accuracy and the factors affecting the construction cost will be indicated based on a 5-point Likert scale, where 1 meant “totally disagree” and 5 stood for “totally agree”.

4.3.3 Validity of instrument

Considering that the factors selected are well established in many research contexts over the years, a professional expert has been consulted to face validate the research constructs content. The pilot study validation step has been excluded in the early stage due to the limitation of the research sample size. Instead, the high rated factors will be cross validated with 10 building construction experts in Ministry of Works and their assigned main contractors. The reliability of the questionnaire will be verified during the survey analysis phase. In case low-reliability values found in the research categories, combining the factors with other categories shall be considered if possible.

4.3.4 Research population and sample

Due to the research specialist in MYS, the most representative sampling technique to collect the data, which will describe the real situation, is nonprobability sampling (convenience or opportunistic sampling). This sampling technique will ensure getting the feedback from the concerned people who dealt and have a clear perception regarding the factors affecting the cost estimation process. The population targeted is 56 participants. The samples divided evenly as follows: 28 participants from client/consultant group and 28 participants from main contractors' group.

Participants will be requested to participate in answering the questionnaire are the client engineers, consultancy offices and main contractors who have a construction contract with MYS in the previous five years. The questionnaire will be distributed among the top management, project managers, architects, site engineers, and quantity surveyors. Gathering survey responses was held over two months period May-June 2019.

Before commencing the study, ethical clearance was sought from MYS. The survey has been distributed to the participants via e-mail. A follow-up call

has been done to ensure filling of the survey. Copy of the distributed survey is in Appendix 5.

The number of participants who answered the questionnaire is 50 participants with a questionnaire response rate of 89%. The percentage response rate is similar for client and consultants' groups (N=12, 86%), and (N=26, 93%) for main contractors.

Despite the sample size recommendation provided by Hair et al. (2013), the research is based on the limited sample size obtained due to construction companies' lockdowns and employees' resignations during the research period. Therefore, the current participants can be considered as representative of the study population.

4.4 Questionnaire Results

The targeted sample actively participated in answering the research questionnaire within the given time limit. Table 4-3 and Table 4-4 include the questionnaire results.

Table 4-3 Factors affecting MYS construction cost estimation (client and consultant perspective)

Category	1.Scope quality							2.Estimator performance			3.Information quality		4.Quality of procedures	
Respondent number	1. Previous client / consultant experience of contract	2. Clear scope definition for the client	3. Level of team integration	4. Buildability of design	5. Level of involvement of client in the estimation procedure	6. Clients priority level of the project	7. Number of other projects of same client	8. Experience of pricing construction	9. Impact of team integration and alignment	10. Estimating method used	11. Completeness of cost information	12. Accuracy and reliability of cost information	13. Time allowed for preparing cost estimate	14. Clear and detail drawings and specification
1	4	3	4	3	4	4	4	4	3	4	4	4	5	4
2	4	4	4	4	3	3	4	4	4	5	4	3	4	4
3	4	5	4	4	4	4	4	4	3	4	5	5	3	4
4	3	3	3	3	3	3	3	3	3	3	3	3	3	3
5	3	5	5	3	4	3	3	5	4	3	4	3	3	5
6	3	4	3	4	4	2	4	4	4	4	4	4	4	4
7	5	3	3	5	5	3	3	4	3	3	3	3	5	4
8	4	4	3	4	5	5	3	5	4	5	5	5	5	5
9	4	4	3	5	3	4	3	4	5	5	5	5	5	5
10	2	3	3	3	2	3	3	3	3	3	2	2	5	3
11	5	5	5	4	5	5	5	5	5	5	5	5	4	5
12	4	5	4	3	4	4	3	4	4	4	4	5	4	5
13	2	5	4	5	5	3	4	5	5	4	5	5	5	5
14	5	2	2	4	5	2	4	3	3	3	3	3	2	3
15	3	5	5	4	5	5	5	3	3	4	3	4	3	3
16	4	4	4	5	4	4	4	5	4	4	5	5	4	5
17	3	5	5	4	4	4	4	4	4	4	4	5	4	4
18	4	5	5	4	5	5	4	5	4	5	5	5	3	5
19	3	5	3	4	4	4	2	4	3	4	5	4	3	5
20	4	5	5	5	5	5	4	4	5	4	4	3	4	5
21	4	5	4	3	2	4	2	4	4	4	4	4	4	4
22	4	4	4	4	5	4	4	5	4	3	4	5	4	5
23	5	5	5	5	5	5	5	5	4	5	5	5	4	5
24	3	4	4	3	4	3	3	3	2	3	3	4	4	4
Frequency	Responses' Frequencies													
Response = 1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Response = 2	2	1	1	0	2	2	2	0	1	0	1	1	1	0
Response = 3	7	4	7	7	3	7	8	5	8	7	5	6	6	4
Response = 4	11	7	9	11	9	9	11	11	11	11	9	6	11	8
Response = 5	4	12	7	6	10	6	3	8	4	6	9	11	6	12

Table 4-3 Factors affecting MYS construction cost estimation (client and consultant perspective) (continued)

Category	5. Project characteristics									6. External factors							7. Uncertainty level						8. Contractual procedures							
Respondent number	15. Type/function (playground, administration building, model youth center, sport hall)	16. Size/gross floor area	17. Type of structures (steel, concrete, masonry)	18. Location (regions; urban)	19. Quality of finishing	20. Priority of project	21. Duration of the project period	22. Clients financial situation and budget	23. Inadequate production of raw material for country	24. Cost of materials	25. Cost of labors	26. Cost of machinery	27. Overhead costs	28. Governments requirements	29. Lack of information and coordination between government agencies	30. Weather	31. Economic insatiability	32. Inflationary pressure	33. Currency exchange	34. Taxation in imported material	35. Monopoly	36. Price fluctuation	37. Clients evaluation and awarding policy	38. Contract conditions	39. Time between project announcement, bid submission and award of contract	40. Availability of other projects for tendering	41. Number of bidders on competitive projects	42. Performance Bond/warranty arrangements	43. Contingencies allowed	44. Tendering duration
1	4	3	4	3	4	3	4	4	3	3	4	4	4	3	2	4	3	4	2	2	2	2	4	4	4	3	4	3	5	5
2	2	2	3	2	3	3	3	3	3	4	4	4	4	2	4	2	4	3	3	4	4	4	4	3	4	3	4	4	3	4
3	4	4	4	4	4	4	4	4	4	4	4	3	3	3	3	2	4	4	3	4	4	2	5	5	5	4	5	4	5	5
4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4	3	3	3	3	3	3	4	4	3	3	3	3	3	3
5	4	4	4	3	5	4	3	3	4	5	4	4	4	4	4	3	3	4	2	2	2	4	3	3	4	3	3	3	3	3
6	3	4	4	4	4	4	4	4	4	4	4	4	4	5	5	3	4	4	2	3	3	3	3	4	4	4	4	4	4	4
7	4	4	4	4	4	3	5	4	3	5	5	5	4	4	4	4	5	5	5	5	4	5	4	4	4	4	4	4	5	5
8	5	4	4	4	5	5	4	5	3	4	3	2	2	3	3	2	3	3	2	2	2	2	3	3	3	2	2	2	2	2
9	4	5	4	4	5	5	5	5	4	4	4	4	4	5	4	3	4	4	4	5	3	3	5	5	5	4	4	5	5	4
10	3	3	3	3	3	4	5	4	2	3	4	4	3	3	3	4	3	3	3	4	3	5	3	3	3	2	3	4	4	5
11	5	4	5	3	5	5	4	5	5	5	4	4	4	4	5	4	4	4	3	4	4	5	3	2	5	4	4	4	4	5
12	4	4	4	3	5	4	5	3	4	5	5	5	5	3	4	4	4	3	3	3	3	4	4	4	3	4	4	3	4	4
13	3	4	5	4	5	4	5	5	5	5	5	5	5	3	4	4	3	4	5	4	5	5	5	5	5	5	4	5	3	4
14	5	5	5	4	5	4	4	4	3	5	4	5	5	4	4	4	3	3	3	3	3	3	4	4	4	4	3	4	4	4
15	5	5	5	5	5	5	5	5	5	5	4	4	4	4	4	5	4	4	3	3	3	3	4	4	4	4	4	4	4	4
16	5	5	5	5	5	5	5	5	5	4	4	5	5	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5
17	4	3	4	5	5	4	3	2	4	4	4	4	3	4	4	3	3	4	3	4	4	4	4	2	4	2	4	2	2	2
18	5	5	5	5	5	5	5	5	5	4	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5
19	5	5	4	3	4	3	3	4	4	4	4	4	3	3	3	3	4	4	3	3	3	4	3	3	4	3	2	3	3	3
20	5	4	4	4	4	4	4	4	5	4	4	4	4	4	4	4	4	4	3	4	4	4	4	4	3	4	2	3	2	3
21	4	3	3	3	4	4	4	5	4	4	4	4	4	3	3	3	4	4	3	3	4	3	3	4	3	3	3	3	3	3
22	5	2	3	1	4	4	4	3	4	4	4	4	4	4	2	2	4	4	1	3	3	4	2	3	2	3	4	4	4	2
23	3	4	5	4	5	5	5	3	3	3	3	3	3	3	3	3	4	3	3	3	2	2	4	4	3	3	5	5	5	5
24	2	3	4	4	2	4	4	3	4	4	4	3	4	3	2	3	3	3	4	5	4	3	3	2	3	4	3	2	4	4
Frequency	Responses' Frequencies																													
Response = 1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Response = 2	2	2	0	1	1	0	0	1	1	0	0	1	1	1	3	4	0	0	4	3	4	4	1	3	1	3	3	3	3	3
Response = 3	5	6	5	8	3	5	5	7	7	4	3	4	6	11	7	8	9	8	13	9	9	7	8	6	8	8	6	7	6	5
Response = 4	8	10	12	10	8	12	10	8	10	13	18	14	13	10	12	11	14	15	4	9	10	9	10	10	9	10	11	9	8	8
Response = 5	9	6	7	4	12	7	9	8	6	7	3	5	4	2	2	1	1	1	2	3	1	4	5	5	6	3	4	5	7	8

Table 4-4 Factors affecting MYS construction cost estimation (main contractor perspective)

Category	2. Estimator performance			3. Information quality		4. Quality of procedures		5. Project characteristics								
Respondent number	8. Experience of pricing construction	9. Impact of team integration and alignment	10. Estimating method used	11. Completeness of cost information	12. Accuracy and reliability of cost information	13. Time allowed for preparing estimate	14. Clear and detail drawings and specification	15. Type/function (playground, administration building, model youth center, sport hall)	16. Size/gross floor area	17. Type of structures (steel, concrete, masonry)	18. Location (regions; urban)	19. Quality of finishing	20. Priority of project	21. Duration of the project period	22. Clients financial situation and budget	23. Inadequate production of raw material for country
1	5	4	4	4	4	5	5	5	5	5	4	5	5	5	5	5
2	4	4	4	4	4	3	3	4	4	4	2	5	5	5	4	5
3	5	5	5	5	5	5	5	5	5	5	5	3	3	3	3	1
4	4	1	4	2	3	5	3	5	5	5	1	2	2	4	3	3
5	5	4	4	5	5	5	4	5	5	5	5	5	4	5	4	4
6	5	3	4	3	5	4	3	3	3	4	3	5	4	4	4	3
7	5	4	4	4	4	4	3	5	5	4	4	5	4	4	4	4
8	4	3	3	4	3	2	3	4	4	3	3	4	4	2	3	2
9	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
10	4	3	5	4	5	4	4	4	3	4	1	4	4	3	5	4
11	5	4	5	5	5	4	5	3	3	3	3	5	4	4	5	5
12	4	3	4	5	4	4	4	5	5	5	5	5	5	5	4	4
13	5	4	4	5	5	4	4	4	5	5	5	5	5	5	4	5
14	3	4	4	3	3	3	3	5	5	5	5	4	4	4	4	2
15	4	4	4	4	4	5	4	3	3	3	3	4	4	4	5	4
16	4	4	4	4	4	4	3	2	3	3	3	3	3	3	1	1
17	5	5	4	5	5	5	4	5	4	5	4	5	4	5	5	5
18	5	5	5	5	5	5	3	5	3	5	5	5	5	5	5	5
19	5	4	5	4	5	4	4	3	4	5	5	5	5	5	4	4
20	5	3	4	4	5	4	3	4	4	5	5	4	4	5	5	3
21	4	3	4	4	4	4	4	4	4	4	1	3	5	4	5	1
22	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5	4
23	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
24	5	3	4	5	5	4	4	5	4	5	4	5	4	4	5	4
25	4	4	4	4	4	4	4	5	5	5	3	2	2	3	5	4
26	3	3	3	3	3	3	3	4	5	5	5	5	5	5	5	3
Frequency	Responses' Frequencies															
Response =1	0	1	0	0	0	0	0	0	0	0	3	0	0	0	1	3
Response =2	0	0	0	1	0	1	0	1	0	0	1	2	2	1	0	2
Response =3	2	8	2	3	4	3	10	4	6	4	6	3	2	4	3	4
Response =4	10	12	17	12	10	13	11	8	8	6	5	6	12	9	9	10
Response =5	14	5	7	10	12	9	5	13	12	16	11	15	10	12	13	7

Table 4-4 Factors affecting MYS construction cost estimation (main contractor perspective) (continued)

Category	6. External factors							7. Uncertainty level						8. Contractual procedures								9. Organization management							
Respondent number	24. Cost of materials	25. Cost of labors	26. Cost of machinery	27. Overhead costs	28. Governments requirements	29. Lack of information and coordination	30. Weather	31. Economic insatiability	32. Inflationary pressure	33. Currency exchange	34. Taxation in imported material	35. Monopoly	36. Price fluctuation	37. Clients evaluation and awarding policy	38. Contract conditions	39. Time between project announcement, bid submission and award of contract	40. Availability of other projects for tendering	41. Number of bidders	42. Performance Bond/warranty arrangements	43. Contingencies allowed	44. Tendering duration	45. Management team (suitability, experience, performance)	46. Financial capability	47. Past relationship with client	48. Payment record of client	49. Relationship with subcontractors and suppliers	50. Experience on similar projects	51. Past loss/profit in similar projects	52. Current workload
1	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
2	5	4	5	4	5	4	3	3	4	2	3	3	3	4	5	4	4	4	5	5	3	4	4	4	3	4	4	3	4
3	3	3	3	5	5	5	5	5	5	5	5	5	5	3	5	3	3	5	3	3	3	5	5	5	5	5	4	5	4
4	4	3	3	3	4	2	2	3	3	3	4	3	5	4	4	2	2	3	3	3	2	4	4	5	2	2	4	3	2
5	5	5	4	5	5	1	3	4	4	2	2	1	4	4	5	4	5	4	5	5	5	4	4	4	3	4	5	4	2
6	4	4	3	3	3	4	2	4	5	3	4	4	4	3	4	3	3	3	2	3	3	4	5	3	4	4	5	2	4
7	4	4	4	4	4	4	4	4	4	4	5	4	4	4	4	4	4	4	4	4	4	4	5	5	4	4	4	4	3
8	3	3	3	1	3	2	1	3	3	3	3	4	2	3	4	2	3	4	3	1	2	4	4	4	4	4	4	3	3
9	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	3	3	3	4	3	3	4	5	5	5	4	4	4	3
10	5	5	4	4	3	3	3	4	4	2	2	2	3	4	4	3	4	3	5	4	3	4	4	5	5	4	5	5	4
11	5	5	5	5	4	4	3	4	4	5	5	5	4	4	4	4	3	3	4	3	3	5	5	5	5	4	5	4	3
12	5	5	5	5	4	4	3	4	4	4	5	3	4	4	5	3	3	4	5	3	3	5	4	4	5	5	5	4	3
13	5	5	5	5	4	4	4	4	4	5	5	4	5	4	5	4	4	4	5	4	4	5	4	4	5	4	5	5	4
14	4	4	4	4	4	4	4	4	4	4	4	4	4	3	4	3	3	3	2	4	3	4	3	3	4	4	4	4	3
15	5	5	4	4	3	5	4	4	4	1	2	5	4	1	1	4	1	5	1	3	3	4	4	5	5	5	3	4	4
16	3	3	3	3	2	2	3	3	3	2	2	2	2	2	2	2	2	1	1	1	4	5	2	5	2	4	5	4	3
17	2	4	4	4	4	4	2	4	4	4	4	5	4	3	4	4	4	3	5	5	3	1	1	5	3	2	1	3	5
18	5	5	5	5	5	5	5	5	5	3	5	3	5	3	3	5	3	5	1	3	5	4	5	5	5	5	5	5	5
19	3	3	3	3	4	3	3	2	3	3	2	3	3	4	4	4	4	4	4	4	4	5	5	5	5	5	5	3	4
20	5	5	4	4	4	3	5	4	4	3	5	3	5	4	4	5	4	5	3	4	5	5	3	4	4	3	5	4	3
21	5	5	5	3	3	5	3	3	5	3	5	5	3	3	5	5	3	5	3	3	3	2	4	1	3	3	2	4	1
22	5	5	5	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	3	5
23	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5
24	5	5	5	5	3	2	2	3	3	2	2	3	2	3	4	2	3	4	2	1	3	5	4	5	5	4	4	4	2
25	3	3	4	5	1	1	5	1	1	3	4	3	4	3	4	3	4	3	3	4	4	3	4	5	5	5	5	5	3
26	5	5	5	5	5	3	3	3	3	3	5	3	3	4	4	4	4	4	4	4	4	5	3	3	4	3	5	3	3
Frequency	Responses' Frequencies																												
Response =1	0	0	0	1	1	2	1	1	1	1	0	1	0	1	1	0	1	1	3	3	0	1	1	1	0	0	1	0	1
Response =2	1	0	0	0	1	4	4	1	0	5	6	2	3	1	1	4	2	0	3	0	2	1	1	0	2	2	1	1	3
Response =3	5	6	6	5	6	4	9	7	6	9	2	9	5	9	1	7	10	8	6	9	12	1	3	3	4	3	1	7	10
Response =4	5	6	9	9	10	9	5	12	12	5	6	6	10	12	14	9	9	9	5	8	6	11	11	6	6	12	8	12	7
Response =5	15	14	11	11	8	7	7	5	7	6	12	8	8	3	9	6	4	8	9	6	6	12	10	16	14	9	15	6	5

The questionnaire results suggest a common agreement between client, consultancy office and main contractor participants on the possible effect of some factors on MYS construction estimation. The ranking of some categories' factors such as: scope quality, estimator performance, information quality, and organization management appeared high compared to remaining categories like contractual procedure and uncertainty level. The questionnaire results suggest existence of factors that may affect MYS construction cost estimation critically, which needed to be further analyzed and investigated.

4.5 Chapter Summary

This chapter demonstrated the procedures followed in collecting data related to research objectives followed by a review of the questionnaire design, research sample. Lastly, an overview of the questionnaire results obtained from the research study has been illustrated.

CHAPTER 5: Data Analysis

5.1 Introduction

5.2 Analysis of Actual Cost Data

5.3 Development of Construction Cost Estimation Model

5.4 Analysis of Factors Affecting Cost Estimation Accuracy

5.5 Modeling of Cost Estimation Accuracy using PLS-SEM Technique

5.6 Satisfaction with Construction Cost Estimation Model

5.7 Chapter Summary

CHAPTER 5: Data Analysis

5.1 Introduction

This chapter aims to report the findings of construction cost estimation accuracy and the factors affecting the cost estimation process in MYS. The chapter will report the hypothesized results related to the relationship between the factors affecting cost estimation accuracy using PLS-SEM for three different perspective models: client and consultant, contractors, and common categories. Furthermore, applying ANN modeling for the selected PLS-SEM model will be reported too.

5.2 Analysis of Actual Cost Data

5.2.1 Descriptive statistics

The data set consist of 44 construction projects. The distribution of projects based on their function are shown in Table 5-1. Majority of the projects are fallen within administration buildings category with a total percentage of 36.36 (N=16), followed by playground projects 31.82% (N=14). Other construction projects are the maintenance and sport halls projects with percentages of 18.18 (N=8) and 13.64 (N=6), respectively.

Table 5-1 Distribution of construction projects based on project function

Num	Project function	Count	Percentage
1	Administration building	16	36.36%
2	Maintenance	8	18.18%
3	Playground	14	31.82%
4	Sport hall	6	13.64%
	Grand Total	44	100.00%

Another criterion used to evaluate the project is the project size and the year of award. Project size has been evaluated through their total amount. The projects distribution as per project size has been tabulated as shown in Table 5-2.

Table 5-2 Distribution of construction project based on project size

	Project size (BD)	N
1	Less than BD225,500	24
2	Between BD225,000 - BD625,000	12
3	Above BD625,000	8
	Grand total	44

With regard to project year of award, four periods are covering the projects awarding period, starting from 2012 to 2019. The period (2014-2015) has shown the most awarding contract with a total of 20 construction projects being awarded. Remaining periods total awarded projects are in a range from (7 to 9) construction projects as shown in Table 5-3.

Table 5-3 Distribution of construction projects based on year of award period

	Awarding period	N
1	2012-2013	8
2	2014-2015	20
3	2016-2017	9
4	2018-2019	7
	Grand Total	44

A comparison between approved budget with contracts awarded amounts descriptive statistics shows a positive variation with respect to awarding amounts statistics as shown in Table 5-4. The difference between the mean estimated costs and the mean awarded contract amounts is BD16,247.569 (4.57%) with a standard deviation of 21,354.677 (4.76%).

Table 5-4 Comparison between projects estimated costs and contract awarded amounts

Variable	N	Mean (BD)	Std. Deviation
Projects estimated costs	44	355,719.502	448,919.588
Projects awarded amounts	44	371,967.071	470,274.265
Difference		16,247.569	21,354.677
% (difference / Projects estimated costs)		4.57	4.76

5.2.2 Cost deviation analysis

Determination of the accuracy of the projects cost estimation done through computing the cost deviation of each project. A detailed descriptive data analysis including projects estimated cost, projects approved contract amounts, and cost deviation for the projects is given in Appendix 6.

Table 5-5 summarizes projects cost deviation descriptive statistics. According to the strict criterion (0% cost deviation), the vast majority (n=26, 59.1%) of the projects' costs were overestimated and the other were underestimated (n=18, 40.9%). The range of cost deviation found between (-39.07%, 52.31%). The mean cost deviation found to be 3.47% with a standard deviation of 21.78.

Table 5-5 Construction projects cost deviation statistics

	overall cost deviation	Underestimate cost deviation	overestimated cost deviation
Mean (%)	3.47	-16.43	17.25
Standard Error	3.28	3.037	2.91
Median	2.27	-13.41	11.90
Standard Deviation	21.78	12.88	14.84
Sample Variance	4.74	1.66	2.20
Kurtosis	-15.32	-85.17	-28.55
Skewness	5.78	-69.36	81.94
Minimum (%)	-39.07	-39.07	0.38
Maximum (%)	52.31	-1.16	52.31
Count	44	18	26
Percentage	100%	40.9%	59.1%

From the short review of construction projects cost deviation results, key findings emerge:

- 1- Underestimated projects: The range of cost deviation found between (-39.07%, -1.16%). The mean cost deviation found to be -16.43% with a standard deviation of 12.88.
- 2- Overestimated projects: The range of cost deviation found between (0.38%, 52.31%). The mean cost deviation found to be 17.25% with a standard deviation of 14.84.

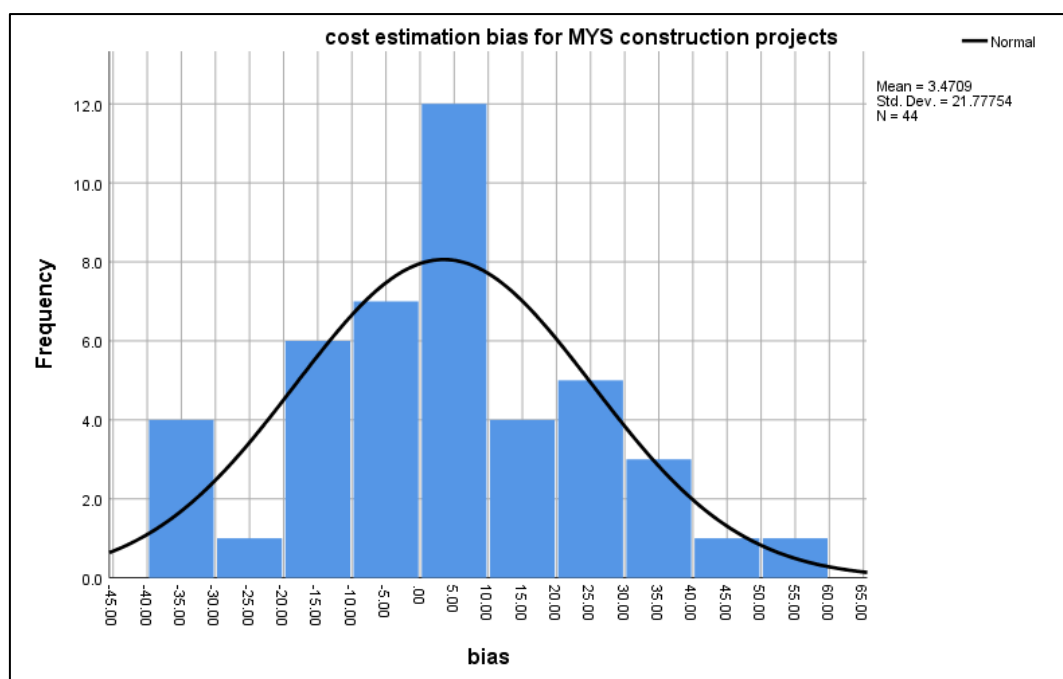


Figure 5-1 Histogram distribution for construction projects' cost deviation

Figure 5-1 shows the histogram distribution for the projects' cost deviation. One sample Kolmogorov-Smirnov normality test has been conducted to test that the cost deviation variable follows the normal distribution. Kolmogorov-Smirnov Normality test statistics indicates that the cost deviation follow a normal distribution, $D(43) = 0.079$, $p = 0.200$. Table 5-6 summarizes the tests results.

Table 5-6 Tests results for Kolmogorov-Smirnov Normality test for overall cost deviation

Variable	Kolmogorov-Smirnov		
	Statistic	df	Sig.
Costs deviation	0.079	44	0.200

In the same way, underestimated and overestimated cost deviations have been tested for normality. The Kolmogorov-Smirnov normality results in Table 5-7 shows that overestimation group, $D(25) = 0.169$, $p = 0.054$, and the underestimation group, $D(17) = 0.144$, $p = 0.200$, were both normal, indicating that the data was normally distributed in both groups.

Table 5-7 Tests results for Kolmogorov-Smirnov Normality test for underestimation and overestimation cost deviation

Estimation type	Kolmogorov-Smirnov		
	Statistic	df	Sig.
Underestimation	0.144	18	0.200
Overestimation	0.169	26	0.054

Classifying the cost deviation based on Permissive criterion (-10% to +15% cost deviation) is shown in Table 5-8. The percentage of optimally estimated project is (47.7%, N=21). According to this classification, the number of under-estimated projects is slightly higher than those were overestimated (12 vs. 11).

Table 5-8 Classifying the cost deviation based on Permissive criterion (-10% to +15% estimation error)

Criteria	N	%
Under-estimated project	12	27.3%
Optimally estimated project	21	47.7%
Over-estimated project	11	25%

5.2.3 Means test for cost deviation difference

To investigate the statistical significance of the mean costs deviation results from zero, one sample t-test has been calculated. The results in Table 5-9 shows no significant difference between overall cost deviation and zero; $t(43) = 1.057, p = 0.296$. However, t-test analysis for overestimation and underestimation cost deviation means individually found to be statistically significant from zero with $t(17) = -5.409, p = 0.000$ and $t(25) = 5.926, p = 0.000$ respectively.

Table 5-9 Student t-test results for overall, underestimation, and overestimation of cost deviation

One-Sample Test						
Variable	(Test Value = 0)					
	t	df	Sig. (2-tailed) (p)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Overall costs deviation	1.057	43	0.296	3.471	-3.150	10.092
Underestimate cost deviation	-5.409	17	0.000	-16.430	-22.838	-10.021
Overestimate cost deviation	5.926	25	0.000	17.248	11.254	23.242

5.2.4 Classification of cost deviation

Categorizing the projects as per project function, project size and year of award in Table 5-10, the following main findings found:

1. With respect to project function, the sport hall had the least cost deviation ($\bar{x}$ =3.48%, SD=13.31) and playgrounds projects which has a noticeable cost deviation ($\bar{x}$ =6.53%, SD=18.59). In other hand, the building maintenance projects had the highest cost deviation ($\bar{x}$ =13.38%, SD=24.6). The exception found in the building administration category whose mean cost deviation tended to be underestimated with a high disparity ($\bar{x}$ =-4.16%, SD= 24.36).
2. Considering the project awarding period, the mean cost deviation for 28 projects (63.64%) performed between (2012-2013, 2014-2015) had a positive value of (12.77%, 6.20%). In contracts, the budget allocation in following years (2016-2017, 2018-2019) faced considerable cost underestimation, as a total of 16 projects (36.36%) performed during that

period were deviated from the allocated budget with a mean cost deviation of (-1.51%, -8.55%).

3. With respect to the project size, the performance of construction cost estimation found to be overestimated with projects below BD 225500 and underestimated with the projects above this amount.

Table 5-10 Descriptive analysis for construction projects cost deviation based on project function, awarding period and project size

Factors	N	Cost deviation (%)	St. Deviation
Project function			
Sport hall	6	3.489	13.331
Administration buildings	16	-4.166	24.363
Playground	14	6.525	18.594
Maintenance	8	13.386	24.656
Total	44	3.471	21.777
Awarding period			
2012-2013	8	12.769	17.130
2014-2015	20	6.201	22.036
2016-2017	9	-1.509	24.047
2018-2019	7	-8.554	20.043
Total	44	3.471	21.777
Project size			
less than BD225,500	24	11.137	20.431
between BD225,000 - BD625,000	12	-3.155	21.358
above BD625,000	8	-9.589	18.794
Total	44	3.471	21.777

5.2.5 Inference analysis (ANOVA) for costs deviation

The statistical effects of following factors: project size, project function, awarding period were explored by using parametric test; One-way ANOVA test as the sub-groups met main assumptions for analysis of variance which are: normality (Shapiro-Wilk normality test (“Sig.” > 0.05), homogeneity (Levene's test “Sig.” > 0.05). The null hypothesis being tested is that the population means are exactly equal. Initially, inferential statistical test results for each subgroup were summarized in Table 5-11.

Table 5-11 Inferential statistical test (ANOVA) results for cost deviation based on project function, awarding period and project size

Factors Function	Mean (%)	St. Deviation	Normality significance	Homogeneity significance	One-way ANOVA	Significance (p)
Sport hall	3.489	13.331	0.7518	$p > 0.6$	1.330	0.2779
Administration buildings	-4.166	24.363	0.4728			
Playground	6.525	18.594	0.3178			
Maintenance	13.386	24.656	0.7469			
Awarding period						
2012-2013	12.769	17.130	0.6078	$p > 0.7$	1.5112	0.2263
2014-2015	6.201	22.036	0.4706			
2016-2017	-1.509	24.047	0.9789			
2018-2019	-8.554	20.043	0.9520			
Project size						
PS< BD225,500	11.137	20.431	0.9753	$p > 0.7$	3.9599	0.0267
BD225,000<PS < BD625,000	-3.155	21.358	0.9412			
PS>BD625,000	-9.589	18.794	0.0490			

As shown in Table 5-11, the only parameter that had a statistically significant difference with error is the project size $F(2,41) = 3.9599, p = 0.0267$. This study did not demonstrate any effect from project function, $F(3,40) = 1.330, p = 0.2779$ and year of award, $F(3,40) = 1.5112, p = 0.2263$.

Since project size factor shows a statistical significance result, post hoc tests represented by Tukey's HSD (Honestly Significant Difference) and estimated effect size have been run. A multiple comparison between 3 mean groups as shown in Table 5-12 indicate that there is a significant difference between the small project and large project error means "Sig." < 0.05 .

Table 5-12 Tukey HSD test results for project size variable

Tukey HSD						
Size of project		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
(I)	(J)					
less than BD225,500	between BD225,000-BD625,000	14.291	7.219	0.130	-3.262	31.845
	above BD625,000	20.725*	8.335	0.044	-0.457	40.994
between BD225,000-BD625,000	less than BD225,500	-14.291	7.219	0.130	-31.845	3.262
	above BD625,000	6.434	9.319	0.770	-16.227	29.095
above BD625,000	less than BD225,500	-20.725*	8.335	0.044	-40.994	-0.457
	between BD225,000-BD625,000	-6.434	9.319	0.770	-29.095	-16.227

The estimated effect size results are tabulated in Table 5-13. Sums of squares for project size represent the variance in cost deviation values accounted by project size. Sums of squares error represents the variance in cost deviation values not accounted for by project size. Partial eta squared (η_p^2) has been derived

by dividing the sums of squares for project size divided by the corrected total sums of squares ($3301.572 / 20393.241 = 0.162$). The aforementioned information indicated that the different project sizes administered account for some 16.2% of the variance in the cost deviation score.

Table 5-13 Effect size for Tukey HSD test result

Tests of Between-Subjects Effects						
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	3301.572 ^a	2	1650.786	3.960	0.027	0.162
Intercept	10.322	1	10.322	0.025	0.876	0.001
Size of project	3301.572	2	1650.786	3.960	0.027	0.162
Error	17091.669	41	416.870			
Total	20923.315	44				
Corrected Total	20393.241	43				

a. Partial eta square (η^2) = 0.162

5.2.6 Time effect on cost deviation

Sensitivity of construction estimation cost toward time imposes the need to investigate the duration spent between project budget request and contract awarding date. The duration of 38 projects were used to describe the current situation. Table 5-14 summarizes the findings.

Table 5-14 Categorization of projects based on duration spent between approving the budget and awarding the contract

	Duration	Mean (days)	N	Std. Deviation	Minimum	Maximum
1	Duration < 1 year	216.500	18	81.542	98.00	336.00
2	1 year < duration < 2 years	555.615	13	107.334	378.00	696.00
3	duration > 2 years	911.285	7	142.625	745.00	1168.00
	Total	460.500	38	284.036	98.00	1168.00

It has been observed that large portion of project (47.37%, N=18) were estimated and executed within one year, with an average lagging of 216.5 days. On the other hand, 52.63% of the projects may be affected by the dynamic market condition and prices change. For more illustration: 34.21% of the projects spent

an average of 555.61 days till contract been awarded (N=13), and the duration of the remaining projects (18.42%) exceeded two years.

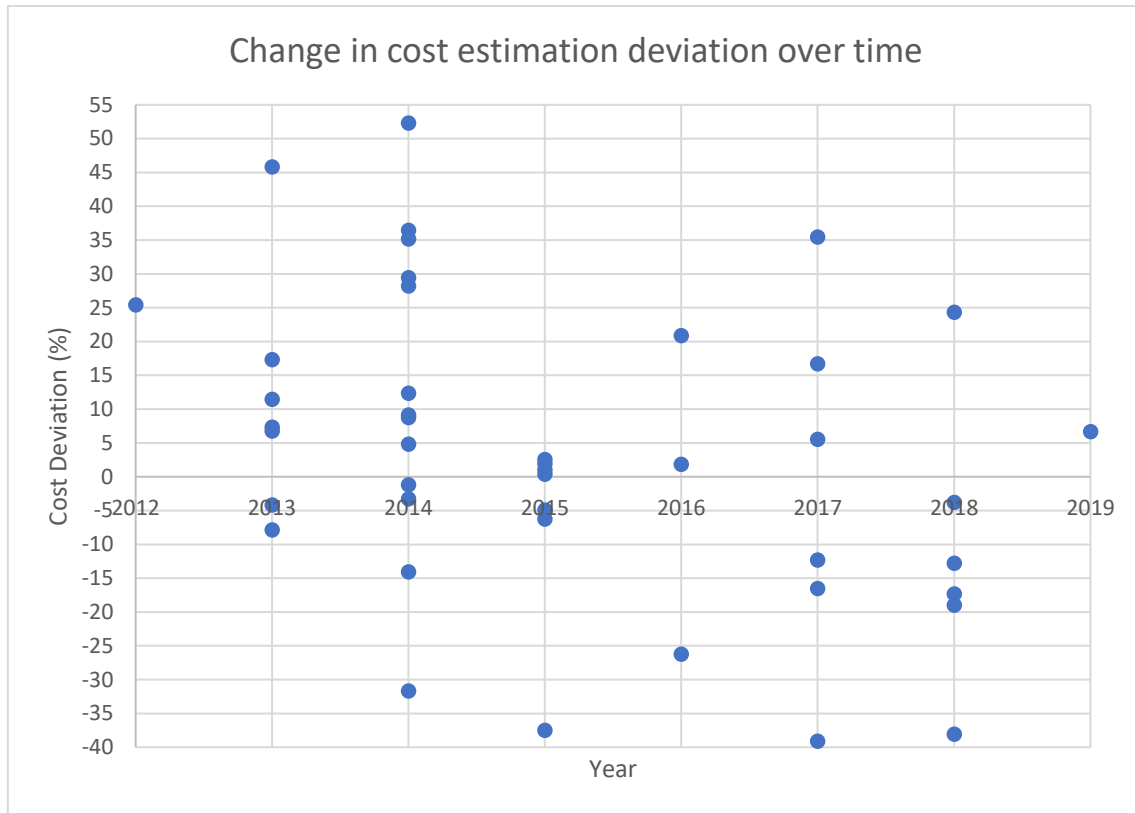


Figure 5-2 Pattern of cost estimation deviation between (2012-2019)

Table 5-15 Mean cost deviation for the period between (2012-2019)

Num	Year of award	Mean	N	Std. Deviation
1	2012	25.433	1	...
2	2013	10.961	7	17.658
3	2014	12.819	13	23.054
4	2015	-6.091	7	14.266
5	2016	-1.145	3	23.708
6	2017	-1.690	6	26.462
7	2018	-11.090	6	20.688
8	2019	6.666	1	...

Plotting cost deviation over time (Figure 5-2), the clustered of costs deviation observed are gradually changing from overestimation to underestimation. Table 5-15 supports the previous finding as the mean cost

deviation has been shifted from a positive cost deviation in years (2013 and 2014) to a negative deviation in the years after (2015 to 2018).

5.3 Development of Construction Cost Estimation Model

5.3.1 Regression model development

Microsoft excel and SPSS software regression model results were not accurate with a large error margin as shown in Appendix 7. However, Matlab software outperformed Microsoft excel and SPSS software. Figure 5-3 shows the history list for model one regression analysis that contains the trained models overall score. Best solution detected for model one, where estimated cost is function of (project type, year of award group, award amount), by using robust linear regression (model 1.3) with RMSE=1.0555E+05. Other model statistics are found as following: R-Square= 0.95, MSE= 1.1141E+10, and MAE= 59606.

Data Browser		
▼ History		
1.1	☆ Linear Regression	RMSE: 1.1315e+05 Last change: Linear 3/3 features
1.2	☆ Linear Regression	RMSE: 1.2768e+05 Last change: Interactions Linear 3/3 features
1.3	☆ Linear Regression	RMSE: 1.0555e+05 Last change: Robust Linear 3/3 features
1.4	☆ Stepwise Linear R...	RMSE: 1.2435e+05 Last change: Stepwise Linear 3/3 features
1.5	☆ Tree	RMSE: 2.602e+05 Last change: Fine Tree 3/3 features
1.6	☆ Tree	RMSE: 3.475e+05 Last change: Medium Tree 3/3 features
1.7	☆ Tree	RMSE: 4.679e+05
▼ Current Model		
Results		
RMSE	1.0555e+05	
R-Squared	0.95	
MSE	1.1141e+10	
MAE	59606	
Prediction speed	~1400 obs/sec	
Training time	10.876 sec	

Figure 5-3 History list for robust linear regression- model 1

The model exhibits considerable prediction error between the true data and predicted data as shown in Figure 5-4. The difference between the true data and its associated prediction is between zero to an approximate 0.5×10^6 maximum error.

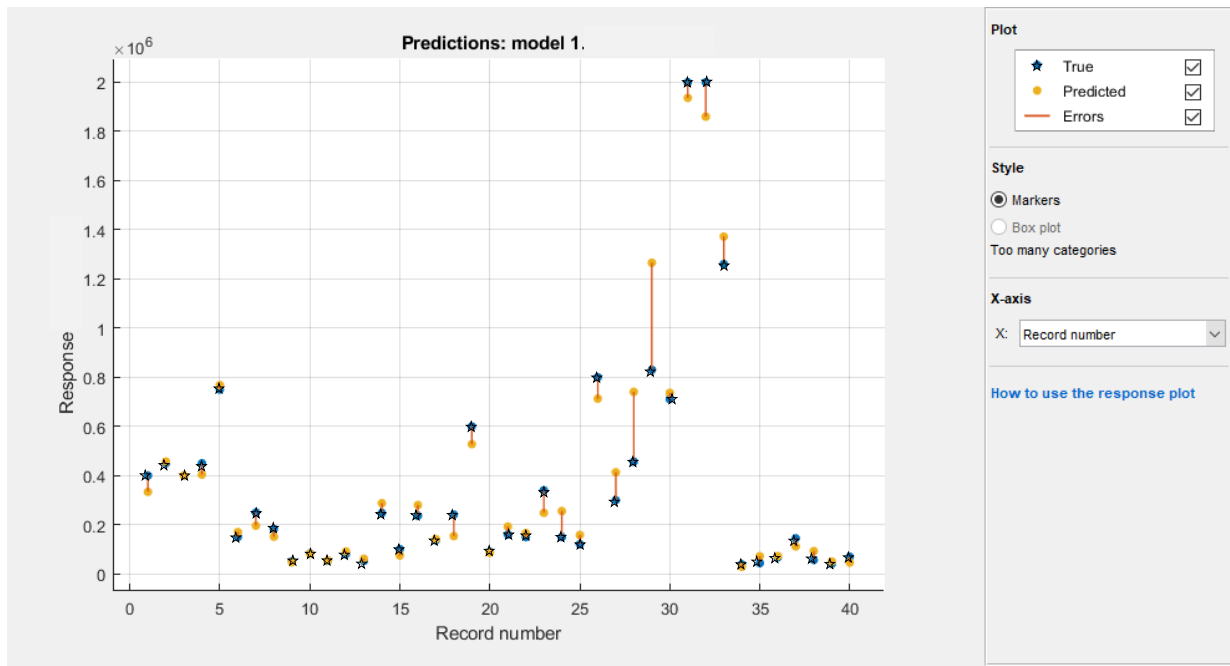


Figure 5-4 Response plot for robust linear regression- model 1

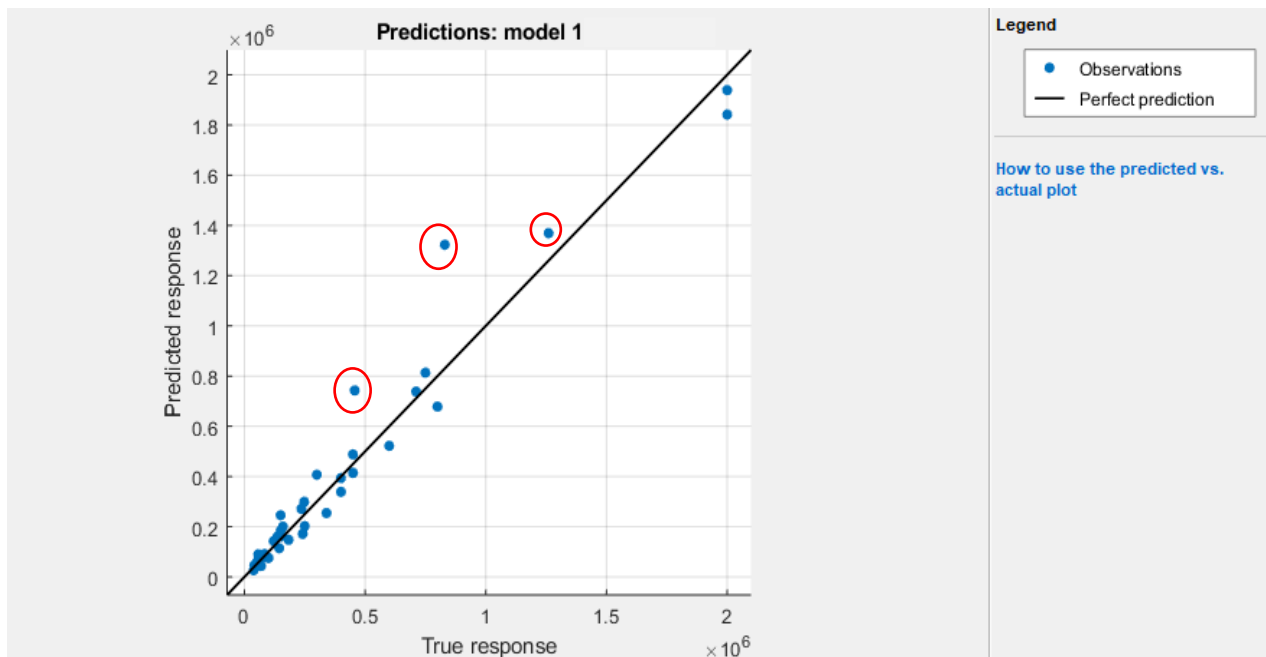


Figure 5-5 Predicted vs true response plot for robust linear

Considering predicted versus true response plot given in Figure 5-5, most of the observations are scattered near the perfect prediction line with a symmetric distribution around it. An exception is found in three points where the error between the observations and perfect regression model is noticeable.

Looking to the residuals plot given in Figure 5-6, the residuals scattered roughly symmetrically around zero and the majority are within $(1, -1) \times 10^5$ limit. An exception is found in four points where the residuals' values between the observations and true response are more than $(-1 \times 10^5, 1 \times 10^5)$.

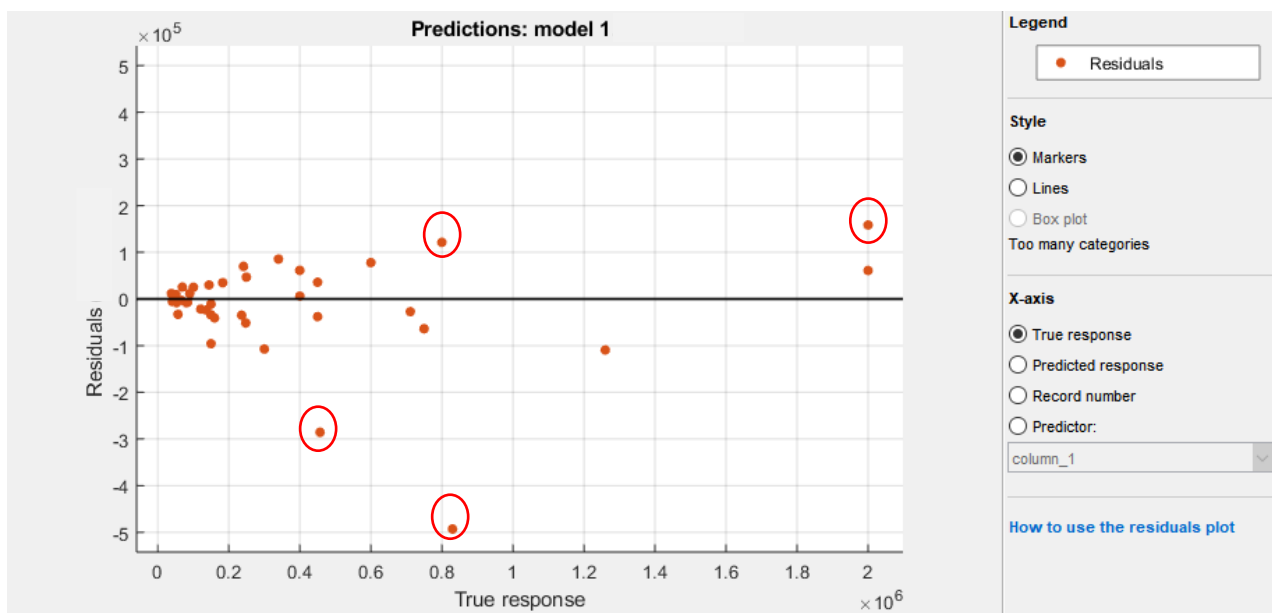


Figure 5-6 Residuals plot for robust linear regression- model 1

Figure 5-7 is the history list for model two that contains the trained models overall score. Best solution detected for model two where estimated cost is function of (project type, year of award group, award amount, cost deviation) by using stepwise linear regression (model 1.4) with RMSE=16.742. Other model statistics are found as following: R-Square= 1.00, MSE= 280.31, and MAE= 11.969.

Data Browser		
▼ History		
1.1	☆ Linear Regression	RMSE: 88314
	Last change: Linear	4/4 features
1.2	☆ Linear Regression	RMSE: 19.385
	Last change: Interactions Linear	4/4 features
1.3	☆ Linear Regression	RMSE: 88495
	Last change: Robust Linear	4/4 features
1.4	☆ Stepwise Linear Regre...	RMSE: 16.742
	Last change: Stepwise Linear	4/4 features
1.5	☆ Tree	RMSE: 2.7419e+05
	Last change: Fine Tree	4/4 features
1.6	☆ Tree	RMSE: 3.4803e+05
	Last change: Medium Tree	4/4 features
1.7	☆ Tree	RMSE: 4.6866e+05
▼ Current Model		
Results		
RMSE	16.742	
R-Squared	1.00	
MSE	280.31	
MAE	11.696	
Prediction speed	~930 obs/sec	
Training time	9.3199 sec	

Figure 5-7 History list for stepwise regression- model 2

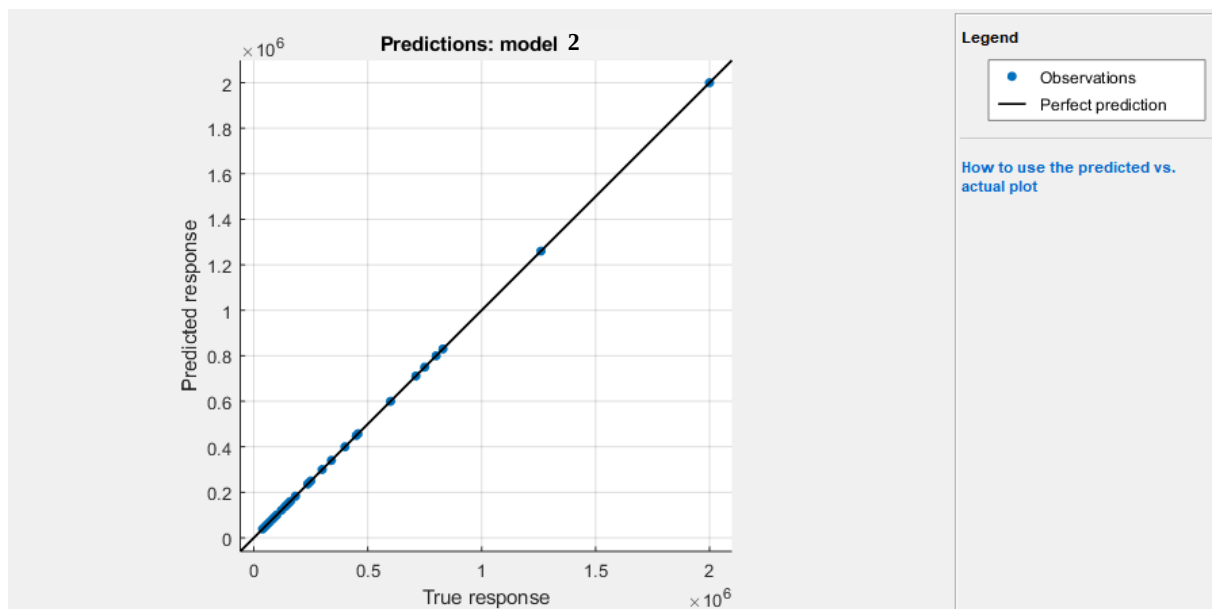


Figure 5-8 Response plot for stepwise regression- model 2

The model exhibits relatively small prediction error between the true data and predicted data as shown in Figure 5-8. Considering predicted response versus true response plot given in Figure 5-9, all the observations are perfectly aligned with perfect prediction regression model line.

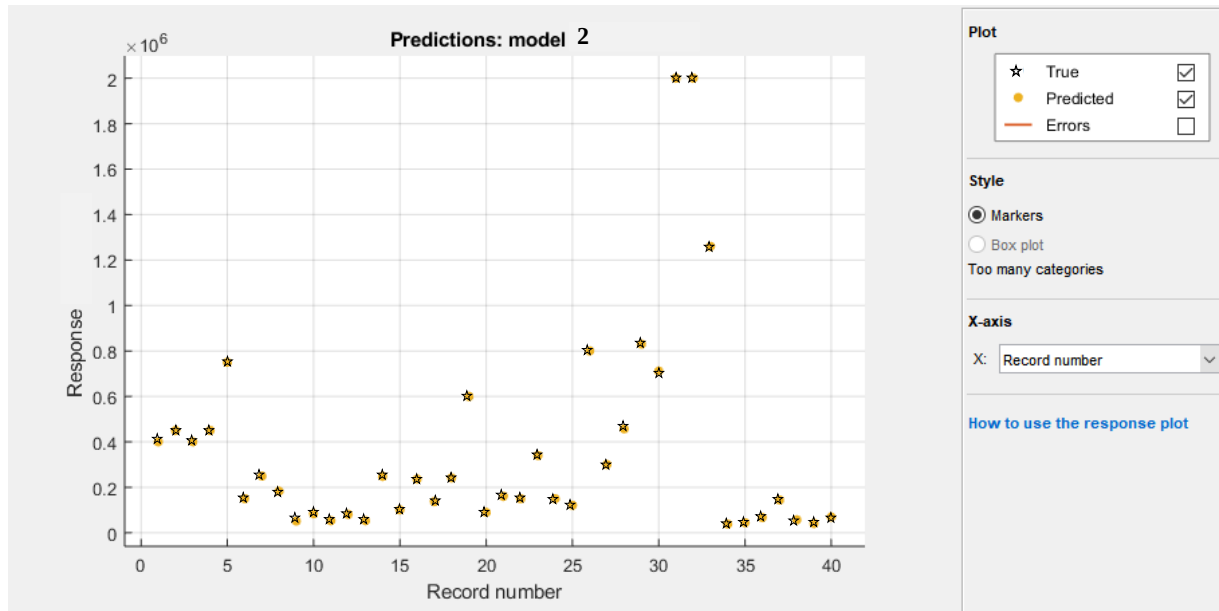


Figure 5-9 Predicted vs true response plot for stepwise regression- model 2

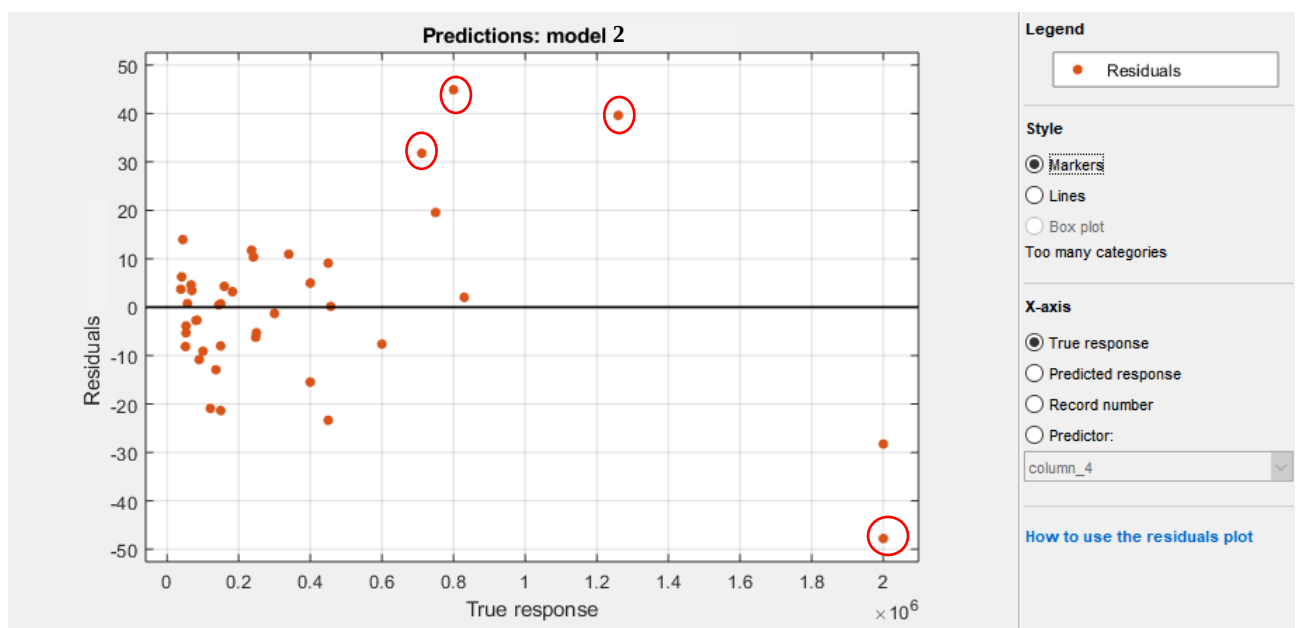


Figure 5-10 Residuals plot for stepwise regression- model 2

Looking to the residuals plot given in Figure 5-10, the residuals scattered roughly symmetrically around 0 and the majority are within (-20, 20) residual limit. An exception is found in four points where the residuals value between the observations and true response are more than (-30, 30) residual limit.

5.3.2 Regression models validation

In order to validate the models, new data was used in the developed regression model to predict the output. The results of the prediction for model one and model two are shown in Table 5-16 and Table 5-17. With respect to model one, robust linear regression estimate model was able to predict the project cost with an error range of (-21.4% to -1.74%).

Table 5-16 Robust linear regression- model 1 validation

ID	Type	Year period	Awarded amount (BD)	Estimated cost (BD)	LR estimation (BD)	Error (%)
6	1	3.00	456,021.195	400,015.000	448,068.884	-1.74
21	3	4.00	208,357.900	222,248.250	187,409.000	-10.05
35	2	4.00	286,592.756	250,000.000	264,731.679	-7.63
45	4	3.00	30,116.000	31,790.000	23,670.536	-21.40

With respect to the second model, stepwise regression estimate model was able to predict the project cost with an error range of (-12.76% to 11.45%).

Table 5-17 Stepwise regression- model 2 validation

ID	Type	Year period	Awarded amount (BD)	Estimated cost (BD)	Stepwise estimation (BD)	Error (%)
6	1	3.00	456,021.195	400,015.000	400,043.322	-12.28
21	3	4.00	208,357.900	222,248.250	232,223.534	11.45
34	2	4.00	286,592.756	250,000.000	250,014.581	-12.76
45	4	3.00	30,116.000	31,790.000	31,790.623	5.56

5.4 Analysis of Factors Affecting Cost Estimation

Accuracy

5.4.1 Demographic and work-related characteristics

The sample consisted of 50 participants of whom the majority were Bachelor degree holders (N=37, 74%) as shown in Table 5-18. Fairly half of the sample were contractors while the remaining participants were either clients or consultants with exact proportions (N=12, 24% for each). Nearly two thirds (N=29, 58%) of the sample hold managerial positions in their organizations with the vast majority of work experience in the field of construction for more than 5 years (N=41, 82%).

Table 5-18 Participants demography

Variable	Categories	N	%
Education	Bachelor degree	37	74%
	Master degree	9	18%
	Professional degree	4	8%
Role in the organization	Client	12	24%
	Consultant	12	24%
	Contractor	26	52%
Position in the organization	Top management	10	20%
	Middle management	2	4%
	Project management	17	34%
	Site engineer	9	18%
	Quantity surveying	6	12%
	Architect	5	10%
	Others	1	2%
Years of experience in construction projects	1 to 5 years	9	18%
	5 to 10 years	14	28%
	10 to 15 years	6	12%
	Over 15 years	21	42%

5.4.2 Cost estimation practices

As summarized in Table 5-19, half of the consultants reported that they are Class A and the other half were Class B. Similar proportion (N=14, 53.8%) of the contractors described themselves as Class A and they were the majority within this group. However, the remaining contractors were ramified to the rest three classes with noticeably variable proportions.

Table 5-19 Summary of the cost estimation practice in participants organization

Variable	Response	Client	Consultant	Contractor
		N (%)	N (%)	N (%)
What is your cost limit for each project in terms of Bahrain (company classification)? <i>Applicable for Consultants and Contactors</i>	Class A	-	6 (50%)	14 (53.8%)
	Class B	-	6 (50%)	8 (30.8%)
	Class C	-	0	3 (11.5%)
	Class D	-	0	1 (3.8%)
Is there an estimating unit in your organization?	Yes	6 (50%)	4 (33.3%)	26 (100%)
	No	6 (50%)	8 (66.7%)	0
How many employees are working in estimating unit in your organization/ firm?	None	6 (50%)	6 (50%)	0
	Less than 3	4 (33.3%)	4 (33.3%)	10 (38.5%)
	3 to 5	2 (16.7%)	2 (16.7%)	12 (46.2%)
	More than 5	0	0	4 (15.4%)
How much you rely on historical data to assess the preparation of your cost estimate?	100%-90%	5 (41.7%)	1 (8.3%)	9 (34.6%)
	89%-70%	2 (16.7%)	10 (83.3%)	9 (34.6%)
	69%-50%	4 (33.3%)	1 (8.3%)	3 (11.5%)
	less than 50%	1 (8.3%)	0	5 (19.2%)
Do you consider labor and equipment productivity during cost estimation?	Yes	6 (50.0%)	11 (91.7%)	23 (88.5%)
	No	6 (50.0%)	1 (8.3%)	3 (11.5%)
Who sets up the firm tendering price?	management level	6 (50.0%)	8 (66.7%)	18 (69.2%)
	estimating level	6 (50.0%)	4 (33.3%)	8 (30.8%)

All contractors reported that they had units dedicated for projects' cost estimation while only 50% of the clients and 33.3% of the consultants reported that they have cost estimation units. Two thirds of the contractors (N=16, 61.6%) had at least three persons working in the cost estimation units while only (16.7% each) of the clients and consultants had similar number of employees.

Table 5-19 shows higher percentage of consultants (N=11, 91.6%) reported that they rely on historical data at least 70% of the time when compared to the rest two groups (N=7, 58.4% and N=18, 69.2%) for the clients and the contractors, respectively. The majority of both consultants and contractors (91.7% and 88.5%) reported that they do consider the labor's productivity and equipment in cost estimation process. With respect to project cost determination, the majority of the participants (N=32, 64%) have reported that the project cost is defined at management level with the following respondents detailed percentage: client (N=6, 50%), consultant (N=8, 66.7%) and contractor (N=18, 69.2%). Whereas the remaining respondents (N=18, 36%) reported that the project cost is developed at estimation level in their organizations.

5.4.3 Satisfaction with pre-tender cost estimation

Only one third of both clients and consultants (N=4, 33.3% for each) stated that they are satisfied with the pre-tender cost estimation accuracy. On the contrary, the majority of the contractors were satisfied (N=18, 69.3%). This difference is statistically significant according to the Chi square likelihood ratio test. The same conclusion supported with the Bayesian multinomial test as it resulted in a Bayesian factor of 20.3, which indicates strong evidence of a significant difference. Table 5-20 summarizes the results.

Table 5-20 Participants degree of satisfaction on pre-tender cost estimation accuracy

Variable	Response	Client	Consultant	Contractor
		N (%)	N (%)	N (%)
Are you satisfied with the current level of pre-tender cost estimate accuracy?	Very unsatisfied	0	1 (8.3%)	1 (3.8%)
	Not satisfied	2 (16.7%)	3 (25.0%)	0
	Neutral	6 (50.0%)	4 (33.3%)	7 (26.9%)
	Satisfied	3 (25.0%)	3 (25.0%)	12 (46.2%)
	Very satisfied	1 (8.3%)	1 (8.3%)	6 (23.1%)
<i>p-value (Bayesian factor)</i>	0.020* (20.3)			

**Statistically significant differences with Chi square Likelihood ratio test and Monte Carlo simulation of 1000 samples ($X^2=13.5$, $df=8$, $p=0.020$) at Alpha 0.05.*

5.4.4 Stakeholders' perception on factors affecting cost estimation

In terms of factors RII scores, a total of 33 factors distributed over 8 scales ranked as critical factors affecting project cost estimation. The breakdown of the critical factors according to their related scales is presented in Table 5-21. The range of the RII scores is between (75.20%-87.69%), which shows that nearly the participants are on an agreement regarding the critical effect of listed factors on both cost estimation practice and construction cost. Appendix 8-Section A contains the detailed RII calculation.

Table 5-21 Critical factors that affect MYS construction cost estimation

Num	Category	Notation	Factor	RII
1	Scope Quality (Limited to Client and Consultants)	SQ	Clear scope definition for the client	85.00%
2			Level of team integration	78.33%
3			Buildability of design	79.17%
4			Level of involvement of client in the estimation procedure	82.50%
5			Client's priority level of the project	75.83%
6	Estimator Performance	EP	Experience of pricing construction	86.00%
7			Impact of team integration and alignment	75.20%
8			Estimating method used	81.60%
9			Time allowed for preparing cost estimate	80.80%
10	Information Quality	IQ	Completeness of cost information	82.80%
11			Accuracy and reliability of cost information	84.40%
12			Clear and detail drawings and specification	81.20%
13	Project Characteristics	PC	Type/function	82.80%
14			Type of structures (steel, concrete, masonry)	85.60%
15			Size/gross floor area	80.80%
16			Quality of finishing	86.00%
17			Priority of project	82.40%
18			Duration of the project period	84.00%
19			Clients financial situation and budget	82.40%
20	External Factors	EF	Cost of materials	84.40%
21			Cost of labors	83.20%
22			Cost of machinery	81.60%
23			Overhead costs	79.60%
24	Uncertainty Level	UL	Inflationary pressure	76.40%
25	Contractual Procedure	CP	Contract conditions	78.40%
26			Number of bidders on competitive projects	75.60%
27	Organization Management (Limited to Contractors)	OM	Management team (suitability, experience, performance)	84.62%
28			Financial capability	81.54%
29			Past relationship with client	87.69%
30			Payment record of client	84.62%
31			Relationship with subcontractors and suppliers	81.54%
32			Experience on similar projects	86.92%
33			Past loss/profit in similar projects	77.69%

5.4.5 Scales reliabilities

Most of the scales showed optimal reliability as measured with Cronbach's α for internal consistency as shown in Table 5-22 (minimal satisfactory reliability is 0.7 as a cutoff point).

Table 5-22 Research category's reliability values

	Category	Number of factors	Sample size	Cronbach's α
1	Scope quality (SQ)	5	24	0.746
2	Estimator's performance (EP)	4	50	0.747
3	Information quality (IQ)	3	50	0.816
4	Project's characteristics (PC)	7	50	0.784
5	External factors (EF)	4	50	0.882
6	Contractual procedure (CP)	2	50	0.551
7	Organization Management (OM)	7	26	0.817

The highest reliability was recorded for the external factors (EF) scale (Cronbach's $\alpha = 0.882$), whereas the lowest was for contractual procedure (CP) with Cronbach's $\alpha = 0.551$. Another low reliability was found in quality of estimation procedure (QEP) category (Cronbach's α lower than 0.5). QEP factors were merged with EP and IQ categories and show optimal reliability as a composite scale with Cronbach's $\alpha = 0.747$ and 0.816. Uncertainty level category has been discharged from further analysis as the group contains one factor only.

5.4.6 Participants perception of the factors affecting cost estimation

1. Scope Quality (limited to clients and consultants)

The participants reported moderate level of agreement ($M(SD)=4\pm0.6$, $M(SD)=4\pm0.6$ for the clients and the consultants, respectively) to the asserted factors that affect cost estimation within the realm of scope quality as 68.3% of the clients' responses and 73.3% of the consultant's responses fell into the agreement side (agree plus totally agree).

Table 5-23 Participants perception on scope quality effect on cost estimation

Factor	Response	Clients N (%)	Consultants N (%)
Clear scope definition for the client	Totally agree	7 (58.3%)	5 (41.7%)
	Agree	3 (25%)	4 (33.3%)
	Total agreement	10 (83.3%)	9 (75%)
	Mean agreement	4.4	4.1
Level of team integration	Totally agree	4 (33.3%)	3 (25%)
	Agree	6 (50%)	3 (25%)
	Total agreement	10 (83.3%)	6 (50%)
	Mean agreement	4.1	3.6
Buildability of design	Totally agree	3 (25%)	3 (25%)
	Agree	3 (25%)	8 (66.6%)
	Total agreement	6 (50%)	11 (91.6%)
	Mean agreement	3.75	4.2
Level of involvement of client in the estimation procedure	Totally agree	4 (33.3%)	6 (50%)
	Agree	5 (41.7%)	4 (33.3%)
	Total agreement	9 (75%)	10 (83.3%)
	Mean agreement	4	4.2
Clients priority level of the project	Totally agree	3 (25%)	3 (25%)
	Agree	3 (25%)	6 (50%)
	Total agreement	6 (50%)	9 (83.3%)
	Mean agreement	3.6	3.9
Scope quality pooled agreement (%)		41 (68.3%)	45 (73.3%)
Scope quality pooled agreement (Mean)		4	4
<i>p-value (Bayesian factor)</i>		<i>0.700* (0.380)</i>	

**Statistically insignificant difference according to Mann-Whitney test and Monte Carlo simulation with 1000 sample (U=65, Z=-0.04, p=0.700). The same conclusion is supported by Bayesian Mann-Whitney test as it resulted in a Bayesian factor of 0.380 which indicates no evidence of significant difference between the two mean ranks.*

On average, the clients had an identical level of perception regarding the effect of scope of quality on the cost estimation to that of the consultants as they had an identical mean score on this scale as shown in Table 5-23. This finding is confirmed with Mann Whitney test as it revealed insignificant differences in their averages (U=65, Z=-0.04, p=0.700). The same conclusion is supported by Bayesian Mann-Whitney test as it resulted in a Bayesian factor of 0.380 which indicates no evidence of significant difference between the two mean ranks.

2. Estimator Performance

The participants reported moderate levels of agreement ($M(SD)=3.9\pm0.5$, $M(SD)=3.9\pm0.6$, $M(SD)=4.1\pm0.5$ for the clients, the consultants, and the contractors, respectively) to the asserted factors that affect cost estimation within the realm of estimation procedure as 72.9% of the clients' responses, 68.7% of the consultant's responses, and 82.7% of the contractors responses fell into the agreement side (agree plus totally agree) as shown in Table 5-24.

Table 5-24 Participants perception on estimator performance effect on cost estimation

Factor	Response	Clients	Consultants	Contractors
		N (%)	N (%)	N (%)
Experience of pricing construction	Totally agree	3 (25%)	5 (41.7%)	14 (53.8%)
	Agree	6 (50%)	5 (41.7%)	10 (38.5%)
	Total agreement	9 (75%)	10 (83.4%)	24 (92.3%)
	Mean agreement	4	4.2	4.4
Impact of team integration and alignment	Totally agree	2 (16.7%)	2 (16.7%)	5 (19.2%)
	Agree	6 (50%)	5 (41.7%)	12 (46.2%)
	Total agreement	8 (66.7%)	7 (58.3%)	17 (65.4%)
	Mean agreement	3.7	3.7	3.7
Estimating method used	Totally agree	2 (16.7%)	4 (33.3%)	7 (26.9%)
	Agree	7 (58.3%)	4 (33.3%)	17 (65.4%)
	Total agreement	9 (75%)	8 (66.6%)	24 (92.3%)
	Mean agreement	3.9	4	4.2
Time allowed for preparing cost estimate is not sufficient	Totally agree	2 (16.7%)	4 (33.3%)	9 (34.6%)
	Agree	7 (58.3%)	4 (33.3%)	13 (50%)
	Total agreement	9 (75%)	8 (66.6%)	22 (84.6%)
	Mean agreement	3.9	3.9	4.1
Estimator performance pooled agreement (%)		35 (72.9%)	33 (68.7%)	86 (82.7%)
Estimator performance pooled agreement (Mean)		3.87	3.95	4.10
<i>p-value (Bayesian factor)</i>		<i>0.504* (0.309)</i>		

**Statistically insignificant difference according to Kurskal-Wallis test and Monte Carlo simulation with 1000 sample ($H=1.37$, $df=2$, $p=0.504$). The same conclusion is supported by Bayesian ANOVA test as it resulted in a Bayesian factor of 0.309 which indicates no evidence of significant difference between the three means*

However, despite the observed differences among their levels of agreement, these differences were not statistically significant as confirmed by the Kruskal-Wallis test ($H=1.37$, $df=2$, $p=0.504$). The same conclusion is supported

by Bayesian ANOVA test as it resulted in a Bayesian factor of 0.309 which indicates no evidence of significant difference between the three means.

3. Information Quality

The participants reported moderate levels of agreement ($M(SD)=4\pm0.6$, $M(SD)=4.3\pm0.9$, $M(SD)=4.1\pm0.6$ for the clients, the consultants, and the contractors, respectively) to the asserted factors that affect cost estimation within the realm of information quality as 75% of the clients' responses, 77.7% of the consultant's responses, and 76.9% of the contractors responses fell into the agreement side (agree plus totally agree) as shown in Table 5-25.

Table 5-25 Participants perception on information quality effect on cost estimation

Factor	Response	Clients	Consultants	Contractors
		N (%)	N (%)	N (%)
Completeness of cost information	Totally agree	2 (16.7%)	7 (58.3%)	10 (38.5%)
	Agree	7 (58.3%)	2 (16.7%)	12 (46.2%)
	Total agreement	9 (75%)	9 (75%)	22 (84.7%)
	Mean agreement	3.9	4.2	4.2
Accuracy and reliability of cost information	Totally agree	3 (25%)	8 (66.7%)	12 (46.2%)
	Agree	5 (41.7%)	1 (8.3%)	10 (38.5%)
	Total agreement	8 (66.7%)	9 (75%)	22 (84.7%)
	Mean agreement	3.9	4.3	4.3
Clear and detail drawings and specification	Totally agree	5 (41.7%)	7 (58.3%)	5 (19.2%)
	Agree	5 (41.7%)	3 (25%)	11 (42.3%)
	Total agreement	10 (93.4%)	10 (93.4%)	16 (61.5%)
	Mean agreement	4.2	4.4	3.8
Information quality pooled agreement (%)		27 (75%)	28 (77.7%)	60 (76.9%)
Information quality agreement (Mean)		4.0	4.3	4.1
p-value (Bayesian factor)		0.238* (0.249)		

**Statistically insignificant difference according to Kurskal-Wallis test and Monte Carlo simulation with 1000 sample ($H=3.87$, $df=2$, $p=0.238$). The same conclusion is supported by Bayesian ANOVA test as it resulted in a Bayesian factor of 0.249 which indicates no evidence of significant difference between the three means*

However, despite the observed differences among their levels of agreement, these differences were not statistically significant as confirmed by the Kruskal-Wallis test ($H=3.87$, $df=2$, $p=0.238$) and Bayesian ANOVA test as it

resulted in a Bayesian factor of 0.249 which indicates no evidence of significant difference between the three means.

4. Project Characteristics

The participants reported moderate levels of agreement ($M(SD)=3.±0.6$, $M(SD)=4.3±0.5$, $M(SD)=4.3±0.6$ for the clients, the consultants, and the contractors, respectively) to the asserted factors that affect cost estimation within the realm of project characteristics as 66.7% of the clients' responses, 86.1% of the consultant's responses, and 82% of the contractors responses fell into the agreement side (agree plus totally agree) as shown in Table 5-26.

Table 5-26 Participants perception on project characteristics effect on cost estimation

Factor	Response	Clients	Consultants	Contractors
		N (%)	N (%)	N (%)
Type/function (playground, admin building, model youth center, sport hall)	Totally agree	2 (16.7%)	7 (58.3%)	13 (50%)
	Agree	4 (33.3%)	4 (33.3%)	8 (30.8%)
	Total agreement	6 (50%)	11 (91.6%)	21 (80.7%)
	Mean agreement	3.5	4.5	4.2
Type of structures (steel, concrete, masonry)	Totally agree	3 (25%)	4 (33.3%)	16 (61.5%)
	Agree	6 (50%)	6 (50%)	6 (23.1%)
	Total agreement	9 (75%)	10 (83.3%)	22 (84.6%)
	Mean agreement	4	4.1	4.4
Size/gross floor area	Totally agree	2 (16.7%)	4 (33.3%)	12 (46.2%)
	Agree	7 (58.3%)	3 (25%)	8 (30.8%)
	Total agreement	9 (75%)	7 (58.3%)	16 (61.5%)
	Mean agreement	3.9	3.7	3.8
Quality of finishing	Totally agree	5 (41.7%)	7 (58.3%)	15 (57.7%)
	Agree	4 (33.3%)	4 (33.3%)	6 (23.1%)
	Total agreement	9 (75%)	11 (91.6%)	21 (80.8%)
	Mean agreement	4.1	4.5	4.3
Priority of project	Totally agree	2 (16.7%)	5 (41.7%)	10 (38.5%)
	Agree	7 (58.3%)	5 (41.7%)	12 (46.2%)
	Total agreement	9 (75%)	10 (83.4%)	22 (84.7%)
	Mean agreement	3.9	4.2	4.1

However, despite the observed differences among their levels of agreement, these differences were not statistically significant as confirmed by the Kruskal-Wallis test ($H=3.3$, $df=2$, $p=0.179$). The same conclusion is supported by Bayesian ANOVA test as it resulted in a Bayesian factor of 0.130 which indicates no evidence of significant difference between the three means.

Table 5-26 Participants perception on project characteristics effect on cost estimation (continued)

Factor	Response	Clients	Consultants	Contactors
		N (%)	N (%)	N (%)
Duration of the project period	Totally agree	4 (33.3%)	5 (41.7%)	12 (46.2%)
	Agree	5 (41.7%)	5 (41.7%)	9 (34.6%)
	Total agreement	9 (75%)	10 (83.4%)	21 (80.8%)
	Mean agreement	4.1	4.2	4.2
Clients financial situation and budget	Totally agree	3 (25%)	5 (41.7%)	13 (50%)
	Agree	3 (25%)	5 (41.7%)	9 (34.6%)
	Total agreement	6 (50%)	10 (83.3%)	21 (80.8%)
	Mean agreement	3.7	4.1	4.2
Project characteristics pooled agreement (%)		48 (66.7%)	62 (86.1%)	128 (82%)
Project characteristics pooled agreement (Mean)		3.9	4.3	4.3
p-value (Bayesian factor)		0.179* (0.130)		

**Statistically insignificant difference according to Kruskal-Wallis test and Monte Carlo simulation with 1000 sample ($H=3.3$, $df=2$, $p=0.179$). The same conclusion is supported by Bayesian ANOVA test as it resulted in a Bayesian factor of 0.130 which indicates no evidence of significant difference between the two mean ranks.*

5. External Factors

The participants reported moderate levels of agreement ($M(SD)=4\pm0.6$, $M(SD)=3.9\pm0.5$, $M(SD)=4.2\pm0.7$ for the clients, the consultants, and the contractors, respectively) to the asserted factors that affect cost estimation within the realm of external factors as 79.1% of the clients' responses, 81.1% of the consultant's responses, and 76.9% of the contractors responses fell into the agreement side (agree plus totally agree) as shown in Table 5-27.

However, despite the observed differences among their levels of agreement, these differences were not statistically significant as confirmed by the

Kruskal-Wallis test ($H=2.2$, $df=2$, $p=0.328$). The same conclusion is supported by Bayesian ANOVA test as it resulted in a Bayesian factor of 0.303 which indicates no evidence of significant difference between the three means.

Table 5-27 Participants perception on external factors effect on cost estimation

Factor	Response	Clients	Consultants	Contactors
		N (%)	N (%)	N (%)
Cost of materials	Agree	5 (41.7%)	8 (66.7%)	5 (19.2%)
	Totally agree	4 (33.3%)	3 (25%)	15 (57.7%)
	Total agreement	9 (75%)	11 (81.7%)	20 (76.9%)
	Mean agreement	4.1	4.1	4.3
Cost of labor	Agree	8 (66.7%)	10 (83.3%)	6 (23.1%)
	Totally agree	2 (16.7%)	1 (8.3%)	14 (53.8%)
	Total agreement	10 (83.4%)	11 (81.7%)	20 (76.6%)
	Mean agreement	4	4	4.3
Cost of machinery	Agree	7 (58.3%)	7 (58.3%)	9 (34.6%)
	Totally agree	2 (16.7%)	3 (25%)	11 (42.3%)
	Total agreement	9 (75%)	10 (83.4%)	20 (76.6%)
	Mean agreement	3.9	4	4.2
Overhead costs	Agree	8 (66.7%)	5 (41.7%)	9 (34.6%)
	Totally agree	2 (16.7%)	2 (16.7%)	11 (42.3%)
	Total agreement	10 (83.4%)	7 (58.4%)	20 (76.6%)
	Mean agreement	4	3.6	4.1
External factors pooled agreement (%)		38 (79.1%)	39 (81.1%)	80 (76.9%)
External factors pooled agreement (Mean)		4	3.9	4.2
p-value (Bayesian factor)		0.328* (0.303)		

**Statistically insignificant difference according to Kurskal-Wallis test and Monte Carlo simulation with 1000 sample ($H=2.2$, $df=2$, $p=0.328$). The same conclusion is supported by Bayesian ANOVA test as it resulted in a Bayesian factor of 0.303 which indicates no evidence of significant difference between the three means.*

6. Contractual Procedure

The participants reported moderate levels of agreement ($M(SD)=3.6\pm0.6$, $M(SD)=3.7\pm0.9$, $M(SD)=4\pm0.7$ for the clients, the consultants, and the contractors, respectively) to the asserted factors that affect cost estimation within the realm of external factors as 66.6% of the clients' responses, 58.3% of the consultant's responses, and 76.9% of the contractors responses fell into the agreement side (Agree plus totally agree) as shown in Table 5-28.

Table 5-28 Participants perception on contractual procedure effect on cost estimation

Factor	Response	Clients	Consultants	Contactors
		N (%)	N (%)	N (%)
Contract conditions	Agree	8 (66.7%)	2 (16.7%)	14 (53.8%)
	Totally agree	1 (8.3%)	4 (33.3%)	9 (34.6%)
	Total agreement	9 (75%)	6 (50%)	23 (88.4%)
	Mean agreement	3.7	3.6	4.1
Number of bidders on competitive projects	Agree	6 (50%)	5 (41.7%)	9 (34.6%)
	Totally agree	1 (8.3%)	3 (25%)	8 (30.8%)
	Total agreement	7 (58.3%)	8 (66.7%)	17 (65.4%)
	Mean agreement	3.5	3.7	3.8
Contractual Procedure pooled agreement (%)		16 (66.6%)	14 (58.3%)	40 (76.9%)
Contractual Procedure agreement (Mean)		3.6	3.7	4
p-value (Bayesian factor)		0.312* (0.333)		

**Statistically insignificant difference according to Kurskal-Wallis test and Monte Carlo simulation with 1000 sample ($H=2.5$, $df=2$, $p=0.312$). The same conclusion is supported by Bayesian ANOVA test as it resulted in a Bayesian factor of 0.333 which indicate no evidence of significant difference between the three means.*

However, despite the observed differences among their levels of agreement, these differences were not statistically significant as confirmed by the Kruskal-Wallis test ($H=2.5$, $df=2$, $p=0.312$). The same conclusion is supported by Bayesian ANOVA test as it resulted in a Bayesian factor of 0.333 which indicates no evidence of significant difference between the three means.

7. Organization Management (limited to contactors)

The participants reported moderate levels of agreement ($M(SD)=4.1\pm0.6$) to the asserted factors that effects cost estimation within the realm of organization management. Table 5-29 shows that 83.3% of the responses fell into the agreement side.

Table 5-29 Participants perception on organization management effect on cost estimation

Factor	Response	Contactors N (%)
Management team (suitability, experience, performance)	Agree	11 (42.3%)
	Totally Agree	12 (46.2%)
	Total agreement	23 (88.5%)
	Mean agreement	4.2
Financial capability	Agree	11 (42.3%)
	Totally Agree	10 (38.5%)
	Total agreement	21 (80.8%)
	Mean agreement	4.1
Past relationship with client	Agree	6 (23.1%)
	Totally Agree	16 (61.5%)
	Total agreement	22 (84.6%)
	Mean agreement	4.3
Payment record of client	Agree	6 (23.1%)
	Totally Agree	14 (53.8%)
	Total agreement	20 (76.9%)
	Mean agreement	4.2
Relationship with subcontractors and suppliers	Agree	12 (46.2%)
	Totally Agree	9 (34.6%)
	Total agreement	21 (80.8%)
	Mean agreement	4.1
Past loss/profit in similar projects	Agree	12 (46.2%)
	Totally Agree	6 (23.1%)
	Total agreement	18 (69.23%)
	Mean agreement	3.8
Experience on similar projects	Agree	8 (30.8%)
	Totally Agree	15 (57.7%)
	Total agreement	23 (88.5%)
	Mean agreement	4.3
Organization Management agreement (%)		148 (81.3%)
Organization Management agreement (Mean)		4.1

5.4.7 Level of management relationship with the perceived factors affecting cost estimation

There is no relationship between the participants' managerial level and their scores on these scales as confirmed with Bayesian Man-Whitney test as shown in Table 5-30.

Despite the causal differences, however they were not statistically significant. A similar conclusion was made with Bayesian Mann-Whitney test as it resulted in a Bayesian factors ranging from 0.292 to 0.954, which indicates no evidence of significant differences among the means.

Table 5-30 Mean categories scores based on participants managerial level

Categories	Senior Managers	Others	p-value (Bayesian factor)
	Mean($\pm$ SD)	Mean($\pm$ SD)	
Scope quality	4.0 ($\pm$ 0.6)	4.1 ($\pm$ 0.6)	0.884 (0.389)
Estimator Performance	3.9 ($\pm$ 0.5)	4.2 ($\pm$ 0.5)	0.101 (0.703)
Information quality	4.0 ($\pm$ 0.7)	4.3 ($\pm$ 0.6)	0.111 (0.735)
Project characteristics	4.2 ($\pm$ 0.6)	4.2 ($\pm$ 0.6)	0.830 (0.292)
External factors	4.4 ($\pm$ 0.7)	4.0 ($\pm$ 0.6)	0.231 (0.572)
Contractual procedure	3.6 ($\pm$ 0.8)	4.1 ($\pm$ 0.6)	0.087 (0.954)
Organization management	4.5 ($\pm$ 0.3)	4.1 ($\pm$ 0.8)	0.296 (0.520)

Senior managers: top and middle managers and project managers with 10 years or more of experience

Others: project managers with less than 10 years' experience, site engineers, quantity surveyors, architects, and others.

5.4.8 Cross validating critical factors with MOW Profession

A modified version of the survey has been introduced to MOW and their contractors in order to cross validate the factors affecting the construction cost estimation. Comparing MOW and MYS results with MYS factors ranking, a considerable agreement has been observed regarding the factors affecting construction cost estimation, as 31 factors (out of the 33 factors stated before) have been recognized as critical factors in both surveyed samples. However, “level of client involvement” was the lowest factor ranked among the list of factors given, and seven additional factors have been recognized as critical in MOW surveyed sample. Table 5-31 lists the additional factors recognized as critical as per MOW sample. The cross validation detailed analysis is presented in Appendix 8 section B.

Table 5-31 MOW Engineers perception on additional factors effect on cost estimation accuracy

No	Category	Factor	RII
1	SQ	Previous client / consultant experience of contract	87%
2	OM	Current workload	80%
3	PC	Inadequate production of raw material for country	87%
4	EF	Governments requirements	82%
5	UL	Price fluctuation	80%
6	EF	Weather	75%
7	CP	Time between project announcement, bid submission and award of contract	78%

5.5 Modeling of Cost Estimation Accuracy using PLS-SEM Technique

According to the results obtained in Section 5.4, the proposed structure of relationships among factors related to cost estimation accuracy will be modeled using the structural equation modeling approach based on the partial least square procedures. Three structural models will be tested: client and consultant model, main contractor model, common categories model. Wrap-PLS software has been used to obtain the models results.

5.5.1 Cost estimation accuracy: client-consultant model

The proposed structure of relationships among factors related to cost estimation accuracy from the client-consultant point of view are modeled as shown in Figure 5-11.

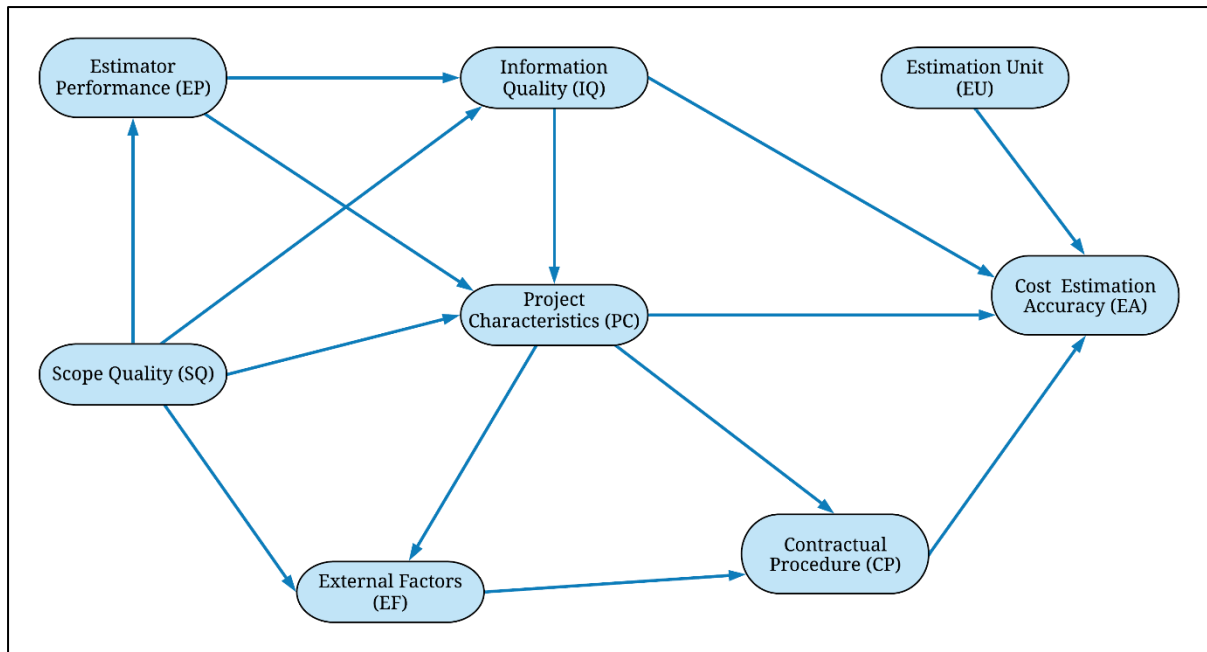


Figure 5-11 Proposed (client/consultant) PLS-SEM model

The proposed model stipulates the following hypotheses:

1. Possession of dedicated cost estimation unit affects the magnitude of cost estimation accuracy (H1).
2. The participants beliefs that the information quality effect on cost estimation has a relationship with the cost estimation accuracy of the projects they involved in previously (H2).
3. The participants beliefs that the project characteristics effect on cost estimation has a relationship with the cost estimation accuracy of the projects they involved in previously (H3).
4. The participants beliefs that the contractual procedure effect on cost estimation has a relationship with the cost estimation accuracy of the projects they involved in previously (H4).
5. The participants beliefs that the information quality effect on cost estimation accuracy has a relationship with their beliefs on the effect of project characteristics on cost estimation (H5).

6. The participants beliefs that the project characteristics effect on cost estimation accuracy has a relationship with their beliefs on the effect of contractual procedure on cost estimation (H6).
7. The participants beliefs that the project characteristics effect on cost estimation accuracy has a relationship with their beliefs on the effect of external factors on cost estimation (H7).
8. The participants beliefs that external factors on cost estimation accuracy has a relationship with their beliefs on the effect of contractual procedure effect on cost estimation (H8).
9. The participants beliefs that the estimator's performance effect on cost estimation accuracy has a relationship with their beliefs on the effect of project characteristics on cost estimation (H9).
- 10.The participants beliefs that the estimator's performance effect on cost estimation accuracy has a relationship with their beliefs on the effect of information quality's on cost estimation (10).
- 11.The participants beliefs that the quality scope effect on cost estimation accuracy has a relationship with their beliefs on the effect of estimator performance on cost estimation (H11).
- 12.The participants beliefs that the quality scope effect on cost estimation accuracy has a relationship with their beliefs on the effect of information quality on cost estimation (H12).
- 13.The participants beliefs that the quality scope effect on cost estimation accuracy has a relationship with their beliefs on the effect of external factors on cost estimation (H13).
- 14.The participants beliefs that the quality scope effect on cost estimation accuracy has a relationship with their beliefs on the effect of project characteristics on cost estimation (H14).

5.5.2 The structural equation modeling results of client-consultant model

The proposed model was statistically significant with most of the proposed relationships have path coefficients (b) with a significance value of ($p < 0.01$) as shown in Figure 5-12 and Table 5-32. Four categories converged on the cost estimation error and had a statistical significance relationships and moderately large effect size of $R^2 = 0.32$, $p < 0.01$. However, six relationships were not statistically significant ($CP \rightarrow EA$, $EF \rightarrow CP$, $PC \rightarrow EF$, $PC \rightarrow EA$, $QS \rightarrow EF$ and $EP \rightarrow PC$). The model has a statistically significant Tenenhaus Goodness of Fit (GoF) value of 0.555.

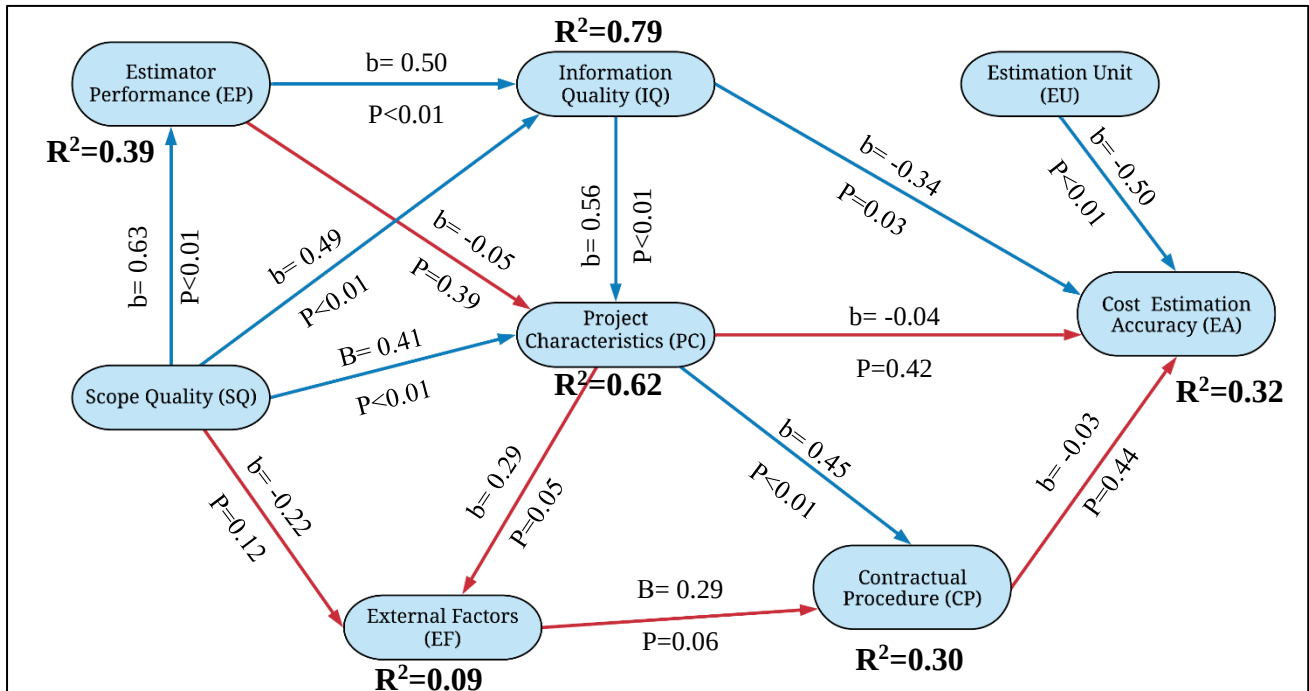


Figure 5-12 Results of client/consultant PLS-SEM model

Table 5-32 Results of client/consultant PLS-SEM model

Hypothesis	b(p-value)	Conclusion
H1: $EU \rightarrow EA$	-0.50(<0.01)	Hypothesis supported
H2: $IQ \rightarrow EA$	-0.34(0.03)	Hypothesis supported
H3: $PC \rightarrow EA$	-0.04(0.42)	Hypothesis unsupported
H4: $CP \rightarrow EA$	-0.03(0.44)	Hypothesis unsupported
H5: $IQ \rightarrow PC$	0.56(<0.01)	Hypothesis supported
H6: $PC \rightarrow CP$	0.45(<0.01)	Hypothesis supported
H7: $PC \rightarrow EF$	+0.29(0.05)	Hypothesis unsupported
H8: $EF \rightarrow CP$	+0.29(0.06)	Hypothesis unsupported
H9: $EP \rightarrow PC$	-0.05(0.39)	Hypothesis unsupported
H10: $EP \rightarrow IQ$	+0.50(<0.01)	Hypothesis supported
H11: $SQ \rightarrow EP$	0.63(<0.01)	Hypothesis supported
H12: $SQ \rightarrow IQ$	0.49(<0.01)	Hypothesis supported
H13: $SQ \rightarrow EF$	-0.22(0.12)	Hypothesis unsupported
H14: $SQ \rightarrow PC$	+0.41(<0.01)	Hypothesis supported

5.5.3 Cost estimation accuracy: contractors model

The proposed structure of relationships among factors related to cost estimation accuracy from the contractor point of view are modeled as shown in Figure 5-13.

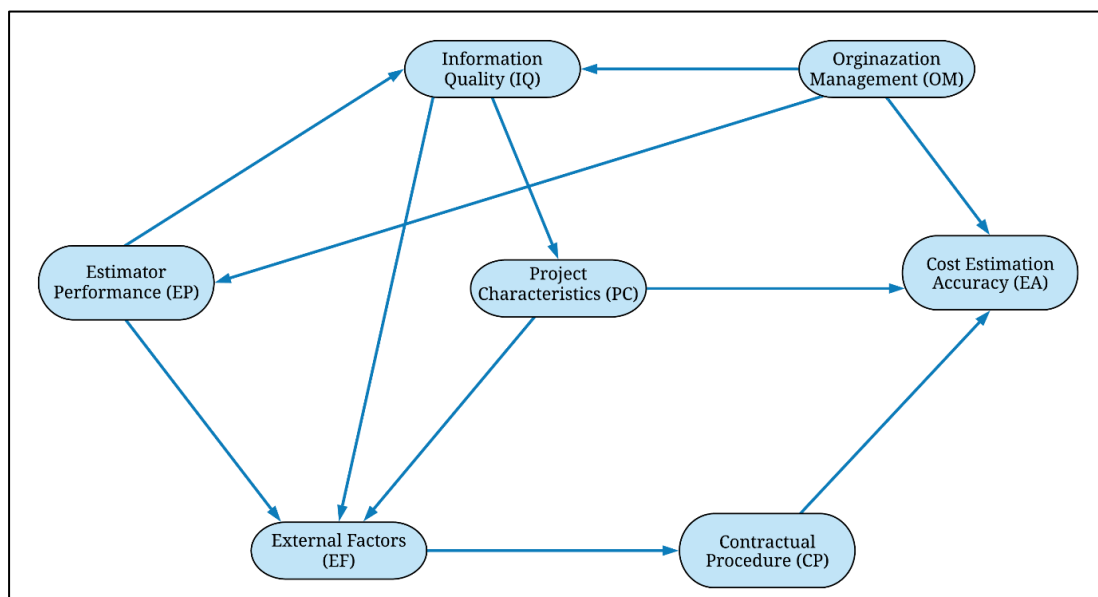


Figure 5-13 Proposed contractor PLS-SEM model

The proposed model stipulates the following hypotheses:

1. The participants beliefs that the organization management effect on cost estimation has a relationship with the cost estimation accuracy of the projects they involved in previously (H1).
2. The participants beliefs that the project characteristics effect on cost estimation has a relationship with the cost estimation accuracy of the projects they involved in previously (H2).
3. The participants beliefs that the contractual procedure effect on cost estimation has a relationship the cost estimation accuracy of the projects they involved in previously (H3).
4. The participants beliefs that the effect of the organization management has a relationship on the effect of information quality on cost estimation accuracy (H4).
5. The participants beliefs that the effect of the organization management has a relationship with their beliefs on the effect of estimator performance on cost estimation (H5).
6. The participants beliefs that the information quality effect on cost estimation has a relationship with their beliefs on the effect of project characteristics on cost estimation (H6).
7. The participants beliefs that the project characteristics effect on cost estimation has a relationship with their beliefs on the effect of external factors on cost estimation (H7).
8. The participants beliefs that the external factors effect on cost estimation has a relationship with their beliefs on the effect of contractual procedure on cost estimation (H8).
9. The participants beliefs that the information quality effect on cost estimation has a relationship with their beliefs on the effect of external factor on cost estimation (H9).

10. The participants' beliefs that the estimator's performance effect on cost estimation has a relationship with their beliefs on the effect of information quality on cost estimation (H10).
11. The participants' beliefs that the estimator's performance effect on cost estimation has a relationship with their beliefs on the effect of external factors on cost estimation (H11).

5.5.4 The structural equation modeling results of contractor model

The proposed model was statistically significant with most of the proposed relationships have path coefficients (b) with a significance value ($p < 0.01$) as shown in Figure 5-14 and Table 5-33. Four categories converged on the cost estimation error and had a statistical significance relationship and effect size of $R^2 = 0.21$, $p < 0.01$. However, two relationships were not statistically significant ($CP \rightarrow EA$, and $PC \rightarrow EA$). The model has a statistically significant Tenenhaus Goodness of Fit (GoF) value of 0.546.

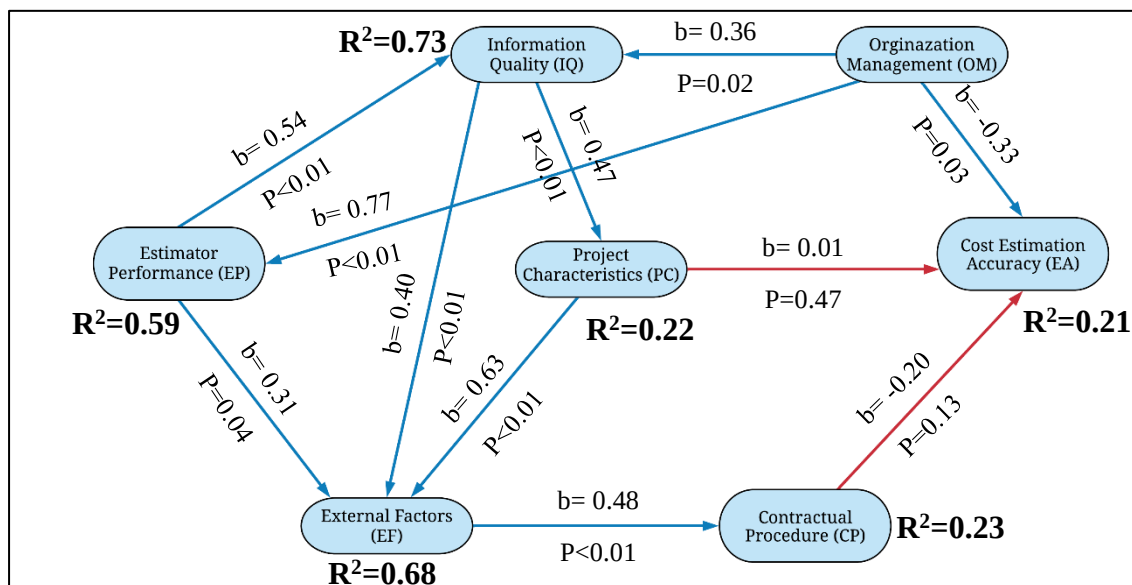


Figure 5-14 Results of contractor PLS-SEM model

Table 5-33 Results of contractor PLS-SEM model

Hypothesis	b(p-value)	Conclusion
H1: $OM \rightarrow EA$	-0.33(0.03)	Hypothesis supported
H2: $PC \rightarrow EA$	+0.01(0.47)	Hypothesis unsupported
H3: $CP \rightarrow EA$	-0.20(0.13)	Hypothesis unsupported
H4: $OM \rightarrow IQ$	+0.36(0.02)	Hypothesis supported
H5: $OM \rightarrow EP$	+0.77(<0.01)	Hypothesis supported
H6: $IQ \rightarrow PC$	+0.47(<0.01)	Hypothesis supported
H7: $PC \rightarrow EF$	+0.63(<0.01)	Hypothesis supported
H8: $EF \rightarrow CP$	+0.48(<0.01)	Hypothesis supported
H9: $IQ \rightarrow EF$	+0.40(<0.01)	Hypothesis supported
H10: $EP \rightarrow IQ$	+0.54(<0.01)	Hypothesis supported
H11: $EP \rightarrow EF$	+0.31(0.04)	Hypothesis supported

5.5.5 Cost estimation accuracy: common categories model

The proposed structure of relationships among factors related to cost estimation accuracy from common categories point of view are modeled as shown in Figure 5-15.

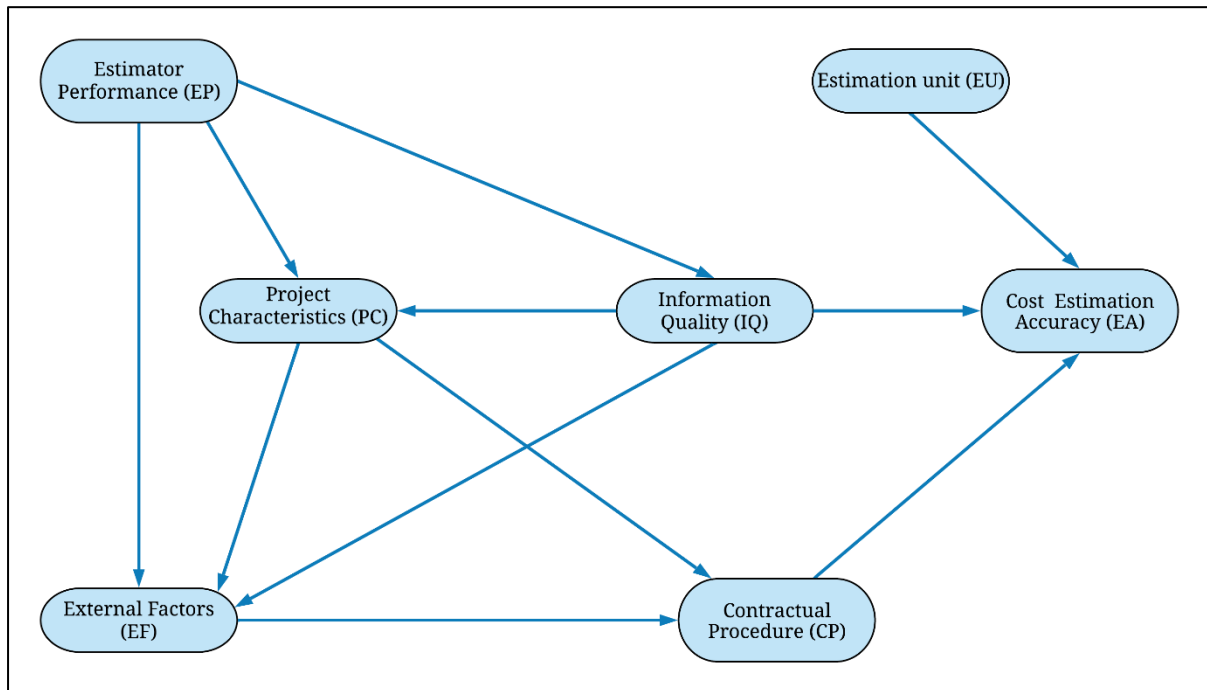


Figure 5-15 Proposed common categories between participants PLS-SEM model

The proposed model stipulates the following hypotheses:

1. Possession of dedicated cost estimation unit affects the magnitude of cost estimation accuracy (H1).
2. The participants beliefs that the information quality effect on cost estimation has a relationship with the cost estimation accuracy of the projects they involved in previously (H2).
3. The participants beliefs that the contractual procedure effect on cost estimation has a relationship with the cost estimation accuracy of the projects they involved in previously (H3).
4. The participants beliefs that the estimator's performance effect on cost estimation has a relationship with their beliefs on the effect of information quality on cost estimation (H4).
5. The participants beliefs that the estimator's performance effect on cost estimation has a relationship with their beliefs on the effect of project characteristics on cost estimation (H5).
6. The participants beliefs that the information quality effect on cost estimation has a relationship with their beliefs on the effect of project characteristics on cost estimation (H6).
7. The participants beliefs that the external factors effect on cost estimation has a relationship with their beliefs on the effect of contractual procedure on cost estimation (H7).
8. The participants beliefs that the project characteristics effect on cost estimation has a relationship with their beliefs on the effect of contractual procedure on cost estimation (H8).
9. The participants beliefs that the estimator performance effect on cost estimation has a relationship with their beliefs on the effect of external factors on cost estimation (H9).

10. The participants' beliefs that the information quality effect on cost estimation has a relationship with their beliefs on the effect of external factors on cost estimation (H10).
11. The participants' beliefs that the project characteristics effect on cost estimation has a relationship with their beliefs on the effect of external factors on cost estimation (H11).

5.5.6 The structural equation modeling results of common categories model

The proposed model was statistically significant with most of the proposed relationships have path coefficients (b) with a significance value ($p < 0.01$) as indicated in Figure 5-16 and Table 5-34. Four categories converged on the cost estimation accuracy and had a statistical significance relationship and moderately large effect size of $R^2 = 0.24$, $p < 0.01$. However, two relationships were not statistically significant ($EF \rightarrow CP$ and $EP \rightarrow EF$). The model has a statistically significant Tenenhaus Goodness of Fit (GoF) value of 0.529.

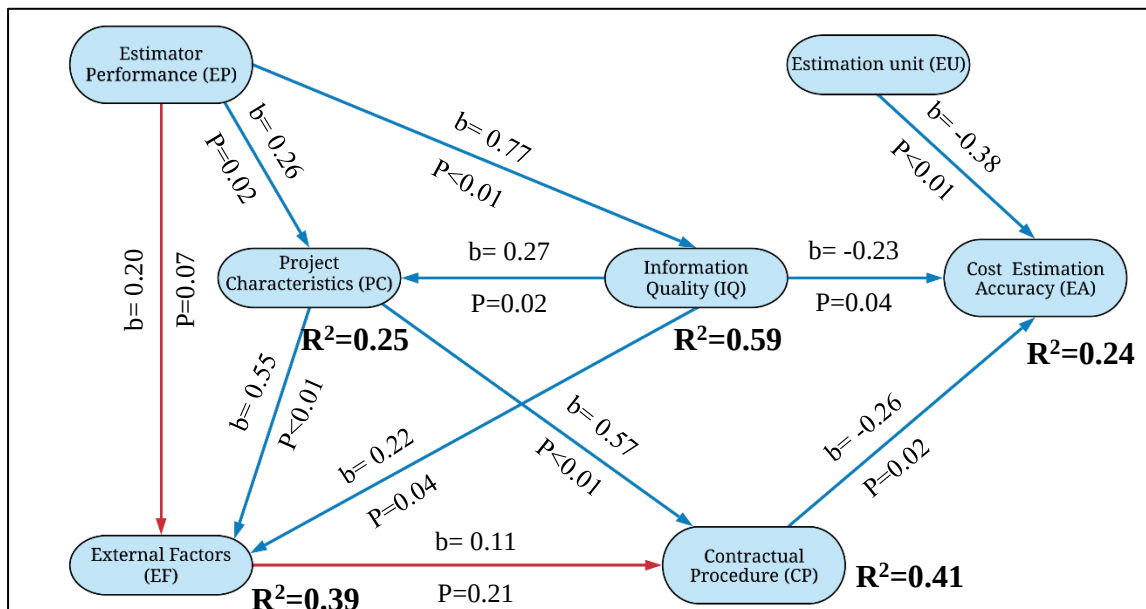


Figure 5-16 Results of common categories between participants PLS-SEM model

Table 5-34 Results of common categories between participants PLS-SEM model

Hypothesis	b(p-value)	Conclusion
H1: $EU \rightarrow EA$	-0.38(<0.01)	Hypothesis supported
H2: $IQ \rightarrow EA$	-0.23(0.04)	Hypothesis supported
H3: $CP \rightarrow EA$	-0.26(0.02)	Hypothesis supported
H4: $EP \rightarrow IQ$	+0.77(<0.01)	Hypothesis supported
H5: $EP \rightarrow PC$	+0.26(0.02)	Hypothesis supported
H6: $IQ \rightarrow PC$	+0.26(0.02)	Hypothesis supported
H7: $EF \rightarrow CP$	+0.11(0.21)	Hypothesis unsupported
H8: $PC \rightarrow CP$	+0.57(<0.01)	Hypothesis supported
H9: $EP \rightarrow EF$	+0.20(0.07)	Hypothesis unsupported
H10: $IQ \rightarrow EF$	+0.22(0.04)	Hypothesis supported
H11: $PC \rightarrow EF$	+0.55(<0.01)	Hypothesis supported

5.5.7 Selection of research model

After Modeling the relationship between the factors affecting construction cost estimation and the observed cost estimation accuracy, primary assessment was implemented to select the most suitable model for further analysis. Table 5-35 contains a comparison between the three proposed models with respect to PLS-SEM design consideration and goodness of fit criteria.

Table 5-35 Comparison between PLS-SEM models

Proposed model	Follow design consideration	Number of significant paths	Number of insignificant paths	R²	GoF
Client and consultant	No	8	6	0.32	0.555
Contractor	No	9	2	0.21	0.546
Common categories	Almost	9	2	0.24	0.529

Despite the high coefficient of determination of client and consultant model value, almost 40% of the relations were found insignificant with GoF model of 0.555. Both contractor and all common categories model had the same number of significant and insignificant relationships, with slightly higher R^2 value found with common categories model ($R^2=0.24$). All the models have a large Tenenhaus GoF value (above criteria threshold 0.36), $GoF>0.360$ indicates large model fit with all relationships highly significant.

Comparing degree of models' fulfillment to the PLS-SEM design consideration, theoretically both client-consultant and contractors models have violated the PLS-SEM model design consideration with respect to sample size required to perform the model. Model three sample size is very closed to the required sample size (59). It can be concluded from the previous facts and the current research circumstances that model three is the most representative proposed model to the research problem. It is needed to be mentioned that scope quality and organization management categories have been eliminated from the common categories model as recommended in design consideration guidance. Having construct with an amount of missing data more than 15% increases the bias in the model and reduce the reliability of the model.

5.5.8 Assessment and validation of PLS-SEM model

For the selected model, Table 5-36 shows the measurement model results along with assessment criteria threshold for: reliability, composite reliability, cross loading, and average variance extracted.

Table 5-36 Assessment of common categories measurement model

	Categories	Factors ID	Cronbach's alpha	Composite reliability	Cross loading	AVE
	Threshold criteria		>0.7	0.7	>0.708	>0.5
1	Estimator performance	EP_1	0.755	0.845	0.807	0.579
		EP_2			0.769	
		EP_3			0.798	
		EP_4			0.66	
2	Information quality	IQ_1	0.816	0.891	0.798	0.733
		IQ_2			0.924	
		IQ_3			0.842	
3	External factors	EF_1	0.889	0.924	0.842	0.752
		EF_2			0.917	
		EF_3			0.911	
		EF_4			0.791	
4	Contractual procedure	CP_2	0.55	0.816	0.83	0.69
		CP_6			0.83	
5	Project characteristics	PC_1	0.789	0.848	0.563	0.454
		PC_3			0.706	
		PC_4			0.343	
		PC_8			0.767	
		PC_9			0.775	
		PC_10			0.778	
		PC_11			0.706	

Generally, model three reflective measurement model meet the relevant assessment criteria. Both internal reliability and composite reliability of latent variables are between 0.755 and 0.924, which is higher than minimum level accepted (above 0.7). An exception found in the CP with Cronbach alfa of 0.55. Low Cronbach alfa value can be justified by the low number of indicators representing the latent variable (contains two factors only).

Concerning the convergent validity, the latent variables AVE values obtained are between 0.454 and 0.752. Screening the values of indicators cross loading, PC_1 (Type/function) and PC_4 (size/gross floor area) cross loadings are below 0.708. Dropping PC_4 (size/gross floor area) from the model increased the PC convergent into 0.518, which is higher than the threshold specified (above

0.5) and slight changes in path coefficients are observed (± 0.01 of paths coefficient values). Therefore, the model convergent becomes acceptable, all indicators cross loading are above 0.708 and other low cross loading values will be accepted as the AVE obtained is larger than 0.5. Estimation unit is not included in the measurement model assessment because it has a single factor, and it will be assessed by the standard evaluation criteria. Thus, it is concluded that the measurement model has been established.

Concerning the discriminant validity, Table 5-37 shows the Locker and Fornell criterion results. The diagonal expresses the square root of the average variance extracted, while the remaining values are the correlation between the constructs. Comparing the values obtained, the results suggest that latent variables are different from each other as each category explains the variance of its own factor rather than the variance of other categories.

A latent construct should explain better the variance of its own indicator rather than the variance of other latent constructs. Therefore, the square root of each construct's AVE should have a greater value than the correlations with other latent constructs.

Table 5-37 Locker and Fornell criterion results for common categories model

	EP	IQ	EF	EA	CP	EU	PC
EP	(0.761)						
IQ	0.761	(0.856)					
EF	0.158	0.204	(0.867)				
EA	0.065	-0.12	0.014	(1)			
CP	0.292	0.338	0.398	-0.19	(0.83)		
EU	-0.3	-0.024	-0.015	-0.336	-0.135	(1)	
PC	0.411	0.425	0.543	-0.105	0.574	-0.097	(0.72)

The evaluation of outer loadings of all factors on the associated categories are tabulated in Table 5-38. The outer loading indicated correlation coefficient between the factor and its related category. It can be observed that the outer loading associated to the factors in one category are greater than all the outer

loadings appeared in remaining categories. Therefore, the cross-loading criteria related to discriminant validity is fulfilled.

Table 5-38 Outer loading results for common categories of the model

	EP	IQ	EF	EA	CP	EU	PC
EP_1	0.807	0.504	-0.089	0.109	0.099	-0.083	0.045
EP_2	0.769	0.123	0.065	-0.036	-0.203	0.057	0.118
EP_3	0.798	0.197	-0.147	-0.152	-0.043	-0.3	0.089
EP_4	0.66	-0.997	0.21	0.092	0.167	0.398	-0.3
IQ_1	0.212	0.798	0.135	-0.119	-0.033	0.232	-0.202
IQ_2	-0.208	0.924	-0.021	0.141	0.024	-0.153	0.046
IQ_3	0.027	0.842	-0.104	-0.042	0.004	-0.052	0.142
EF_1	-0.289	0.186	0.842	-0.194	-0.245	-0.131	0.152
EF_2	0.111	-0.072	0.917	-0.023	-0.037	0.023	-0.039
EF_3	-0.04	-0.053	0.911	0.107	0.124	0.062	0.013
EF_4	0.226	-0.054	0.791	0.11	0.162	0.042	-0.132
EA	0	0	0	1	0	0	0
CP_2	-0.304	0.164	-0.143	0.21	0.83	-0.075	0.108
CP_6	0.304	-0.164	0.143	-0.21	0.83	0.075	-0.108
EU	0	0	0	0	0	1	0
PC_1	-0.527	0.431	-0.188	0.391	0.111	-0.12	0.563
PC_3	-0.019	-0.068	-0.182	0.134	0.333	0.004	0.706
PC_8	-0.273	0.479	0.012	0.027	-0.313	0.059	0.767
PC_9	-0.096	0.123	-0.032	-0.393	-0.254	-0.12	0.775
PC_10	0.756	-0.817	0.109	-0.014	0.238	0.43	0.778
PC_11	0.008	-0.031	0.233	-0.029	-0.065	-0.314	0.706

Considering the structural model assessment, Table 5-39 indicates the results of variance inflation factors (VIF) and model predictive relevance. VIF values concluded that the data set does not counter a collinearity issue as all the VIF values are below the threshold (VIF< 5).

Table 5-39 VIF and Q results for common categories of the model

	Threshold criteria	EP	IQ	EF	EA	CP	EU	PC
VIF	<5	3.039	2.86	1.486	1.249	1.638	1.381	2.042
Q²	>0	N/A	0.6	0.327	0.246	0.371	N/A	0.249

The findings related to Q-squared (Q^2) criteria found in Table 5-39 are acceptable and reflects a good model with acceptable predictive relevance. The Q^2 coefficients for both EP and EU have not being generated in the software because these values are provided only for endogenous latent variables.

Table 5-40 Summary of path coefficient, R^2 , effect size for results for common categories of the model

Hypothesis	b (p-value)	R^2	Path contribution (f^2 effect size)	Effect size
H1: $EU \rightarrow EA$	-0.38 (<0.01)	0.24	0.128	Small
H2: $IQ \rightarrow EA$	-0.23 (0.04)		0.040	Small
H3: $CP \rightarrow EA$	-0.26 (0.02)		0.070	Small
H4: $EP \rightarrow IQ$	+0.77 (<0.01)	0.59	0.595	Large
H5: $EP \rightarrow PC$	+0.27 (0.02)	0.24	0.124	Small
H6: $IQ \rightarrow PC$	+0.26 (0.02)		0.118	Small
H7: $EF \rightarrow CP$	+0.14 (0.14)	0.38	0.060	Small
H8: $PC \rightarrow CP$	+0.53 (<0.01)		0.324	medium
H9: $EP \rightarrow EF$	+0.20 (0.07)	0.39	0.032	Small
H10: $IQ \rightarrow EF$	+0.22 (0.05)		0.054	Small
H11: $PC \rightarrow EF$	+0.56 (<0.01)		0.304	medium

Table 5-40 summarized the key quality criteria in the structural model assessment. Considering the explanatory power of the independent variables connected to dependent variable, the combined effect of estimation unit, information quality and contractual procedure is found significant and it explains 24% of the variance occurring in cost estimation error. It has been observed that estimation unit has the strongest effect on cost estimation accuracy, followed by contractual procedure and information quality with f^2 values near to zero.

When analyzing the key predictors, the relationship between estimator performance and information quality is found significant and it explains 59.0% of the variance occurring in information quality. The effect size of estimator performance is significantly high. On the other hand, the combined effect of estimator performance and information quality is found significant and it explains 24.0% of the variance occurring in project characteristics. However, f^2 values of the path contribution is less than 0.15 which indicate small effect.

An overview for the relationship related to contractual procedure, the combined effect of external factors and project characteristics explains 38.0 % of the variance occurring in contractual procedure. The results substantiate that the effect of project characteristics is significant and has a medium effect on CP, while the external factors do not have significant effect at the 5% probability level.

Finally, for the relationship related to external factors, the combined effect of estimator performance, information quality, and project characteristics explains 39.0% of the variance occurring in external factor. The results substantiate that the effects of project characteristics and information quality are significant, and both have a medium and small effect on external factors, respectively, while the estimator performance does not have significant effect at the 5% probability level.

In this model the global fit index is 0.529 as stated in page 122, which indicates that the empirical data fits the model very well.

5.6 Satisfaction with Construction Cost Estimation Model

5.6.1 Neural network model development

The neural network has been trained with three different learning algorithms and different hidden neurons in order to discover the network structure

that's derives a sufficient predictive power with optimum neuron number. Seventy percent of the data set has been used to training the neural network model. Whereas the remaining thirty percent of the data has been used for validation and testing the developed model.

ANN networks modeled using both Bayesian generalization and Conjugate Gradient Algorithms exhibit low predictive models with coefficient of correlation and MSE values within the ranges shown in Table 5-41.

Table 5-41 The results of best neural network performance using Bayesian generalization and Scaled Conjugate Gradient Algorithms

Bayesian generalization Algorithm						
Hidden layer structure	algorithm	MSE train	MSE testing	Rtrain	Rtest	Rall
1-8N	BG	0.845	1.872	0.39	0.145	0.306
1-9N		0.938	1.014	0.342	0.153	0.291
1-10N		1.11	0.502	0.262	0.544	0.289
Scaled Conjugate Gradient Algorithms						
1-7	SCG	0.66	0.806	0.669	0.471	0.562
1-9		0.656	0.805	0.649	0.2	0.578
1-11		0.633	0.401	0.677	0.9	0.671

Better performance results are obtained from performing Levenberg-Marquardt algorithm. Table 5-42 performance demonstrates the MSE values obtained during training the network with different number of neurons.

Table 5-42 The results of best neural network performance using Levenberg-Marquardt learning algorithms neural networks

Hidden layer structure	MSE train	MSE validate	MSE testing	R train	R validation	R test	R all
LM-9N	0.219	0.556	1.07	0.892	0.427	0.82	0.817
LM-10N	0.1915	0.75	0.606	0.908	0.569	0.803	0.844
LM-11N	0.172	0.584	0.758	0.921	0.781	0.461	0.824

Reviewing the results, it can be concluded that the network trained by Levenberg-Marquardt backpropagation algorithm with 10 hidden neurons (LM-10N) has the best performance network as it outperforms the other networks in predicting an accurate target value with an overall coefficient of correlation of $R=0.844$ as shown in Figure 5-17.

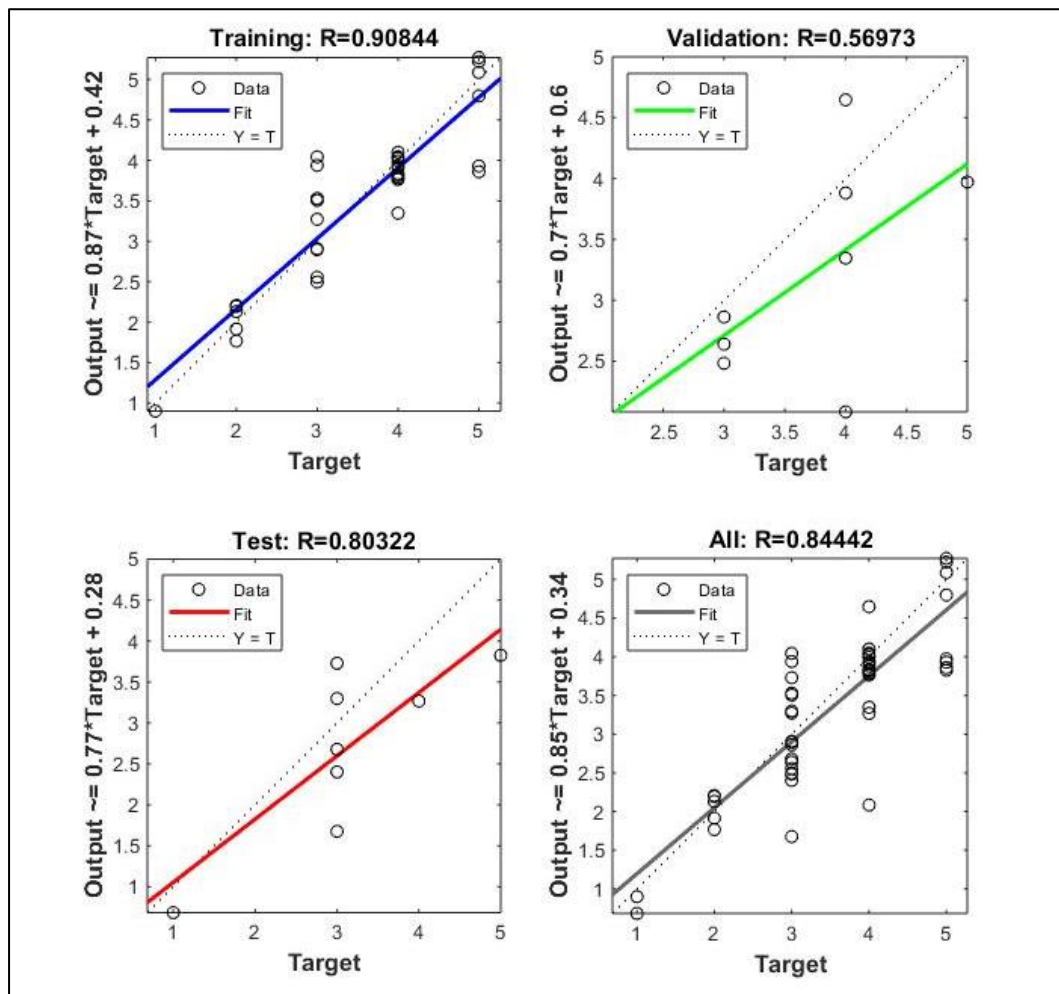


Figure 5-17 Coefficient of correlation of LM-10N neural network model

Figure 5-17 shows the plot of the actual network outputs in terms of the associated target values in different model development phase. Initially, the model exhibits an accurate prediction of target value in the training phase with $R=0.908$. Although that the R value reduced to 0.508 in the validation phase, but the output prediction performance becomes better in the testing phase with R value equal to 0.804.

The best performance in validation set has been recognized at epoch 6 with a performance of 0.750 as shown in Figure 5-18. For more illustration, after 6 training epochs, the error started to increase on the validation data set as the network starts overfitting the training data. Therefore, the training stops after the increase in validation error, and the best performance is taken from the epoch 6 that has the lowest validation error.

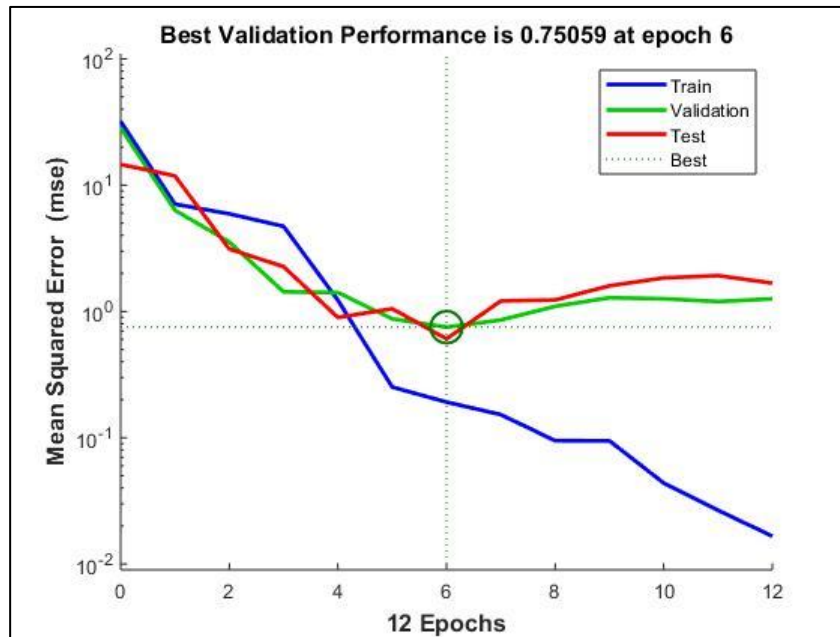


Figure 5-18 Validation performance for LM-10N model

Figure 5-19 shows the error sizes distribution of the developed model. Generally, the output error ranges between $(-0.96, 1.84)$ with most errors clustered near zero. The model mean square errors calculated using the Equation 3-10 and MSE values found are 0.191, 0.75 and 0.606 in training, validation and testing stages, respectively. This result suggest that the neural network had a high predictive ability in training phase, and it is reduced in validation and testing phase. It can be interpreted by the small sample size used in developing the model.

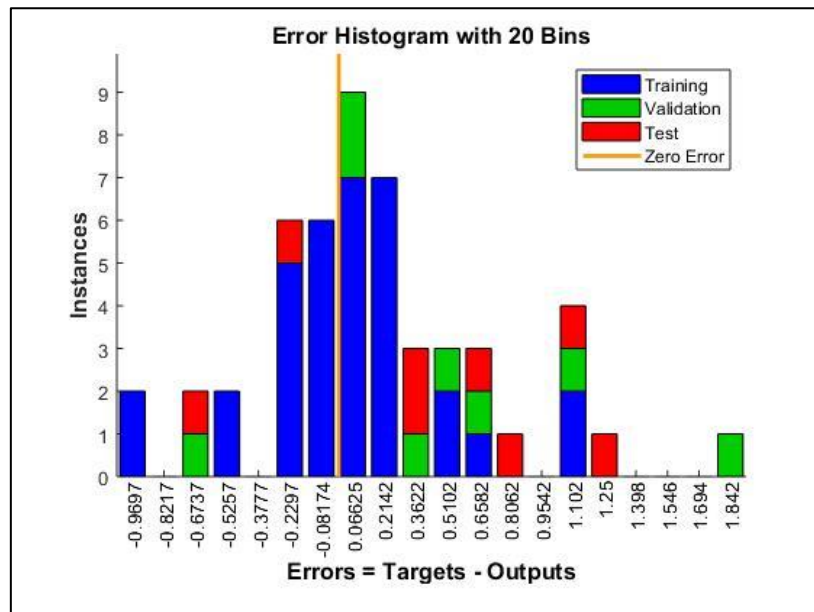


Figure 5-19 Error distribution for LM-10N model

The network architect is illustrated in Figure 5-20 and Figure 5-21. The architecture schematics of the developed neural network consists of five inputs trained by Levenberg-Marquardt backpropagation in 10 hidden neurons to predict one output.

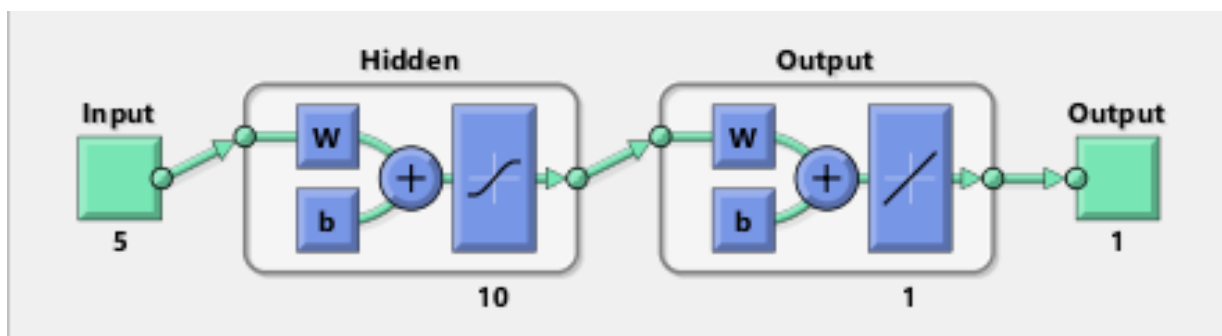


Figure 5-20 ANN diagram for model 5-10-1

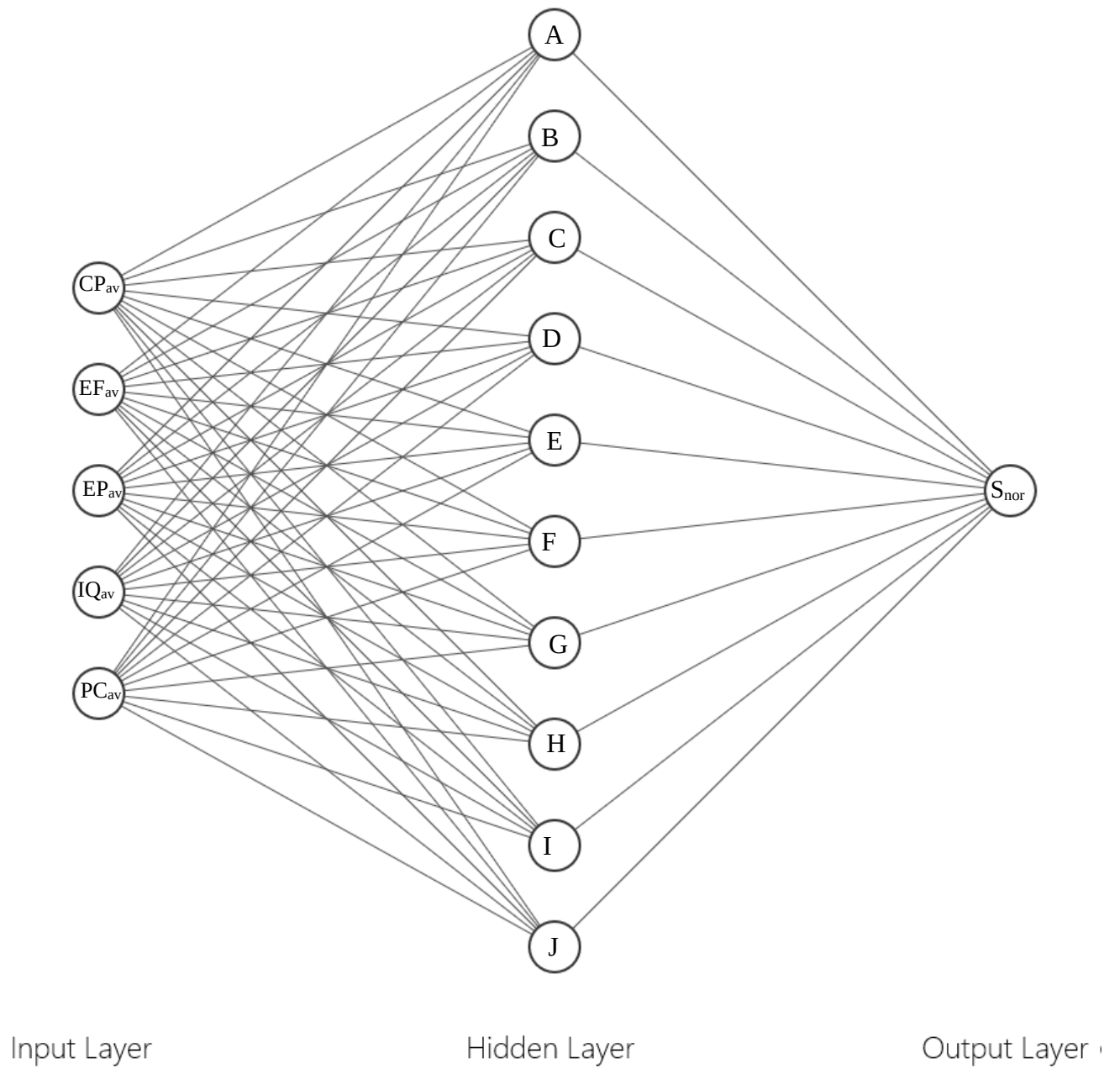


Figure 5-21 The architecture schematics for 5-10-1 ANN model

5.6.2 Neural network mathematical model development

In order to develop the mathematical equation that predicts the level of participant satisfaction on his construction cost estimation, the hidden weights of input, output, and bias has been extracted from the best performed model, i.e., LM-10N, and tabulated in Table 5-43.

Table 5-43 (5-10-1) ANN hidden input, output, and bias connection weight

Hidden node	Input hidden weights					Hidden output weight	bias	
	CP _{av}	EF _{av}	EP _{av}	IQ _{av}	PC _{av}	satisfaction	hidden	output
A	-3.074	1.222	1.854	0.707	-3.114	1.390	-1.692	-0.794
B	-1.733	4.105	1.182	-3.482	-1.552	-0.526	-1.533	
C	4.331	-0.015	-2.972	-0.988	-1.062	-0.438	0.573	
D	0.626	-4.115	-1.153	2.807	-0.689	-0.055	-1.325	
E	-0.516	1.635	2.495	0.247	-4.928	-0.891	-0.123	
F	-1.159	-0.331	-0.445	1.289	-0.285	0.006	-1.234	
G	1.237	1.255	-0.790	-0.607	-0.784	0.982	1.093	
H	-1.155	1.908	-0.899	3.707	2.004	-0.046	1.278	
I	2.429	-1.071	-0.437	-1.623	-0.402	1.858	2.825	
J	1.475	0.844	-1.157	-1.813	-0.681	-1.109	2.358	

The hidden nodes equations are generated by substituting connection weights and bias values of the model given in Table 5-43 in the mathematical expression of Equation 3-6 as follows:

$$A = -3.074 \text{ CP}_{av} + 1.222 \text{ EF}_{av} + 1.854 \text{ EP}_{av} + 0.707 \text{ IQ}_{av} - 3.114 \text{ PC}_{av} - 1.692$$

$$B = -1.733 \text{ CP}_{av} + 4.105 \text{ EF}_{av} + 1.182 \text{ EP}_{av} - 3.482 \text{ IQ}_{av} - 1.552 \text{ PC}_{av} - 1.533$$

$$C=4.331 CP_{av}-0.015EF_{av}-2.972 EP_{av}-0.988 IQ_{av} -1.062 PC_{av} +0.573$$

$$D=0.626 CP_{av}-4.115 EF_{av}-1.153EP_{av}+2.807 IQ_{av}-0.689 PC_{av}-1.325$$

$$E=-0.516 CP_{av} +1.635 EF_{av} +2.495 EP_{av} +0.247 IQ_{av} -4.928 PC_{av} -0.123$$

$$F=-1.159 CP_{av} -0.331 EF_{av} -0.445 EP_{av} +1.289 IQ_{av} -0.285 PC_{av} -1.234$$

$$G=1.237 CP_{av} +1.255 EF_{av} -0.790 EP_{av} -0.607 IQ_{av} -0.7845 PC_{av} +1.093$$

$$H=-1.155 CP_{av} +1.908 EF_{av} -0.899 EP_{av} +3.707 IQ_{av} +2.004 PC_{av} +1.278$$

$$I=2.42917 CP_{av} -1.071 EF_{av} -0.437 EP_{av} -1.623 IQ_{av} -0.402 PC_{av} +2.825$$

$$J=1.475 CP_{av} +0.844 EF_{av} -1.157 EP_{av} -1.8131 IQ_{av} -0.681 PC_{av} +2.358$$

Equation 5-1

Where A, B, C, ..., J are the hidden nodes, Cp_{av} , EF_{av} , EP_{av} , IQ_{av} and PC_{av} are normalized scores obtained from the questionnaire survey.

An important step before the mathematical computation of the nodes values is to standardize the input values between (-1,1) using Equation 3-6. The results of a standardized input variables used to establish the hidden nodes equation are shown in Table 5-44.

Table 5-44 Standardization of input variables to establish the ANN hidden nodes equations

	CP_av	EF_av	EP_av	IQ_av	PC_av
Input value	5	4	4.25	5	5
Minimum	1.5	2.5	2.75	2.333	2.5
Maximum	5	5	5	5	5
Normalized input	1.000	0.200	0.333	1.000	1.000

The multiplication result of the normalized input variables and connection nodes are tabulated in Table 5-45.

Table 5-45 Multiplication results of hidden connection weight by the standardized input variables

Hidden nodes	hidden input connection weight*standardized input variables				
	CP_av	EF_av	EP_av	IQ_av	PC_av
A	-3.074	0.244	0.618	0.707	-3.114
B	-1.733	0.821	0.394	-3.482	-1.552
C	4.331	-0.003	-0.991	-0.988	-1.062
D	0.626	-0.823	-0.384	2.807	-0.689
E	-0.516	0.327	0.832	0.247	-4.928
F	-1.159	-0.066	-0.148	1.289	-0.285
G	1.237	0.251	-0.263	-0.607	-0.784
H	-1.155	0.382	-0.300	3.707	2.004
I	2.429	-0.214	-0.146	-1.623	-0.402
J	1.475	0.169	-0.386	-1.814	-0.681

After that, nodes values are obtained by summing the values of hidden neuron values as mentioned in Table 5-46.

Table 5-46 Results of sum hidden neuron values with hidden bias and connection weight to the output neuron

Node	Sum hidden neuron values with hidden bias	Hidden output connection
A	-6.311	1.390
B	-7.085	-0.526
C	1.860	-0.438
D	0.211	-0.055
E	-4.162	-0.891
F	-1.604	0.006
G	0.926	0.982
H	5.916	-0.046
I	2.869	1.858
J	1.122	-1.109

The nodes coefficients have been summed and the value obtained has been transferred using tan-sigmoid transfer function. The multiplication result of the transferred sum hidden layer values and output connection nodes is tabulated in Table 5-46 and presented in Figure 5-22.

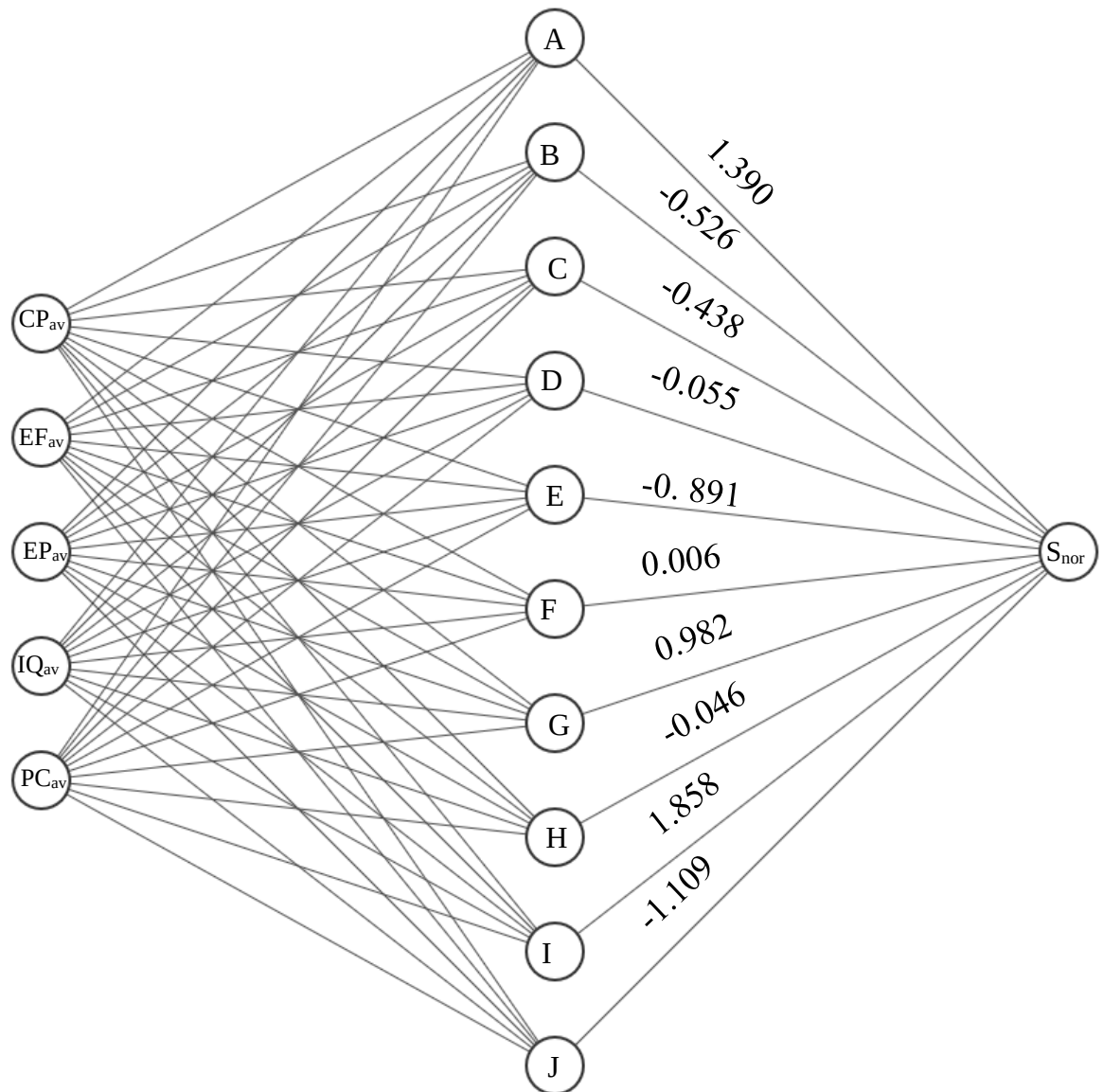


Figure 5-22 Multiplication result of the transferred sum hidden layer values and output connection nodes for 5-10-1 ANN model

The derivation of the target value (S_{nor}) using the transfer function Tanh in normalized form is expressed by the following equation:

$$S_{nor} = 1.390 \tanh(A) - 0.526 \tanh(B) - 0.438 \tanh(C) - 0.055 \tanh(D) - 0.891 \tanh(E) + 0.006 \tanh(F) + 0.982 \tanh(G) - 0.046 \tanh(H) + 1.858 \tanh(I) - 1.109 \tanh(J) - 0.794 \quad \text{Equation 5-2}$$

The S_{nor} equation of participant satisfaction level to be de-normalized to predict the level of cost estimation satisfaction. The mathematical equation derived is: **Satisfaction, $S_a = 1.75S_{nor} + 3.25$** . The calculation sheet of the previous computations is shown in Appendix 9- Table 2.

5.6.3 Neural network model cross-validation.

The results of MOW's questionnaire have been used to cross-validate the performance of LV-10N neural network model. MOW data introduced to predict the satisfaction of construction cost estimation. Figure 5-23 shows the plot of the actual network outputs in terms of the associated target values. The model predicted the cross-validation data output with a value of $R=0.69$.

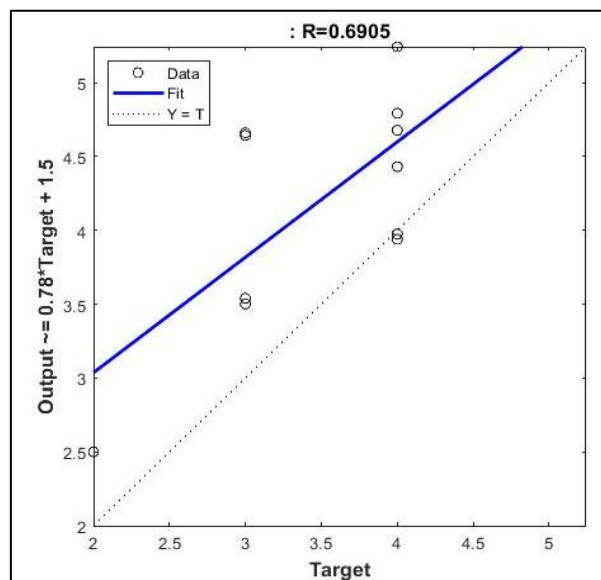


Figure 5-23 Coefficient of correlation for LM-10N model using cross validation data

Figure 5-24 shows the error sizes distribution of the developed model. The output error ranges between (-1.617, 0.018). The model mean square error found to be MSE=0.82. The result suggest that the neural network had an acceptable predictive ability.

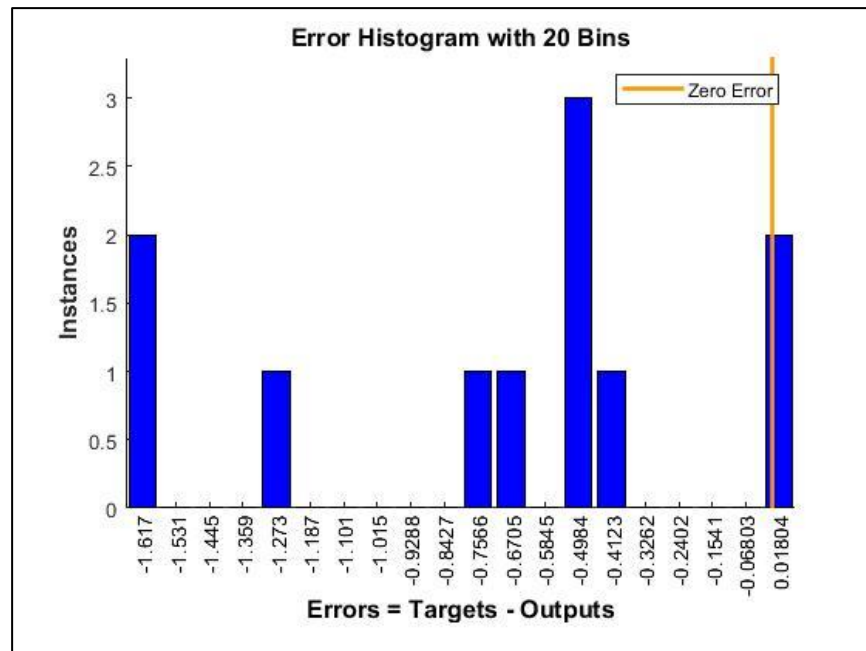


Figure 5-24 Error distribution for LM-10N model using cross validation data

5.6.4 Sensitivity analysis

With an attention to define the factors affecting MYS cost estimation accuracy, sensitivity analysis has been performed to address this objective. The results of factors affecting the MYS construction estimation ranking using connection weight method represented in Table 5-47. The findings emphasized that the information quality has the greatest effect on satisfaction level with a value of (1) followed by estimator performance (0.93). Also, the high influence of project characteristics is observed with a relative importance value of (0.87). Whilst lower impact has been noticed for the remaining input parameters: contractual procedure (0.605) and external factors (0).

The type of influence was found negative with respect to external factors and contractual procedure which implies indirect relation to participant satisfaction level on cost estimation. Whilst positive influence was detected for estimator performance, information quality, and project characteristics on satisfaction on construction cost estimation. The calculation sheet for implementation of connection weight method can be found in Appendix 10.

Table 5-47 Categories ranking as per normalized connection weight method

Num	Categories	Sum of importance	Normalized	Rank	Type of influence
1	Contractual procedure	-0.694	0.606	4	Negative
2	External factors	-3.469	0.000	5	Negative
3	Estimator performance	0.834	0.940	2	Positive
4	Information quality	1.111	1.000	1	Positive
5	Project characteristics	0.524	0.872	3	positive

5.7 Chapter Summary

This chapter aims to report the findings of the level of construction cost estimation accuracy in MYS, construction cost model development, and the factors affecting the cost estimation process in MYS. The chapter reported the hypothesized results about the relationship between the factors and cost estimation accuracy using PLS-SEM three different perspective models: client/consultant, contractors, all participants. The accepted model is used to develop a mathematical model that quantifies the influence of factors on cost estimation satisfaction using ANN.

CHAPTER 6: Results and Discussion

6.1 Introduction

6.2 Level of Accuracy of MYS Cost Estimation Process

6.3 Factors Affecting Cost Estimation Accuracy at MYS

6.4 Factors Affecting Cost Estimation Model

6.5 Chapter Summary

CHAPTER 6: Results and Discussion

6.1 Introduction

This chapter aims to summarize the findings of the three main objective of the dissertation: level of cost estimation accuracy in MYS, the factors affecting the cost estimation process in MYS, and the proposed construction cost estimation model for MYS. The chapter will compare the results with prior studies.

6.2 Level of Accuracy of MYS Cost Estimation Process

Considering the data set used for the evaluation of cost estimation accuracy in MYS construction projects, the performance of construction cost estimation process is within the international accuracy criteria in the pre-tender stage with an overestimation tendency (mean =3.47%). The overestimation pattern of construction cost estimation has been observed in the past literature, MYS performance substantially better than prior studies: 5.13% in USA, 4.29% in Australia, 11.18% in Malaysia but not better than Nigeria, where the mean cost deviation found is 2.43%.

Nevertheless, the extensive results demonstrate existence of inconsistency in cost estimation determination, as overestimation and underestimation bias proportion found to be 59.1% and 40.9%, respectively, with a mean cost deviation that is significant from zero. This evidence confirms that the cost deviation is not a random event. A similar status has been reported by (Aibinu and Pasco, 2008), and other recent GCC literature confirmed the deficiency of cost estimation with a large cost escalation, which generalizes a lack of cost performance and control in the region.

The Construction project data exhibit a disparity between the cost deviation and its associated standard deviation as shown in Table 5-5. This can be

interpreted due to the widespread of projects cost deviation values in each criterion (overall, underestimation, overestimation), which leads to a more flatter data distribution curve.

The construction projects data gathered suggests that MYS cost estimation performance is affected by the project size as shown in Table 5-11. The performance of construction cost estimation found to be overestimated with projects below BD225,500 and underestimated with the projects above this amount.

Furthermore, MYS tends to overestimate sport halls projects, playgrounds projects with noticeable cost deviation (3.48% and 6.53%). Interestingly, the estimation performance gets substandard in maintenance projects with a large deviation from the estimated cost (13.38%). This can be interpreted through MYS reliance on international design standards in the main decisions related to projects' significant cost items during the design phase of the playgrounds and sport halls, which give them a better opportunity to predict the projects costs. The exception found in the administration building category which has mean cost deviation tended to be underestimated with a high disparity (cost deviation - 4.16%, standard deviation 24.36%).

Considering the project awarding period, cost estimation for the projects performed between 2012-2013 and 2014-2015 has experienced a cost deviation with average of 12.77% and 6.20%, respectively. This may be interpreted as a good practice in considering a management buffer during the cost estimation process. In contracts, the budget allocation in following years (2016-2017, 2018-2019) faced considerable cost underestimation (-1.51%, -8.55%). This result suggests a change in cost estimation practice in MYS.

When the effect of projects function, project size and year of awards have been investigated, no credible evidence found regarding the effect of project

function, size, and year of award on cost estimation as shown in Table 5-11. The present study confirmed the effect of project size on construction cost estimation, as the results showed a statistical significance between two categories (less than BD225,500 and above BD625,000). This basic finding is consistent with researches of Oladokun et al. (2011) and Aibinu and Pasco (2008) showing the possible effect of project size on the construction cost estimation accuracy.

Administration time between budget reservation and project awarding in MYS showed that approximately 47% of the construction projects spent an average duration of 216 days, 34% spent an average duration of 555 days, 18% spent an average duration of 911 days till the project has been awarded as given in Table 5-14. This finding illustrates cost estimator complex responsibility in forecasting an approximate project cost during the concept development phase and concurrently considering the possible effect of time on the cost during project life cycle.

With respect to construction cost estimation model, it is quite a surprise to find that both IBM SPSS statistics and Microsoft Excel software derived linear regression equations with large margin of error compared with the actual project cost estimation amount as shown in Appendix 7. These results suggest that extracting adequate construction cost models learning algorithm is complex, and the built-in algorithms in the mentioned software are not sufficient for the current data.

However, the regression models obtained from Matlab software exhibit high predictive ability. Stepwise regression model outperforms robust linear regression model. The visual plots and models statistic indicate more cost estimation accuracy using the following independent variables: project type, year of award group, award amount and cost deviation. Regression models validation confirms the previous conclusion, as stepwise regression model exhibit cost predication deviation range between -12.76 % to 5.56% while a cost deviation

range between -21.4% to -1.74% is obtained with the robust linear regression model.

Comparing stepwise regression model cost deviation range with (AACE) International Recommended Practice (No. 17R-97), the cost deviation is near to the acceptable cost deviation range which lies between -10 and +15%.

6.3 Factors Affecting Cost Estimation Accuracy in MYS

With respect to RII, there are similarities in participant perception about the critical factors affecting the cost estimation process between the present study and those reported in literature (Alumbugu et al., 2016; Hatamleh et al., 2018; Famiyeh et al., 2017).

Unexpectedly, this research has not found a significant influence for uncertainty level factors as found in recent studies (Akhund et al., 2019; Al-Sabah and Refaat, 2019). The only factor that has been recognized as critical was the price inflation. This inconsistency maybe due to stakeholders' experience with market condition, well understanding of construction delivery method, and the economic stability which makes the price fluctuation predictable and easy to be accounted through the contingency amount added in contract.

In addition, the low contribution of external factors related to governmental and environmental factors on construction cost estimation is found in line with the past studies reported (Al-Sabah and Refaat, 2019; Akhund et al., 2019; Amusan et al., 2018; Ali et al., 2018; Sridarran et al., 2017; Famiyeh et al., 2017). Although these factors are time consuming and have a direct influence on the construction cost, the results affirmed that external factors have minimal effect on construction cost estimation in MYS projects.

Concentrating on the critical factors' findings, one promising finding is the study highlight of the effect of stakeholders on the construction cost estimation process. The high pooled agreement found in scope quality and organization

status factors gives generally accepted explanation to the extent of management interference on determination of construction cost. The fact has been validated as 64% of the participant admitted that the cost estimation is being setup by the management level. This finding is consistent with research showing the significant influence of participants characteristics on cost performance (Lu et al., 2017; Abd El-Karim et al., 2017; Famiyeh et al., 2017).

The factors related to project characteristics (PC) had the highest pooled mean agreement, this results ties well with past study performed by (Oke, 2017; Sridarran et al., 2017). Nevertheless, comparing PC advance ranking obtained with similar researches performed in GCC region (Mahamid et al., 2014) and Palestine (Mahamid, 2018; Enshassi et al., 2013) indicates projects stakeholders awareness to the unique nature of construction projects executed in MYS, in consequence the parties are enforced to draw attention to the cost significant items driven from project characteristics.

The advanced rank of information quality has been in conformity with findings reported (Bakr, 2019). The results are accepted as the majority of participants reported a high dependency on historical data during the construction cost estimation (30% of participants relies on historical data with a percentage of 100%-90%, 42% of the participants relies on historical data with a percentage of 89%-70%) as shown in Table 5-19.

With slight differences in mean pooled agreements, quality of estimation, external factors and estimation process have same importance ranking as per participants opinions. Generally, these constructs already are interrelated and their influence start at project early stage. The amount of effort both client and consultant spend in preparation the contract documents will directly influence main contractor ability to accurately estimate the actual construction cost.

Surprisingly, the pool agreement for the contractual procedure is low with only two factors considered critical (number of bidders and contract condition).

Apart from the current study, contractual factors have advanced ranking affecting the cost estimation in many research (Mahamid et al., 2014; Enshassi et al., 2013). It may be argued that participants mitigate the risk of contractual factors by embedding their financial influence within bill of quantity clauses (i.e., preliminary clause and standard contingency amount), which ultimately represent a minor amount from the total project cost.

Comparing the mean categories scores based on participant managerial level, it was observed that the difference is not significant between management level and operational level perception on the factors affecting the cost estimation as shown in Table 5-30. Therefore, it can be argued that the management intervention in setting up the final project cost will affect the reliability of the construction cost estimation as declared by participants in Table 5-19.

The findings of structural model developed using PLS-SEM for the critical factors show clear support for possible relationships between the factors affecting the accuracy of construction cost estimation. The results suggest that dispossession a dedicated cost estimation unit will affect the magnitude of cost estimation error. This conclusion is generally accepted and verified through the existence of cost estimation unit in the organization and stakeholder satisfaction on construction cost estimation. Considering the main contractors category who already declared having a cost estimation unit, the awarded contract amount is satisfactory for almost 70% of the contractors who participated in the survey. Meanwhile, 30% is the degree of both client and consultant satisfaction of the construction cost estimation with a percentage of 58% admitted the absence of cost estimation unit in their organization. The consequences of lack of estimation unit in design development stage will impose a significant risk on projects budget allocation, which eventually will affect the project cost performance during its life cycle.

The results demonstrate the participants beliefs about the mediating role the information quality plays between the relationship of estimator's performance and cost estimation error occurred in the projects they involved in previously. The scales concatenation is logical, and it supports the workflow.

The combined effect of estimator performance and information quality on the project characteristics is found to be significant. Similarly, the effect of project characteristics is significant on external factors and contractual procedure. The results are quite expected and support the influence of the factors included in the categories.

With regard the insignificant relationships, the direct relationships between estimator performance on external factors and the external factors on contractual procedure are insignificant. The researcher speculates that this may be interpreted by the structure of relationship interaction the categories have on each other. The current model is more supporting complex mediating effects between the different categories than the direct relationships. Thus, the model suggests that the estimator performance interplays mediation effect on external factors through the information quality. With regard to the external factors and contractual procedure, the relationship may contain a mediating variable, which is not included in the current study.

The results excluded the scope quality and organization management from the significant categories, the relation of these important scales on the cost estimation accuracy remains unanswered. This exclusion due to the model limitation as PLS-SEM requires minimum number of missings at random values to reduce the bias in the model.

Although the model selected explains the variance in the cost estimation error with a percentage of 21% which is small to moderate, the model fits the data

with GoF of 0.529 and meets both measurement and structural model performance criteria.

6.4 Factors Affecting Cost Estimation Model

It has been confirmed the ability of ANN to detect the nonlinear relationship between the study variables. It was observed that Levenberg learning algorithm outperform both Bayesian and scale gradient learning algorithm in predicting the level of satisfaction. Applying the Levenberg algorithm on the ANN network resulted in better prediction to the satisfaction level using the network structure (5-10-1), which brings more trust in the trained ANN algorithm prediction capability. The trained data converged with relatively low error range (-0.96, 1.84) and prediction accuracy of 84.4% of the subject's satisfaction scores as given in Figure 5-19 and Figure 5-17, respectively.

The nonlinear complex equation between the satisfaction level and the study categories has been formulated. Using the connecting weights for inputs and output parameters and generated bias values obtained from the best performing ANN enable to express the satisfaction level as given in Equation 5-2. ANN model validation results values are inconsistent with the developed model. The R-value is found lower compared to the developed model, but the MSE value is higher than the developed model. It is suspected that this may be due to the small data sample which affects ANN learning algorithm potentiality in providing an accurate model.

The sensitivity analysis used to prioritize model input parameters contribution on the output suggests that the information quality and estimator performance have the greatest influence on participant's satisfaction level on the cost estimation with a positive influence on the output. The excitatory behavior of these inputs on the neural network supports the critical role of invisible factors, especially the quality of human interference activities on determining the

construction cost estimate. Similarly, project characteristics have a positive influence on the cost estimation process and placed in third ranking among the research categories variables.

Meanwhile, the sensitivity analysis indicated an inhibitory effect for both the contractual procedure and external factors on satisfaction level. This finding reflects the ranking of these categories too, as their ranks were the least of the research categories variables.

The present findings support the theoretical conceptual framework suggested by the researcher in this study as given by in Figure 3-1. It provides evidence of intangible factors impact on defining the construction cost estimate. Despite the visible and granted influence of external factors and project characteristics on the determination of construction cost, it can be argued that the quality of activities and information will either increase or decrease the occurrence of cost deviation from the allocated budget. Therefore, it is highly recommended that stakeholders provide extra attention to invisible factors related to the quality of resources (i.e.: estimator performance, information quality) used during the construction cost estimation. Their inherent influence either will work toward predicting the actual construction cost and granting the management satisfaction on project financial performance or it will cause the opposite scenario.

6.5 Chapter Summary

The chapter summarized the findings related to the following research objectives: level of cost estimation accuracy in MYS, construction cost prediction, factors affecting the cost estimation process in MYS, and modeling these factors. Generally, most findings related to significant factors affecting cost estimation follows the existing literature. The modeling results of factors

affecting the cost estimation have been discussed and aligned with the research third objective.

CHAPTER 7: Conclusions, Recommendations and Further Studies

7.1 Conclusions

7.2 Recommendations

7.3 Limitation of the Study

7.4 Further Recommended Studies

CHAPTER 7: Conclusions, Recommendations and Further Studies

7.1 Conclusions

The research aimed to demonstrate the level of cost estimation accuracy in MYS, investigate the factors affecting the cost estimation process to develop a model that illustrates the influenceable categories that can be controlled to improve the construction cost estimation process at MYS.

Although the overall level of cost deviation in MYS construction project is low, the disparity of project costs deviation suggests MYS suffers from a deficiency in determination of construction cost estimation.

The results of factors affecting the construction cost estimation promoted management contribution in setting up project cost which disturbs the construction cost estimation process and can cause cost deviations from the actual cost.

One promising finding is confirming the relationships between the research categories: availability of cost estimation unit, information quality, estimator performance, project characteristics, external factors and contractual procedure on the construction cost estimation accuracy.

The research objectives have been achieved as a cost model has been derived from the actual data to predict the estimation cost of MYS construction projects. The neural network model results support the conceptual framework proposed by the researcher. The effect of invisible activities on cost estimation accuracy: information quality and estimator performance have been recognized as first and second factors affecting the satisfaction level of participants' construction cost estimation.

7.2 Recommendations

The research findings suggest several courses of action that MYS can adopt to improve its construction cost estimate. First, it is highly recommended to reduce management interference in the construction cost estimation process. Second, using the information obtained in this research, it is advisable that MYS studies the causes of each project cost deviation and work toward establishing a target cost estimation accuracy range for future construction projects. MYS must overcome the deficiency of construction cost estimation in the planning phase through the implementation of recent design software instead of traditional construction cost estimation. The construction cost model developed can be used as a guideline in preliminary planning/design stage, but it is highly recommended to improve it through introducing other variables not limited to area size and level of quality.

The significant relationships found between the research categories and cost estimation accuracy suggest a number of important implications for future practices such as: develop complete construction cost database, enhance design team cost estimation skills, provide sufficient time for tender document preparation, confirm the clarity of contract drawings and tender document, document the progressive construction cost estimation, and specify a percentage from the contract amount in the cost estimate to encounter the uncertainty associated with the external factors influence.

7.3 Limitation of the Study

The results of this research are mainly derived for MYS as a case study. Limiting the research with MYS embedded constraining the data collection solely from MYS, this includes the cost data related to construction projects executed between (2013-2019) and limiting the survey participant to the stakeholders worked in MYS construction projects between (2013-2019). The researcher was unable to approach all the main contractors listed due to some companies being

bankrupted. As a result, a limited study sample size was obtained to perform the survey. Another limitation is the survey design, limiting the scope quality and organization management constructs to certain participants affect the model identification and its analysis using PLS-SEM. This constrain hinders the researcher from integrating the study variables in a comprehensive model that explains the relationships between them.

7.4 Further Recommended Studies

Based on the research limitations, the future studies recommended are:

- 1- To benchmark the degree of construction cost estimation accuracy in different construction disciplines (i.e., infrastructure projects like roads construction projects, sanitary construction projects, and mega projects).
- 2- To refine the research survey design by introducing a construct that measures the management influence on construction cost estimation and suits all the survey participants. This will integrate the interactional influence of management with other factors.
- 3- To validate the research findings regarding the factors affecting construction cost estimation in another construction directorate in Kingdom of Bahrain including a larger sample size.
- 4- To improve the construction cost model through introducing other independent variables such as (project characteristics, gross area, structural frame, project duration, quality of finishes, ..., etc.) to quantify a correction factors that helps in deriving better project construction cost prediction.
- 5- To explore improvement in construction cost estimation through adoption of new technologies such as: BC, Internet Of Thing (IOT), Building Information Management (BIM), google trends, and knowledge management.

References:

1. Abbasi, G. Y., and Al-Mharmah, H. (2000). Project management practice by the public sector in a developing country. *International Journal of Project Management*, 18(2), 105–109. reverived from [https://doi.org/10.1016/S0263-7863\(98\)00074-X](https://doi.org/10.1016/S0263-7863(98)00074-X)
2. Abd El-Karim, M. S. B. A., Mosa El Nawawy, O. A., and Abdel-Alim, A. M. (2017). Identification and assessment of risk factors affecting construction projects. *HBRC Journal*, 13(2), 202–216. reverived from <https://doi.org/10.1016/j.hbrcj.2015.05.001>
3. AbouRizk, S. M., Babey, G. M., and Karumanasseri, G. (2002). Estimating the cost of capital projects: an empirical study of accuracy levels for municipal government projects. *Canadian Journal of Civil Engineering*, 29(5), 653-661.
4. Abusafiya, H. A. M., and Suliman, S. M. A. (2017). Causes and Effects of Cost Overrun on Construction Project in Bahrain: Part I (Ranking of Cost Overrun Factors and Risk Mapping). *Modern Applied Science*, 11(7), 20. reverived from <https://doi.org/10.5539/mas.v11n7p20>
5. AACE International Recommended Practice No. 17R-97 (2003) COST ESTIMATE CLASSIFICATION SYSTEM TCM Framework: 7.3 – Cost Estimating and Budgeting , retrieved from <http://www.anvari.net/Risk%20Analysis/17r-97.pdf>
6. AACE International Recommended practice 10S-90 (10 october 2019) reverived from <http://www.aacei.org/docs/default-source/rps/10s-90.pdf?sfvrsn=42>
7. Adafin, J., Rotimi, J. O. B., Wilkinson, S., and Windapo, A. O. (2017). *Achieving Design-Stage Elemental Cost Planning Accuracy : Case Study of New Zealand. In International Conference on Civil and Construction Engineering (ICCCE).*

8. Aibinu, A. A., and Pasco, T. (2008). The accuracy of pre-tender building cost estimates in Australia. *Construction Management and Economics*, 26(12), 1257–1269. Retrieved from <https://doi.org/10.1080/01446190802527514>
9. Akadiri, O. P. (2011). Development of a multi-criteria approach for the selection of sustainable materials for building projects. *PhD Thesis - University of Wolverhampton*, 1–437. Retrieved from [http://wlv.openrepository.com/wlv/bitstream/2436/129918/1/Akadiri_PhD thesis.pdf](http://wlv.openrepository.com/wlv/bitstream/2436/129918/1/Akadiri_PhD%20thesis.pdf)
10. Akanni, P. O., Oke, A. E., and Akpomiemie, O. A. (2015). Impact of environmental factors on building project performance in Delta State, Nigeria. *HBRC Journal*, 11(1), 91–97. Retrieved from <https://doi.org/10.1016/j.hbrcj.2014.02.010>
11. Akhund, M. A., Raza Khoso, A., Shahzaib Khan, J., Usama Imad, H., and Muhammad Memon, K. (2019). Prompting cost overrun factors during PCP in construction projects. *Indian Journal of Science and Technology*, 12(4), 1–7. Retrieved from <https://doi.org/10.17485/ijst/2019/v12i4/140936>
12. Akintoye, A. (2010). Analysis of factors influencing project cost estimating practice *Construction Management & Economics*, 18(1), 77-89. Retrieved from <https://doi.org/10.1080/014461900370979>
13. Akintoye, A., and Fitzgerald, E. (2000). A survey of current cost estimating practices in the UK. *Construction Management and Economics*. Retrieved from <https://doi.org/10.1080/014461900370799>
14. Al bubshait, A., and ALjuwairah, Y. (2002). Factors contributing to construction costs in Saudi Arabia. *Cost Engineering, Cost engineering*, 44(5), 30.
15. Alharthi, H. (2019). Predicting physicians' satisfaction with electronic medical records using artificial neural network modeling. *Saudi Journal*

for *Health Sciences*, 8(2), 105. Retrieved from

https://doi.org/10.4103/sjhs.sjhs_14_19.

16. Al-khalidi, Z. (1990). Factors Affecting the Accuracy of Construction Costs Estimating in Saudi Arabia. (*Doctoral dissertation, King Fahd University of Petroleum and Minerals*).
17. Ali, Z., Zhu, F., and Hussain, S. (2018). Identification and assessment of uncertainty factors that influence the transaction cost in public sector construction projects in Pakistan. *Buildings*, 8(11). Retrieved from <https://doi.org/10.3390/buildings8110157>
18. Alkooheji, H. A., and Suliman, S. M. (2017) “Root Causes of Cost Deviation in Highway Construction Projects in Bahrain” *International Journal of Transport, Highways & Railways Engineering Vol.3 : Issue:1*, 36 – 53 Retrieved from <https://www.seekdl.org/journals/paper/details/8872>
19. Al-Sabah, R., and Refaat, O. (2019). Assessment of construction risks in public projects located in the state of Kuwait. *Journal of Engineering Research (Kuwait)*, 7(3), 13–32.
20. Alumbugu, P. O., Saidu, W. A. O., Abdullahi, I., Muhammed, M., and Abdulazeez, A. (2016). Assessment of the Factors Affecting Accuracy of Pre-tender Cost Estimate in Assessment of the Factors Affecting Accuracy of Pre-tender Cost Estimate in Kaduna State , Nigeria. *IOSR Journal of Environmental Science, Toxicology and Food Technology (IOSRJESTFT)*, 8(5), 19-27 Retrieved from <https://doi.org/10.9790/2402-08541927>
21. Amusan, L. M., Afolabi, A., Ojelabi, R., Omuh, I., and Okagbue, H. I. (2018). Data exploration on factors that influences construction cost and time performance on construction project sites. *Data in Brief*, 17, 1320–1325. Retrieved from <https://doi.org/10.1016/j.dib.2018.02.035>
22. Annamalaisami, C. D., and Kuppuswamy, A. (2019). Reckoning

- construction cost overruns in building projects through methodological consequences. *International Journal of Construction Management*, 1–11.
 Reverived from <https://doi.org/10.1080/15623599.2019.1683689>
23. Arif, F., Lodi, S. H., and Azhar, N. (2015). Factors influencing accuracy of construction project cost estimates in Pakistan : Perception and reality. *International Journal of Construction Management*, 15(1), 59–70.
 Reverived from <https://doi.org/10.1080/15623599.2015.1012141>
 24. Ashuri, B. and Lu, J. (2010). Time series analysis of ENR construction cost index. *Journal of Construction Engineering and Management*, 136(11), 1227-1237.
 25. Ashuri, B., Shahandashti, S. M., and Lu, J. (2012). Empirical tests for identifying leading indicators of ENR construction cost index. *Construction Management and Economics*, 30(11), 917-927.
 26. Ashworth, A., and Skitmore, R. M. (1982). Accuracy in estimating. *London: Chartered Institute of Building*.
 27. ASTM international. Section A.21 Precision and Bias. Retrieved from
 Reverived from <https://www.astm.org/ILS/precisionbias.html>
 28. Azrai, M., and Ismail, S. (2012). The accuracy of preliminary cost estimates in Public Works Department (PWD) of Peninsular Malaysia. *International Journal of Project Management*, 31(7), 994-1005
 Reverived from <https://doi.org/10.1016/j.ijproman.2012.11.008>
 29. Badawy, M. (2016). Critical factors affecting cost variation in residential construction projects in Egypt.
 30. Bakr, G. A. (2019). Identifying crucial factors affecting accuracy of cost estimates at the tendering phase of public construction projects in Jordan. *International Journal of Civil Engineering and Technology*, 10(1), 1335–1348.
 31. Bley, A. F. S. (1990). Improved conceptual estimating performance using a knowledge-based approach.
 32. Chan, A. P. C., and Tam, C. M. (2006). Factors affecting the quality of

building projects in Hong Kong. *International Journal of Quality & Reliability Management*, 17(4-5), 423-442.

33. Chandanshive, V. B., and Kambekar, A. R. (2019). Estimation of Building Construction Cost Using Artificial Neural Networks. *Journal of Soft Computing in Civil Engineering*, 3(1), 91–107. Reverived from <https://doi.org/10.22115/SCCE.2019.173862.1098>
34. Cheng, Y. (2014). An exploration into cost-in fl uencing factors on construction projects. *International Journal of Project Management* 32(5), 850–860.
35. Cho, K. M., Hong, T. H., and Hyun, C. T. (2009). Effect of project characteristics on project performance in construction projects based on structural equation model. *Expert Systems with Applications*, 36(7), 10461–10470. Reverived from <https://doi.org/10.1016/j.eswa.2009.01.032>
36. Cho, H. G., Kim, K. G., Kim, J. Y., and Kim, G. H. (2013). A comparison of construction cost estimation using multiple regression analysis and neural network in elementary school project. *Journal of the Korea Institute of Building Construction*, 13(1), 66-74.
37. Coffie, G. H., Aigbavboa, C. O., and Thwala, W. D. (2019). Modelling construction completion cost in Ghana public sector building projects. *Asian Journal of Civil Engineering*, 20(7), 1063-1070. Reverived from <https://doi.org/10.1007/s42107-019-00165-7>
38. Council, N. research. (1990). *Improving the Accuracy of Early Cost Estimates for Federal Construction Projects*. Reverived from <https://doi.org/10.17226/1693>
39. Davison, A. C., and Hinkley, D. V. (1997). Bootstrap Methods and their Application. In *Bootstrap Methods and their Application*. Reverived from <https://doi.org/10.1017/cbo9780511802843>
40. de Oña, J., de Oña, R., and Garrido, C. (2017). Extraction of attribute

importance from satisfaction surveys with data mining techniques: a comparison between neural networks and decision trees. *Transportation Letters*, 9(1), 39–48. Retrieved from <https://doi.org/10.1080/19427867.2015.1136917>

41. DeVellis, R. F. (2016). Scale development: Theory and applications Retrieved from <https://books.google.com.bh/books?id=9yfYCwAAQBAJ&lpg=PP1&ots=HoJBVkJWaik&lr&pg=PA145#v=onepage&q&f=false>
42. Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., and Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: a predictive validity perspective. *Journal of the Academy of Marketing Science*, 40(3), 434-449.
43. Dysert, L. (2007). Is “Estimate Accuracy ” an Oxymoron ? *Cost Engineering*, 49(1), 32.
44. Efron, B., and Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, 1(1), 54–75. Retrieved from <https://doi.org/10.1214/ss/1177013815>
45. Elhag, T. M. S., Boussabaine, A., and Ballal, T. (2005). Critical determinants of construction tendering costs : Quantity surveyor’s standpoint. *International journal of project management*, 23(7), 538–545. Retrieved from <https://doi.org/10.1016/j.ijproman.2005.04.002>
46. Elmousalami, H. H. (2019). Comparison of Artificial Intelligence Techniques for Project Conceptual Cost Prediction.*arXiv preprint arXiv:1909.11637*.
47. Enshassi, A., Mohamed, S., and Abdel-hadi, M. (2013). Factors affecting the accuracy of pre-tender estimates in the gaza strip. 18(1), 73–94.
48. Famiyeh, S., Amoatey, C. T., Adaku, E., and Agbenohevi, C. S. (2017). Major causes of construction time and cost overruns: A case of selected educational sector projects in Ghana. *Journal of Engineering, Design and*

Technology, 15(2), 181–198. Retrieved from

<https://doi.org/10.1108/JEDT-11-2015-0075>

49. Flanagan, R., and Norman, G. (1983). The accuracy and monitoring of quantity surveyors price forecasting for building work. *Construction Management and Economics*, Vol. 1, pp. 157–180. Retrieved from <https://doi.org/10.1080/01446198300000012>
50. Flyvbjerg, B. (2008). Curbing optimism bias and strategic misrepresentation in planning: Reference class forecasting in practice. *European Planning Studies*. Retrieved from <https://doi.org/10.1080/09654310701747936>
51. Flyvbjerg, B., and Holm, M. S. (2002). Works Projects. *Journal of the American Planning Association*, 68(3), 279–295. Retrieved from <https://doi.org/10.1080/01944360208976273>
52. Flyvbjerg, B. (Ed.). (2017). The Oxford handbook of megaproject management. *Oxford University Press*.
53. Flyvbjerg, B., Holm, M. S., and Buhl, S. (2002). Underestimating costs in public works projects: Error or lie?. *Journal of the American planning association*, 68(3), 279-295.
54. Gunduz, M., and Sahin, H. B. (2015). An early cost estimation model for hydroelectric power plant projects using neural networks and multiple regression analysis. *Journal of Civil Engineering and Management*, 21(4), 470–477. Retrieved from <https://doi.org/10.3846/13923730.2014.890657>
55. Goh, A. T. C., Kulhawy, F. H., and Chua, C. G. (2005). Bayesian neural network analysis of undrained side resistance of drilled shafts. *Journal of Geotechnical and Geoenvironmental Engineering*, 131(1), 84–93. Retrieved from [https://doi.org/10.1061/\(ASCE\)1090-0241\(2005\)131:1\(84\)](https://doi.org/10.1061/(ASCE)1090-0241(2005)131:1(84))
56. Gunner, J., and Skitmore, M. (1999). Comparative analysis of building

price forecasting errors: a sample of singapore data. *Construction Management and Economics*, 17(5), 635–646.

57. Gzik, M., Tkacz, E., Paszenda, Z., and Piętka, E. (Eds.). (2017). Innovations in biomedical engineering. *Springer International Publishing*. Retrieved from shorturl.at/txMS1
58. Hair, J. F., Ringle, C. M., and Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing theory and Practice*, 19(2), 139-152.
59. Hair Jr, J. F., Hult, G. T. M., Ringle, C., and Sarstedt, M. (2013) A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM) Retrieved from <https://books.google.com.bh/books?id=IFiarYXE1PoC&lpg=PP1&dq=pls%20sem&pg=PP1#v=onepage&q=pls%20sem&f=false>
60. Hair, J. F., Risher, J. J., Sarstedt, M., and Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24. Reverived from <https://doi.org/10.1108/EBR-11-2018-0203>.
61. Hatamleh, M. T., Hiyassat, M., Sweis, G., and Sweis, R. J. (2018). Factors affecting the accuracy of cost estimate : Case of Jordan Engineering , *Construction and Architectural Management Article information* : (January). Reverived from <https://doi.org/10.1108/ECAM-10-2016-0232>
62. Hecht-Nielsen, R.(1989). Theory of the backpropagation neural network. *International Joint Conference on Neural Networks (Vol. 2, pp. 593-605)*.Reverived from <https://doi.org/10.1109/ijcnn.1989.118638>
63. Heravi, G., and Mohammadian, M. (2017). Cost Overruns and Delay in Municipal Construction Projects in Developing Countries. 1(1), 31–38. Reverived from <https://doi.org/10.22060/ceej.2017.12189.5163>
64. Islam, R., Musabbir, S. R., Ahmed, I. U., Hadiuzzaman, M., Hasnat, M., and Hossain, S. (2016). Bus service quality prediction and attribute ranking using probabilistic neural network and adaptive neuro fuzzy inference

- system. *Canadian Journal of Civil Engineering*, 43(9), 822–829. Reverived from <https://doi.org/10.1139/cjce-2016-0119>
65. Jeffreys, H. (1961). *Theory of Probability*. Oxford: Oxford University Press.
 66. Juszczuk, M., Leśniak, A., and Zima, K. (2018). ANN Based Approach for Estimation of Construction Costs of Sports Fields. *Complexity*, 2018. Reverived from <https://doi.org/10.1155/2018/7952434>
 67. Kadiri, D. S. (2014). An assessment of current preliminary cost estimating practice in Nigeria. *Journal of Environmental Design and Management*, 6(1&2), 97-111.
 68. Kerzner, H. (2017). *Project management: a systems approach to planning, scheduling, and controlling*. John Wiley & Sons. Retrieved from shorturl.at/djuz4
 69. Khan, A. N., and Ali, A. (2018). Factors Affecting Retailer's Adopti on of Mobile Payment Systems: A SEM-Neural Network Modeling Approach. *Wireless Personal Communications*, 103(3), 2529–2551. Reverived from <https://doi.org/10.1007/s11277-018-5945-5>
 70. Kock, N. (2010). Using WarpPLS in e-collaboration studies: An overview of five main analysis steps. *International Journal of e-Collaboration (IJeC)*, 6(4), 1-11.
 71. Kock, N. (2014). Single missing data imputation in PLS-SEM. (December), 20. Reverived from <https://doi.org/10.13140/2.1.4967.8409>
 72. Kock, N. (2015). A note on how to conduct a factor-based PLS-SEM analysis. *International Journal of E-Collaboration*, 11(3), 1–9. Reverived from <https://doi.org/10.4018/ijec.2015070101>
 73. Kock, N. (2016). Hypothesis testing with confidence intervals and P values in PLS-SEM. *International Journal of E-Collaboration*, 12(3), 1–6. Reverived from <https://doi.org/10.4018/IJeC.2016070101>

74. Kock, N. (2017). WarpPLS user manual: Version 6.0. *ScriptWarp Systems: Laredo, TX, USA*.
75. Koushki, P. A., Al-Rashid, K., and Kartam, N. (2005). Delays and cost increases in the construction of private residential projects in Kuwait. *Construction Management and Economics*, 23(3), 285–294. Reverived from <https://doi.org/10.1080/0144619042000326710>
76. Kroonenberg, P. M., and Lohmoller, J.-B. (1990). Latent Variable Path Modeling with Partial Least Squares. *Journal of the American Statistical Association*. Reverived from <https://doi.org/10.2307/2290049>
77. Latan, H., and Noonan, R. (2017). Partial least squares path modeling: Basic concepts, methodological issues and applications. In *Partial Least Squares Path Modeling: Basic Concepts, Methodological Issues and Applications*. Reverived from <https://doi.org/10.1007/978-3-319-64069-3>
78. Leśniak, A., and Zima, K. (2018). Cost calculation of construction projects including sustainability factors using the Case Based Reasoning (CBR) method. *Sustainability (Switzerland)*, 10(5). Reverived from <https://doi.org/10.3390/su10051608>
79. Ling, Y. Y., Hwee, J., Boo, S., Ling, Y. Y., Hwee, J., and Boo, S. (2010). Improving the accuracy estimates of building of approximate projects Improving the accuracy of approximate. *Building Research & Information*, 29(4), 312-318. Reverived from <https://doi.org/10.1080/09613210122440>
80. Loy, J. (2019). Neural Network Projects with Python: The ultimate guide to using Python to explore the true power of neural networks through six projects. *Packt Publishing Ltd*. Retrieved from <https://rb.gy/pe2tif>.
81. Lu, W., Hua, Y. Y., and Zhang, S. J. (2017). Logistic regression analysis for factors influencing cost performance of design-bid-build and design-build projects. *Engineering, Construction and Architectural Management*, 24(1), 118–132. Reverived from <https://doi.org/10.1108/ECAM-07-2015->

82. Mahamid, I. (2018). Critical Determinants of Public Construction Tendering Costs. *International Journal of Architecture, Engineering and Construction*, 7(1), 34–42. Reverived from <https://doi.org/10.7492/ijaec.2018.005>
83. Mahamid, I., Al-Ghonamy, A., and Aichouni, M. (2014). Factors affecting accuracy of pretender cost estimate: Studies of Saudi Arabia. *International Journal of Applied Engineering Research*, 9(1), 21–36.
84. Marcoulides, G. A., and Saunders, C. (2006). Editor's comments: PLS: a silver bullet?. *MIS quarterly*, iii-ix.
85. Morrison, N., and Stevens, S. (1981). Cost planning and computers. *Report for the Property Services Agency*.
86. McCaffer, R. (1976). Contractors' bidding behaviour and tender price prediction. *A Doctoral Thesis. Submitted in Partial Fulfillment of the Requirements for the Award of Doctor of Philosophy of Loughborough University.*, 26(3), 348–358. Reverived from <https://doi.org/10.3390/buildings5020405>
87. Mislick, G. K., and Nussbaum, D. A. (2015). Cost estimation: methods and tools. *John Wiley & Sons*.
88. Monecke, A., and Leisch, F. (2012). SemPLS: Structural equation modeling using partial least squares. *Journal of Statistical Software*, 48, 1–32. Reverived from <https://doi.org/10.18637/jss.v048.i03>
89. Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). Introduction to linear regression analysis (Vol. 821). *John Wiley & Sons*.
90. Moon, T. and Shin, D. H. (2018). Forecasting construction cost index using interrupted time-series. *KSCE Journal of Civil Engineering*, 22(5), 1626-1633.

91. Moon, T. and Shin, D. H. (2018). Forecasting model of construction cost index based on VECM with search query. *KSCE journal of civil engineering*, 22(8), 2726-2734.
92. Muhammad T. Hatamleh, Mohammed Hiyassat, Ghaleb Jalil Sweis, and Rateb Jalil Sweis, (2018) "Factors affecting the accuracy of cost estimate: case of Jordan", *Engineering, Construction and Architectural Management*, Vol. 25 Issue: 1, pp.113-131, Retrieved from <https://doi.org/10.1108/ECAM-10-2016-0232>
93. Mulla, S. S. and Waghmare, A. P. (2015). A Study of Factors Caused for Time & Cost Overruns in Construction Project & their Remedial Measures. *Journal of Engineering Research and Applications Wwww.Ijera.Com ISSN*, 5(16), 2248–962248. Retrieved from Retrieved from www.ijera.com
94. National Research Council. (1990). Improving the accuracy of early cost estimates for federal construction projects. *National Academies Press*.
95. Newton, S. (2012). *Improving a Cost Estimate using Reference Class Analytics*. 73–83.
96. Offei-Nyako, K., Ohene Tham, L. C., Bediako, M., Adobor, C. Dela, and Oduro Asamoah, R. (2016). Deviations between Contract Sums and Final Accounts: The Case of Capital Projects in Ghana. *Journal of Construction Engineering*, 2016, 1–8. Retrieved from <https://doi.org/10.1155/2016/2814126>
97. Ogunlana, S. O. (1989). Accuracy in design cost estimating. (*Doctoral dissertation, Loughborough University*).
98. Oke, A. E. (2017). Appraisal of Qualitative Factors Affecting Contractors Tender Price in a Developing Country. (July).
99. Oladokun, M. G., Oladokun, A. A., and Odesola, I. A. (2011). Accuracy of Pre-Tender Cost Estimates of Consultant Quantity Surveyors in Nigeria. *Science*, 1(1).
100. Olden, J. D., Joy, M. K., and Death, R. G. (2004). An accurate

comparison of methods for quantifying variable importance in artificial neural networks using simulated data. *Ecological Modelling*, 178(3–4), 389–397. Retrieved from

<https://doi.org/10.1016/j.ecolmodel.2004.03.013>

101. Pillai, M. P. (2008). Validation — success key in lump-sum projects: contractor's perspective. *Paper presented at PMI® Global Congress 2008—Asia Pacific, Sydney, New South Wales, Australia. Newtown Square, PA: Project Management Institute.*
102. Pires dos Santos, R., Dean, D. L., Weaver, J. M., and Hovanski, Y. (2019). Identifying the relative importance of predictive variables in artificial neural networks based on data produced through a discrete event simulation of a manufacturing environment. *International Journal of Modelling and Simulation*, 39(4), 234-245.
103. PMI A Guide to the Project Management Body of Knowledge (PMBOK® Guide); *fifth edition*(2013).
104. PMI Lexicon of Project Management Terms Version 3.2 (2017) <https://www.pmi.org/pmbok-guide-standards/lexicon>
105. Reza, M. C. (2020). Blockchain Technology in the Construction Industry. California Polytechnic State University. <https://digitalcommons.calpoly.edu/cmisp/410/>
106. Rivera, A., Le, N., Kapsikar, K., Kashiwagi, J., and Yasir, A. (2017). Identifying the Global Performance of the Construction Industry. *53rd ASC Annual International Conference Proceedings*, 567–575.
107. Ross, J., and Staw, B. M. (1993). Organizational escalation and exit: Lessons from the Shoreham nuclear power plant. *Academy of Management journal*, 36(4), 701-732.
108. Saidu, I., and Shakantu, W. (2017). An investigation into cost overruns for ongoing building projects in Abuja, Nigeria. *Acta Structilia*, 24(1), 53–72. Retrieved from <https://doi.org/10.18820/24150487/as24i1.3>

109. Sarstedt, M., and Cheah, J. H. (2019). Partial least squares structural equation modeling using SmartPLS: a software review. *Journal of Marketing Analytics*, 7(3), 196–202. Reverived from <https://doi.org/10.1057/s41270-019-00058-3>
110. Sarstedt, M., Diamantopoulos, A., and Salzberger, T. (2016). Should we use single items? Better not. *Journal of Business Research*, 69(8), 3199–3203. Reverived from <https://doi.org/10.1016/j.jbusres.2016.02.040>
111. Sarstedt, M., Ringle, C. M., and Hair, J. F. (2019). Partial Least Squares Structural Equation Modeling. In *Handbook of Market Research*. Reverived from <https://doi.org/10.1007/978-3-319-05542-8>.
112. Schottlander, D. E. (2006). How Accurate Are Your Estimates ? *AACE International Transactions*, ES171.
113. Senouci, A., Ismail, A., and Eldin, N. (2016). Time Delay and Cost Overrun in Qatari Public Construction Projects. *Procedia Engineering*, 164(June), 368–375. Reverived from <https://doi.org/10.1016/j.proeng.2016.11.632>.
114. Sharma, S. K., Gaur, A., Saddikuti, V., and Rastogi, A. (2017). Structural equation model (SEM)-neural network (NN) model for predicting quality determinants of e-learning management systems. *Behaviour and Information Technology*, 36(10), 1053–1066. Reverived from <https://doi.org/10.1080/0144929X.2017.1340973>.
115. Sharma, P. N., and Kim, K. H. (2013). A Comparison of PLS and ML Bootstrapping Techniques in SEM: A Monte Carlo Study. *Springer Proceedings in Mathematics and Statistics*, 56, 201–208. Reverived from https://doi.org/10.1007/978-1-4614-8283-3_13.
116. Shih, T. C., and Chang, H. T. (2015). Establishment of a evaluation system for photovoltaic power generation using neural network. *Neuron*, 22(11), 1.
117. Skitmore, M. (1985). The influence of professional expertise in

construction price forecasts. *The University of Salford*.

118. Skitmore, M. (1991). Early stage construction price forecasting a review of performance. *RICS Books*.
119. Skitmore, M., and Ashworth, A. (1983). Accuracy in estimating. *London: Chartered Institute of Building*.
120. Skitmore, M., and Gunner, J. (2000). Comparative Analysis of Pre-Bid Forecasting of Building Prices Based on Singapore Data. *Construction Management and Economics*, 17(5), 635-646.
121. Skitmore, m., and picken, d. (2000). The accuracy of pre-tender building price forecasts: an analysis of USA data. *Institute of Quantity Surveyors Refereed Journal*, 4(1), 33–39.
122. Skitmore, M., and NG, T. (2000). *The variance of elemental estimates*.
123. Sridarran, P., Keraminiyage, K., and Herszon, L. (2017). Improving the cost estimates of complex projects in the project-based industries. *Built Environment Project and Asset Management*, 7(2), 173–184. Reverived from <https://doi.org/10.1108/BEPAM-10-2016-0050>.
124. Staw, B. M., and Ross, J. (1987). Behavior in Escalation Situations: Antecedents, Prototypes and Solutions. *Research in Organizational Behavior*, 9, 39–78.
125. Tatiya, A., Zhao, D., Syal, M., Berghorn, G. H., and LaMore, R. (2018). Cost prediction model for building deconstruction in urban areas. *Journal of Cleaner Production*, 195, 1572–1580. Reverived from <https://doi.org/10.1016/j.jclepro.2017.08.084>.
126. Trost, S. M., Asce, M., Oberlender, G. D., and Asce, F. (2003). Predicting Accuracy of Early Cost Estimates Using Factor Analysis and Multivariate Regression. *Journal of construction Engineering and Management*, 129(2), 198-204.
127. Weisberg, S. (2005). Applied linear regression (Vol. 528). *John Wiley & Sons*.

128. Willaby, H. W., Costa, D. S. J., Burns, B. D., MacCann, C., and Roberts, R. D. (2015). Testing complex models with small sample sizes: A historical overview and empirical demonstration of what Partial Least Squares (PLS) can offer differential psychology. *Personality and Individual Differences*, 84, 73–78. Retrieved from <https://doi.org/10.1016/j.paid.2014.09.008>.
129. Wong, K. K. K. (2019). Mastering partial least squares structural equation modeling (PLS-SEM) with Smartpls in 38 hours. Retrieved from shorturl.at/lnLO8
130. Xia, B., Xiong, B., Skitmore, M., Wu, P., and Hu, F. (2016). Investigating the Impact of Project Definition Clarity on Project Performance: Structural Equation Modeling Study. *Journal of Management in Engineering*, 32(1), 04015022. Retrieved from [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000386](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000386).
131. Zeinalizadeh, N., Shojaie, A. A., and Shariatmadari, M. (2015). Modeling and analysis of bank customer satisfaction using neural networks approach. *International Journal of Bank Marketing*, 33(6), 717–732. Retrieved from <https://doi.org/10.1108/IJBM-06-2014-0070>.
132. Zuna, H. T., Hadiwardoyo, S. P., and Rahadian, H. (2016). Developing a model of toll road service quality using an artificial neural network approach. *International Journal of Technology*, 7(4), 562–570. Retrieved from <https://doi.org/10.14716/ijtech.v7i4.2612>

Appendices

Appendix 1: Summary of top and lowest ranked factors affecting construction cost estimation from recent publications

Appendix 2: Quantitative measures and methodologies used to evaluate projects cost estimation accuracy

Appendix 3: Interpretation of Bayes Factors as evidence for null hypothesis (H0) and alternative hypotheses (H1)

Appendix 4: Mathematical expressions for PLS-SEM quality criteria

Appendix 5: Questionnaire survey used to measure factors affecting construction cost estimation

Appendix 6: Descriptive data analysis for cost deviation in MYS construction projects for the period between (2012-2019)

Appendix 7: SPSS and Microsoft Excel regression model development

Appendix 8: Questionnaire Survey - RII calculation

Appendix 9: ANN model results and mathematical equation development

Appendix 10: Categories ranking as per normalized connection weight method

Appendix 1:

**Summary of top and lowest ranked factors affecting
construction cost estimation from recent publications**

	Authors	Top ranked	Lowest Ranked
1	(Alumbugu, Saidu, Abdullahi, Muhammed, & Abdulazeez, 2016)	Experience and skill level of the consultants, Project teams experience on the construction type, clear and detail drawings and specification, completeness of cost information, accuracy and reliability of cost information.	Inflationary Pressures, Level of involvement of the project manager Level of competition (number of competitors), Time between project announcement, bid submission and award of contract, Content of specification and arbitration clause Social and cultural requirements
2	(Oke, 2017)	For project characteristics: buildability of the project, construction technique, project size, type of structure and the location of the project , Type of construction , Scale and scope of construction client characteristics: type of client, client requirement on quality and the financial ability of the client. contractors characteristics factors: experience on similar project, Tendering situation, consultant and design characteristics :incomplete design/incorrect design, Procurement method , Tendering method , Type of contract. external factors and market condition : labour cost, interest rate, inflation rate and market stability.	For project characteristics: Project duration, Assess to site and storage client characteristics: Certainty of the project brief , Deadline requirement contractors characteristics factors: management team and planning capability of the contractor. Past relationship with client, Current work load , Need for work. Tendering situation, consultant and design characteristics :, completeness of projects information External factors and market condition : Market stability , Level of competition and level of construction activities , number of builders the inaccuracy of cost estimating factors :Poor tender document, insufficient analysis of tender

			document, poor requirement stand and insufficient tendering time
3	(Amusan, Afolabi, Ojelabi, Omuh, & Okagbue, 2018)	Contractors Project inexperience, Inadequate planning, Inflation Incessant variation order, Change in project design, Project complexity Shortening of contract period	Material Price fluctuations , Site conflicts , Poor workmanship Inadequate financial provision, Contractors inefficiency Unsteady material supply Unpredictable weather condition, Breach of local regulation, Lack of executive capacity by employer
4	(Famiyeh, Amoatey, Adaku, & Agbenohevi, 2017).	Client-related cost factors Consultant related cost factors Contractor-related cost factors Material-related cost factors	Government-related cost factors Contractual-related cost factors External-related cost factors
5	(Samuel Famiyeh, Charles Teye Amoatey and Ebenezer Adaku, Collins Sena Agbenohevi,2016).	Client related cost factor: Financial difficulty by client , Delay in contract awards , Poor communications , Delay in decision making Consultant related cost factors Poor feasibility and project analysis , Inaccurate estimates Government-related cost factors Slow permits by government agencies Material-related cost factors	Government actions related delay factors: Political interference, Building Regulations Contractual-related delay factors Inappropriate type of contract used (Procurement process), Mistakes and discrepancies in contract documents , Inappropriate overall organizational structure linking all parties to the project External-related delay factors Changes in laws and regulations

		Material price fluctuations , Shortage of materials External-related cost factors Material price changes (inflation) , High cost of labour	Strikes , External work due to public agencies, Problems with neighbours
6	(Offei-Nyako, Ohene Tham, Bediako, Adobor, & Oduro Asamoah, 2016)	Price fluctuation , Late material delivery Changes in the scope of work , Changes in design	Bureaucracy in tendering method Economic stability
7	(Enshassi, Mohamed, & Abdel-hadi, 2013)	Factors related to consultants, design parameters and information top category : Material (prices/availability/supply/quality/imports) Project team's experience in the type of construction , Experience and skill level of the consultant , Having clear and detailed drawings and specifications , Quality of information flow, Completeness of cost information , Accuracy and reliability of cost information , Currency exchange fluctuation , Clear contract conditions	Factors related to project characteristics , Factors related to contract requirements and procurement methods Volume of consultant's workload during estimation , Content of disputes resolution methods clause , Weather effects , Type of project , government regulations requirement, project location, type of structures, size of estimating team and social and cultural impact.

8	(Cheng, 2014)	Clearly define the scope of project in the contract, Contract dispute (unclear drawings or guidelines/regulations) High fluctuation in commodity , The gap between the construction plan and the reality is too great, Material shortage or supply delay , Practical experience Modifications to the scope of construction The level of demand on quality Project team (coordination capability and the understanding of operational procedure)	Procurement contract , Geology, topography , Climate factor Natural disaster
9	(Mulla & Waghmare, 2015)	Inaccurate estimate of time and cost due to: Faulty designs , Land acquisition problem , Poor bidding , Irregular flow of finance , Delay in payment of work done , Deficiencies in management Delay in decision by Client/ Architect Lack of coordination between different parties involved , Change in scope of work	Not mentioned
10	(Lu, Hua, & Zhang, 2017)	For DBB project, owner ability plays a significant role in influencing cost performance,	while for DB projects, contractor ability is found to be critical in explaining cost performance.

11	(Mahamid, Al-Ghonamy, & Aichouni, 2014)	Level of competitors in the tendering Material price changes , Communications with suppliers , Communications with client , Estimating method used Labor productivity , Clear contract conditions , Estimator's experience Tender selection method , Client's evaluation and awarding policy	Location of project , Communications and coordination between designers Equipment conditions , Volume of consultant's workload during estimation Construction techniques requirements Timing of advertisement (weather effects) , Site conditions (topography, access,..etc) ,Equipment costs Material shortage , Type of project (residential, commercial, industrial,...etc) Lowest groups : Factors related to labors, material and equipments Factors related to project characteristics
12	(Arif, Lodi, & Azhar, 2015)	client's level of experience , the construction type, level of expertise of contractor, priority of project, initial scale of project, applicability of cost information, accuracy and reliability of cost information, completeness of tender documents, and labour productivity	Not mentioned
13	(Abd El-Karim, Mosa El Nawawy, & Abdel-Alim, 2017)	political situation , engineering and design contractor , project management and site	Environment, equipment, Labor Owner related factors Were among lowest ranked

		location financial related factors were among the highest groups	
14	(Badawy, 2016)	Amount of specialist work Scale & scope of construction Complexity of design Quality of design & specification Constructability of design Availability of qualified management staff Type of construction	Risk acc. To politic of country Working time (8 - 12) hours Working shifts (8 - 16 - 24) hours Level of uncertainty of soil condition Risk acc. To labor variation cost Tendering method Risk acc. To legal rule of country
15	(Abusafiya & Suliman, 2017)	Design and documentation Related Factors Contractor's site Management Related Factors Project Management and Contract Administration Related Factors	Financial Management Related Factors Non-Human Resource Related Factors Environmental factors
16	(Mahamid, 2018)	client's evaluation and awarding policy, labor skills, tender selection method (open, selected, negotiation, etc.), method of procurement (design bid build, design and build, project management, etc.), clear scope definition for the client, material price changes,	Timing of advertisement (weather effects), Project size , Material shortage Location of project , Labor availability Designer's experience level , Equipment availability , Equipment conditions , Equipment costs

		estimating method used, and time between project design and project execution. Top: Factors related to clients , Factors related to cost estimating	Lowest group : Factors related to consultants and design Factors related to project characteristics
17	(Hatamleh, Hiyassat, Sweis, & Sweis, 2018)	clear and detail drawings and specification, pricing experience of construction projects, perception of estimation importance, equipment (cost/availability/performance), project complexity, clear scope definition, accuracy and reliability of cost information, site, constraints (access, storage, services), material availability, financial capabilities of the client, and availability of database of bids on similar project (historical data).	The estimating method used, Level of involvement of the project Manager, Quality of information and flow requirements, Market conditions / economic climate , The padding estimates , Contractor bidding strategy
18	(Ali, Zhu, & Hussain, 2018).	Competitive tendering , Incomplete design and specifications , Late payments Conflict management , Delay possession of sites , Force majeure , Work acceleration , Impractical tender estimation , Owner negligence Contractor's characteristics	Inadequate investigation, Inclement weather conditions, Differing site conditions, Labour strike Contractor-subcontractor relationship Dispute on claim evaluation, Incentive payments, Unclear project scope , Failure to investigate contractor's bidding

19	(Bakr, 2019)	The level of experience of estimation team, Client's financial capabilities, experience of the project's team, time allowed for estimation, complexity and constructability of the project, and the availability and prices for labor and material. Top group :Client's characteristics related factor , Consultant related factors	Awarding policy and evaluation by the client Constraints at site (storage, access, services, etc.) Advertisement timing and effect of weather , Effect of government policies Pressure of inflation , Currency exchange rate changes , Impact of cultural and social matters , Resources and market conditions , Lowest: External and tendering requirements related factors
20	(Akhund, Raza Khoso, Shahzaib Khan, Usama Imad, & Muhammad Memon, 2019)	wrong or improper (poor/inappropriate) design, wrong estimation, increase in scope of work, poor resource management, change orders, exchange rate /fluctuation of prices ,construction methods ,material and labor wage escalation (inflation), financial instability in markets and poor professional construction management	Permits (urban planning bureau & order of engineers) and access facilities Unreasonable project time frame Uncooperative owners , Changes in laws and government regulations Inadequate definition of substantial completion , Improper selection of subsequent consultants , Political situation , Monopoly
21	(Al-Sabah & Refaat, 2019)	Price Inflation, Resources Restrictions , Resources, Quality , Public Authorities Regulation and Requirements , Permits and Licenses , Contract Harsh Clauses, and Inclement Weather .	stability risks: : Political Variability , Authority Performance (Override) , Safety and Crimes , Bribe and Corruption , Public Authorities Strikes.

		The top project risks were Performance of Subcontractors Performance of Preselected Subcontractors , Insufficient Project Duration, Subcontractors/Suppliers Coordination , Transmittal Process, and Major Changes	
22	(Annamalaisami & Kuppuswamy, 2019)	Design change , Change in specification Incomplete drawing , Change in the scope Practice in assigning the lowest bidder, Unpredictable weather condition , Increase in material price	Insurance problem , Poor technical problems , Inaccurate site investigation Lack of experience of project type Lack of experience of project location Fluctuation in money exchange rate Deficiencies in infrastructure Labor cost increase, Lack of constructability
23	(Sridarran, Keraminiyage, & Herszon, 2017).	Product and project size , Time frame , Organizational capability , Project management maturity level , Uncertainty , Budgetary constraints , Knowledge and experience are the high-impact complexity dimensions on cost estimation,	Economic uncertainty, Stakeholder interaction , Political influence (politics) ,External environment constraint , Project description , Pace/speed to market , Cultural resistance and differences , Environmental and safety impact . Impact on society
24	(Adafin, Rotimi, Wilkinson, & Windapo, 2017)	market conditions, inadequate tender documentation, incomplete design information and site investigation information.	Not mentioned

Appendix 2:

**Quantitative measures and methodologies used to evaluate
projects cost estimation accuracy**

Table: Previous quantitative measures and methodologies to evaluate projects cost estimation accuracy (Skitmore et al., 1991)

Quality Measures	Researcher	Indicator used	Statistics
Bias	(McCaffer,1976),(De Neufville et al.,1977)	Lowest bid/forecast ratio	Arithmetic mean
	Wilson et al.,1987		Median
	Runeson and Bennett (1983) Hanscomb Associates (1984)		Pearsons r
	McCaffer,1976		Spearman's rho
	Harvey (1979), Hanscomb (1984)		Coefficient of regression
Associates	Jupp and McMillan (1981)	Forecast/lowest bid ratio	Arithmetic mean
	Morrison and Stevens (1980) (Skitmore,1987) , Skitmore and Tan(1988)	Percentage forecast exceeds lowest bid	Arithmetic mean
	Tan (1988)		Pearsons r
	Flanagan and Norman (1983)	Lowest bid-estimate	Coefficient of regression

Consistency	McCaffer,1976	Lowest bid/forecast ratio	Standard deviation
	McCaffer,1976 , McCaffer et al., 1984 Ross (1983)		Coefficient of variation
	Jupp and McMillan (1981)	Forecast/lowest bid ratio	Standard deviation
	(Skitmore,1985) , Skitmore and Tan (1988)	Percentage forecast exceeds lowest bid	Standard deviation
	Morrison and Stevens (1980)		Coefficient of variation
	Skitmore and Tan (1988)	Percentage forecast exceeds forecast	Standard deviation
	Bennett (1987)	Number of 'serious' errors	Total
Accuracy	Morrison and Stevens(1980)	Modulus Percentage forecast exceeds lowest bids	Arithmetic mean
	Flanagan and Norman (1983)		Number observations in ranges
	Skitmore (1985) , Skitmore and Tan (1988)		Arithmetic mean and standard deviation
	Tan (1988)		Pearsons r

	Skitmore and Tan (1988)	Modulus Percentage lowest bid exceeds forecast	Arithmetic mean and standard deviation
	Tan (1988)		Pearsons r
	Skitmore (1985)	Squared Percentage forecasts exceeds lowest bids	Square root of Arithmetic mean

Appendix 3:

Interpretation of Bayes Factors as evidence for null hypothesis (H_0) and alternative hypotheses (H_1)

Table: Interpretation of Bayes Factors as evidence for null hypothesis (H₀) and alternative hypotheses (H₁)

Bayes factor	Interpretation
>100	Decisive evidence for H ₁
30-100	Very strong evidence for H ₁
10-30	Strong evidence for H ₁
3-10	Substantial evidence for H ₁
1-3	Anecdotal evidence for H ₁
1	No evidence
$1/3-1$	Anecdotal evidence for H ₀
$1/10-1/3$	Substantial evidence for H ₀
$1/30-1/10$	Strong evidence for H ₀
$1/100-1/30$	Very strong evidence for H ₀
$<1/100$	Decisive evidence for H ₀

Appendix 4:

Mathematical expressions for PLS-SEM quality criteria

Table: Mathematical expressions for PLS-SEM quality criteria

Model	quality criteria	Equation
Measurement model (Reflective)	Internal consistency: composite reliability	$\frac{N \cdot C_{av}}{v_{av} + (N-1) \cdot C_{av}}$ Where: • N = the number of items. • C_{av} = average covariance between item-pairs. • v_{av} = average variance.
	Convergent validity	$AVE = \frac{\sum_{i=1}^k \lambda_i^2}{\sum_{i=1}^k \lambda_i^2 + \sum_{i=1}^k Var(e_i)}$ Where, • K: is the number of items, • λ_i: the factor loading of item • $Var(e_i)$: the variance of the error of item .
	Discriminant validity	The Fornell–Larcker criterion and thus the requirements for distinctiveness between two latent variables X and Y are fully met if the AVE of X and Y are both higher than the variance that X and Y share with each other. $AVE_x > \phi_{xy}^2 \text{ and } AVE_y > \phi_{xy}^2$ where ϕ_{xy} is the squared correlation between X and Y.
The structural model	Coefficient of determination (explained variance)	$R = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2] [n \sum y^2 - (\sum y)^2]}}$ Where: X: values in first set of data Y: data in second set of data

	Collinearity assessment	$VIF_i = \frac{1}{1 - R_i^2}$ Where R_i^2 is the unadjusted coefficient of determination for regressing the ith independent variable on the remaining ones
	Effect size	$F^2 = \frac{R_{\text{included}}^2 - R_{\text{excluded}}^2}{1 - R_{\text{included}}^2}$
	Model predictive relevance (Stone-Geisser)	$Q^2 = 1 - \frac{\sum_D E_D}{\sum_D O_D}$ Where: E = The sum of squares of prediction error O = The sum of squares error using the mean for prediction D = Omission distance

Appendix 5:

**Questionnaire survey used to measure factors affecting
construction cost estimation**

UNIVERSITY OF BAHRAIN

College of Engineering



Department of Mechanical Engineering

Objective:

The objective of this questionnaire is to identify the factors affecting the construction cost estimation process. It is a part of Masters thesis in University of Bahrain (UOB).

Design:

This questionnaire contains three sections which will be considered in return:

Sections	Subject
Section (1): Part A	General information
Section (2): Part B	Factors affecting cost estimation

Guidance:

As client / consultant : you are requested to answer scope quality category with other categories (Estimator performance, Information quality , Project characteristics External factors, Contractual procedure)

As Main contractors : you are requested answer the contractor organization status with other categories (Estimator performance, Information quality , Project characteristics , External factors, Contractual procedure) .

We will be most grateful to you for answering this questionnaire.

Part 1: General Information

- **Kindly mention your company name**
- **Circle one of the following to indicate your role in construction**
a) Consultant b) Contractor c) Owner
- **Circle one of the following to indicate your position in the company**
b) Project Manager b) site engineer c) Quantity surveyor d) accountant e) others
- **What is your cost limit for each project in terms of Bahrain (company classification)?**
a) Class A b) class B c) class C d) class D
- **What is the average job duration for each project?**
a) Up to 6 months b) 6-12 months c) 12-24 months d) 24-36 months e) over 36 months
- **How many years of working experience has your firm been in construction?**
a) 1 to 5 years b) 5 to 10 years c) 10 to 15 years d) over 15 years
- **Are you satisfied with the current level of pretender cost estimate accuracy?**
a) Very satisfied b) satisfied c) Neutral d) not satisfied e) very not satisfied
- **Is there an estimating unit in your organization?**
a) Yes b) No
- **How many employees are working in estimating unit in your organization/ firm?**
a) Less than 3 b) 3-5 c) 4 and more
- **Who sets up the firm tendering price?**
a) At the management level (owner, manager)
b) At the estimating level (cost engineer, estimator)
- **How much you rely on historical data to assess the preparation of your cost estimate?**
a) 100%-90% b) 89%-70% c) 69%-50% d) less than 50%
- **Do you consider the productivity of labor and equipment estimating the cost?**
a) Yes b) No
- **If you choose no, what do you think the reasons for failure of not considering the productivity?**
a) Lack of productivity standards
b) It is too complex to deal with
c) It needs specialist.
d) Others

Part 2: factors are affecting on the accuracy of construction cost estimate.

- **Kindly indicate which factors are affecting construction cost estimation performance (i.e accuracy of project costs determination)**
- **Scope quality** (*Limited to client / consultant*)

Num	Factor	Totally disagree	Disagree	Neutral	Agree	Totally agree
1	Clear scope definition for the client					
2	Previous client / consultant experience of contract					
3	Level of team integration					
4	Buildability of design					
5	Number of other projects with client					
6	Level of involvement of client in the estimation procedure					
7	Client's priority level of the project					

- **Contractor Organization status** (*limited to main contractor*)

Num	Factor	Totally disagree	Disagree	Neutral	Agree	Totally agree
1	Management team (suitability, experience, performance)					
2	Financial capability					
3	Past relationship with client					
4	Past loss/profit in similar projects					
5	Payment record of client					
6	Relationship with subcontractors and suppliers					
7	Current work load					
8	Experience on similar projects					

- **Estimator performance:**

Num	Factor	Totally disagree	Disagree	Neutral	Agree	Totally agree
1	Experience of pricing construction					
2	Impact of team integration and alignment					
3	Estimating method used					
4	Time allowed for preparing cost estimate					

- **Information quality**

Num	Factor	Totally disagree	Disagree	Neutral	Agree	Totally agree
1	Completeness of cost information					
2	Accuracy and reliability of cost information					
3	Clear and detail drawings and specification					

- **Project characteristics**

Num	Factor	Totally disagree	Disagree	Neutral	Agree	Totally agree
1	Type/function of the building					
2	Size/gross floor area					
3	Type of structures (steel, concrete, masonry)					
4	Quality of finishing					
5	Priority of project					
6	Duration of the project period					
7	Inadequate production of raw material for country					
8	Clients financial situation and budget					
9	Location (regions; urban)					

- **External factors**

Num	Factor	Totally disagree	Disagree	Neutral	Agree	Totally agree
1	Cost of materials					
2	Cost of labors					
3	Cost of machinery					
4	Overhead costs					
5	Governments requirements (permits)					
6	Weather					
7	Lack of information and coordination between government agencies					
8	Price fluctuation					
9	Inflationary pressure					
10	Economic insatiability					
11	Currency exchange					
12	Taxation in imported material					
13	Monopoly					

- **Contractual procedure**

Num	Factor	Totally disagree	Disagree	Neutral	Agree	Totally agree
1	Clients evaluation and awarding policy					
2	Contract conditions					
3	Number of bidders on competitive projects					
4	Contingencies allowed					
5	Tendering duration					

6	Time between project announcement, bid submission and award of contract					
7	Availability of other projects for tendering					
8	Performance Bond/warranty arrangements					

Kindly list any other factors may affect the project cost and not been mentioned above.

.....

.....

.....

.....

Appendix 6:

Data analysis for cost deviation in MYS construction projects for the period between (2012-2019)

Table: Data analysis for cost deviation in MYS construction projects for the period between (2012-2019)

ID	Type	Year of award	Year group	Approved Budget (BD)	Project size	Contract amount (BD)	Cost deviation (%)	Restricted criteria	Pessimistic criteria
1	1	2012	1.00	400000.00	2.00	318894.900	25.43	over	over
2	1	2014	2.00	450000.00	2.00	465000.000	-3.23	under	optimal
3	1	2013	1.00	400000.00	2.00	374636.273	6.77	over	optimal
4	1	2014	2.00	450000.00	2.00	412285.500	9.15	over	optimal
5	1	2015	2.00	750000.00	3.00	788700.000	-4.91	under	optimal
6	1	2017	3.00	400015.00	2.00	456021.195	-12.28	under	under
7	3	2013	1.00	150000.00	1.00	162801.870	-7.86	under	optimal
8	3	2014	2.00	249700.00	1.00	192898.156	29.45	over	over
9	3	2014	2.00	182879.40	1.00	142627.012	28.22	over	over
10	3	2014	2.00	51395.30	1.00	37666.464	36.45	over	over
11	3	2014	2.00	83611.00	1.00	79750.000	4.84	over	optimal
12	3	2015	2.00	53000.00	1.00	52800.000	0.38	over	optimal
13	3	2014	2.00	81000.00	1.00	81950.000	-1.16	under	optimal
14	3	2015	2.00	53000.00	1.00	51981.490	1.96	over	optimal
15	3	2017	3.00	247800.00	2.00	296875.700	-16.53	under	under
16	3	2016	3.00	100357.40	1.00	83000.000	20.91	over	over
17	3	2018	4.00	236314.00	2.00	285835.000	-17.33	under	under
18	3	2018	4.00	136722.00	1.00	168769.986	-18.99	under	under
19	3	2018	4.00	241321.00	1.00	194075.761	24.34	over	over
20	3	2019	4.00	222248.25	1.00	208357.900	6.67	over	optimal
21	2	2013	1.00	600000.00	2.00	538368.380	11.45	over	optimal
22	2	2013	1.00	90000.00	1.00	61711.400	45.84	over	over
23	2	2013	1.00	160000.00	1.00	166954.000	-4.17	under	optimal

24	2	2013	1.00	150000.00	1.00	139700.000	7.37	over	optimal
25	2	2014	2.00	340000.00	2.00	251490.146	35.19	over	over
26	2	2015	2.00	150000.00	2.00	240000.000	-37.50	under	under
27	2	2014	2.00	121000.00	1.00	140800.000	-14.06	under	under
28	2	2014	2.00	800000.00	3.00	711975.000	12.36	over	optimal
29	2	2016	3.00	300000.00	2.00	406594.131	-26.22	under	under
30	2	2017	3.00	457000.00	3.00	750000.000	-39.07	under	under
31	2	2018	4.00	830000.00	3.00	1339739.445	-38.05	under	under
32	2	2018	4.00	711168.15	3.00	738921.750	-3.76	under	optimal
33	2	2016	3.00	2000000.00	3.00	1963347.000	1.87	over	optimal
34	2	2018	4.00	250000.00	2.00	286592.756	-12.77	under	under
35	2	2015	2.00	2000000.00	3.00	1978900.223	1.07	over	optimal
36	2	2015	2.00	1260000.00	3.00	1343700.000	-6.23	under	optimal
37	4	2014	2.00	38500.00	1.00	25277.900	52.31	over	over
38	4	2013	1.00	44000.00	1.00	37503.250	17.32	over	over
39	4	2014	2.00	67000.00	1.00	61594.000	8.78	over	optimal
40	4	2017	3.00	144436.60	1.00	123759.471	16.71	over	over
41	4	2014	2.00	57000.00	1.00	83385.145	-31.64	under	under
42	4	2015	2.00	41000.00	1.00	39965.145	2.59	over	optimal
43	4	2017	3.00	69400.00	1.00	51228.800	35.47	over	over
44	4	2017	3.00	31790.00	1.00	30116.000	5.56	over	optimal

Appendix 7:

SPSS and Microsoft Excel regression model development

This Appendix contains the following sections

Section A: SPSS regression model development and validation (model 1)

Section B: SPSS regression model development and validation (model 2)

Section C: Microsoft Excel regression model development and validation (model 1)

Section D: Microsoft Excel regression model development and validation (model 2)

Section A: SPSS regression model development and validation

Budget= F (type,year group,award amount)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.978 ^a	.957	.955	98713.61392
2	.981 ^b	.963	.961	92546.72960

a. Predictors: (Constant), awarded amount

b. Predictors: (Constant), awarded amount, year group

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	9432.135	19946.707		.473	.639	-30947.863	49812.133		
	Awarded amount	.934	.032	.978	28.925	.000	.869	.999	1.000	1.000
2	(Constant)	94647.084	38919.553		2.432	.020	15788.580	173505.588		
	Awarded amount	.951	.031	.996	30.645	.000	.888	1.014	.951	1.051
	Year group	-41249.649	16522.322	-.081	-2.497	.017	-74727.054	-7772.244	.951	1.051

a. Dependent Variable: budget

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	8152838836692.183	1	8152838836692.183	836.671	.000 ^b
	Residual	370286347791.425	38	9744377573.459		
	Total	8523125184483.607	39			
2	Regression	8206223989605.315	2	4103111994802.658	479.061	.000 ^c
	Residual	316901194878.292	37	8564897158.873		
	Total	8523125184483.607	39			

a. Dependent Variable: budget

b. Predictors: (Constant), awarded amount

c. Predictors: (Constant), awarded amount, year group

Validation using the following data: (estimated budget=9432.135+0.934* awarded amount)						
ID	Type	Year group	Awarded amount	Approved Budget	LR estimation (stepwise)	Error
6	1	3.00	456021.195	400015.00	435355.9311	-4.53%
21	3	4.00	208357.900	222248.25	204038.4136	-2.07%
34	2	4.00	286592.756	250000.00	277109.7691	-3.31%
45	4	3.00	30116.000	31790.00	37560.479	24.72%

Note: Linear regression estimates were able to predict the project cost with an error range of (-4.53 to 24.72)%.
Further data is required to improve the model prediction ability.

Validation using the following data: $\text{cost} = 94647.048 + 0.951 \times \text{awarded amount} - 41249.649 \times \text{year group}$						
ID	Type	Year group	Awarded amount	Approved Budget	LR estimation (stepwise)	Error
6	1	3.00	456021.195	400015.00	404574.2634	1.14%
21	3	4.00	208357.900	222248.25	127796.8229	-42.50%
34	2	4.00	286592.756	250000.00	202198.171	-19.12%
45	4	3.00	30116.000	31790.00	-461.577	-101.45%

Note: Linear regression estimates are lower than the actual approved budget, the model is worse than actual

Section B: SPSS regression model development and validation

Budget= F(type,year group,award amount,cost deviation)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.978 ^a	.957	.955	98713.61392
2	.987 ^b	.974	.972	78009.57000

a. Predictors: (Constant), awarded amount

b. Predictors: (Constant), awarded amount, cost dev

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	9432.135	19946.707		.473	.639	-30947.863	49812.133		
	Awarded amount	.934	.032	.978	28.925	.000	.869	.999	1.000	1.000
2	(Constant)	-17130.813	16675.237		-1.027	.311	-50918.053	16656.427		
	Awarded amount	.973	.027	1.018	36.408	.000	.918	1.027	.912	1.096
	Cost dev	2830.702	579.661	.137	4.883	.000	1656.198	4005.205	.912	1.096

a. Dependent Variable: budget

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	8152838836692.183	1	8152838836692.183	836.671	.000 ^b
	Residual	370286347791.425	38	9744377573.459		
	Total	8523125184483.607	39			
2	Regression	8297961943082.880	2	4148980971541.440	681.782	.000 ^c
	Residual	225163241400.727	37	6085493010.830		
	Total	8523125184483.607	39			

a. Dependent Variable: budget

b. Predictors: (Constant), awarded amount

c. Predictors: (Constant), awarded amount, cost dev

Validation using the following data: (Cost=9432.135+0.934* awarded amount)							
id	type	Year group	Awarded amount	Cost deviation	Approved Budget	LR estimation	Error
6	1	3.00	456021.195	-12.28	400015.00	435355.9311	-4.53%
21	3	4.00	208357.900	11.45	222248.25	204038.4136	-2.07%
34	2	4.00	286592.756	-12.77	250000.00	277109.7691	-3.31%
45	4	3.00	30116.000	5.56	31790.00	37560.479	24.72%

Note: Linear regression estimates were able to predict the project cost with an error range of (-2.07 to 24.72)%. Further data is required to improve the model prediction ability.

Section C: Microsoft Excel regression model development and validation (model 1)

$$\text{Budget} = F(\text{type}, \text{year group}, \text{award amount})$$

SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	0.981359
R Square	0.963066
Adjusted R Square	0.959988
Standard Error	93510.64
Observations	40

ANOVA

	Df	SS	MS	F	Significance F
Regression	3	8.21E+12	2.74E+12	312.9043746	7.8506E-26
Residual	36	3.15E+11	8.74E+09		
Total	39	8.52E+12			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	73405.07	58460.89	1.25562	0.217340	-45159.115	191969.24	-45159.115	191969.2
type	9378.573	19098.72	0.49105	0.626365	-29355.41	48112.564	-29355.419	48112.56

Section D: Microsoft Excel regression model development and validation (model 2)

Budget= F(type,year group,award amount, cost deviation)

SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	0.987784288
R Square	0.9757178
Adjusted R Square	0.972942692
Standard Error	76897.00484
Observations	40

ANOVA

	df	SS	MS	F	Significance F
Regression	4	8.31616E+12	2.07904E+12	351.596	9.99126E-28
Residual	35	2.0696E+11	5913149354		
Total	39	8.52313E+12			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	42610.853	48612.220	0.876546	0.3867099	-56077.200	141298.9068
type	-2723.2338	15959.155	-0.1706377	0.8654912	-35122.042	29675.574
Year group	-23944.559	15261.301	-1.5689723	0.1256521	-54926.648	7037.529
Awarded amount	0.9766431	0.0303977	32.128749	1.44589E-27	0.91493	1.03835
cost deviation	2572.7209	602.45895	4.2703671	0.0001418	1349.664	3795.777

Validation using the following data:(Cost=42610.853-2723.2338*Type-23944.559*Year group+0.9766431* awarded amount+2572.7209* cost deviation)						
type	Year group	Awarded amount	Cost deviation	Approved Budget	LR estimated	Error
1	3.00	456021.195	-12.28	400015.00	381830.9155	-16.27%
3	4.00	208357.900	11.45	222248.25	171611.8891	-17.64%
2	4.00	286592.756	-12.77	250000.00	188431.3601	-34.25%
4	3.00	30116.000	5.56	31790.00	3601.153313	-88.04%

Note: Linear regression estimates are lower than the actual approved budget, the model is worse than actual.

Appendix 8:
Questionnaire Survey - RII calculation

This Appendix contains the following sections

Section A: RII calculation for MYS participants

Section B: MOW RII calculation (cross-validation)

Section A: RII calculation for MYS

num	group	factor	totally disagree	disagree	Neutral	agree	totally agree	n	RII
factor rating			1	2	3	4	5		
1	SQ	Previous client / consultant experience of contract	0	2	7	11	4	24	74.17%
2		Level of involvement of client in the estimation procedure	0	2	3	9	10	24	82.50%
3		Number of other projects of same client	0	2	8	11	3	24	72.50%
4		Level of team integration	0	1	7	9	7	24	78.33%
5		Clients priority level of the project	0	2	7	9	6	24	75.83%
6		Clear scope definition for the client	0	1	4	7	12	24	85.00%
7		Buildability of design	0	0	7	11	6	24	79.17%
8	EP	Experience of pricing construction	0	0	7	21	22	50	86.00%
9		Impact of team integration and alignment	1	1	16	23	9	50	75.20%
10		Estimating method used	0	0	9	28	13	50	81.60%
11	IQ	Completeness of cost information	0	2	8	21	19	50	82.80%
12		Accuracy and reliability of cost information	0	1	10	16	23	50	84.40%
13	QEP	Time allowed for preparing cost estimate	0	2	9	24	15	50	80.80%
14		Clear and detail drawings and specification	0	0	14	19	17	50	81.20%
15	OM	Management team (suitability, experience, performance)	1	1	1	11	12	26	84.62%
16		Financial capability	1	1	3	11	10	26	81.54%
17		Experience on similar projects	1	1	1	8	15	26	86.92%
18		Current work load	1	3	10	7	5	26	69.23%
19		Past relationship with client	1	0	3	6	16	26	87.69%
20		Payment record of client	0	2	4	6	14	26	84.62%
21		Relationship with subcontractors and suppliers	0	2	3	12	9	26	81.54%
22		Past loss/profit in similar projects	0	1	7	12	6	26	77.69%
23	PC	Type/function	0	3	9	16	22	50	82.80%
24		Size/gross floor area	0	2	12	18	18	50	80.80%
25		Type of structures (steel, concrete, masonry)	0	0	9	18	23	50	85.60%
26		Location (regions; urban)	4	2	14	15	15	50	74.00%
27		Quality of finishing	0	3	6	14	27	50	86.00%
28		Priority of project	0	2	7	24	17	50	82.40%
29		Duration of the project period	0	1	9	19	21	50	84.00%
30		Clients financial situation and budget	1	1	10	17	21	50	82.40%
31		Inadequate production of raw material for country	3	3	11	20	13	50	74.80%
32	EF	Cost of materials	0	1	9	18	22	50	84.40%
33		Cost of labors	0	0	9	24	17	50	83.20%
34		Cost of machinery	0	1	10	23	16	50	81.60%

35		Overhead costs	1	1	11	22	15	50	79.60%
36		Governments requirements	1	2	17	20	10	50	74.40%
37		Lack of information and coordination between government agencies	2	7	11	21	9	50	71.20%
38		Weather	1	8	17	16	8	50	68.80%
39	UL	Economic insatiability	1	1	16	26	6	50	74.00%
40		Inflationary pressure	1	0	14	27	8	50	76.40%
41		Currency exchange	2	9	22	9	8	50	64.80%
42		Taxation in imported material	0	9	11	15	15	50	74.40%
43		Monopoly	1	6	18	16	9	50	70.40%
44		Price fluctuation	0	7	12	19	12	50	74.40%
45	CP	Clients evaluation and awarding policy	1	2	17	22	8	50	73.60%
46		Contract conditions	1	4	7	24	14	50	78.40%
47		Time between project announcement, bid submission and award of contract	0	5	15	18	12	50	74.80%
48		Availability of other projects for tendering	1	5	18	19	7	50	70.40%
49		Number of bidders on competitive projects	1	3	14	20	12	50	75.60%
50		Performance Bond/warranty arrangements	3	6	13	14	14	50	72.00%
51		Contingencies allowed	3	3	15	16	13	50	73.20%
52		Tendering duration	0	5	17	14	14	50	74.80%

Section B: MOW RII calculation (cross-validation)

Num	Factors	Sample	1	2	3	4	5	MOW RII	Study RII
1	Clear scope definition for the client	6	0	0	1	1	4	90%	85%
2	Previous client / consultant experience of contract	6	0	0	1	2	3	87%	74.17%
3	Level of team integration	6	0	0	2	0	4	87%	78.33%
4	Buildability of design	6	0	1	1	2	2	77%	79.17%
5	Number of other projects with client	6	0	0	2	2	2	80%	72.5%
6	Level of involvement of client in the estimation procedure	6	0	2	1	3	0	63%	82.5%
7	Client's priority level of the project	6	0	0	1	2	3	87%	75.83%
8	Management team (suitability, experience, performance)	5	0	0	0	4	1	84%	84.62%
9	Financial capability	5	0	0	0	4	1	84%	81.54%
10	Experience on similar projects	5	0	0	0	1	4	96%	86.92%
11	Current work load	5	0	0	1	3	1	80%	69.23%
12	Past relationship with client	5	0	0	1	1	3	88%	87.69%
13	Payment record of client	5	0	0	1	2	2	84%	84.62%
14	Relationship with subcontractors and suppliers	5	0	0	0	2	3	92%	81.54%
15	Past loss/profit in similar projects	5	0	0	0	3	2	88%	77.69%
16	Experience of pricing construction	11	0	0	2	3	6	87%	86%
17	Impact of team integration and alignment	11	0	0	1	6	4	85%	75.2%
18	Estimating method used	11	0	0	2	7	2	80%	81.6%
19	Time allowed for preparing estimate	11	0	0	1	8	2	82%	80.8%
20	Completeness of cost information	10	0	0	0	5	5	90%	82.8%
21	Accuracy and reliability of cost information	11	0	0	0	5	6	91%	84.4%
22	Clear and detailed drawings and specification	11	0	0	0	5	6	91%	81.2%
23	Type/function (playground, administration building, model youth center, sport hall)	11	0	1	2	6	2	76%	82.8%
24	Type of structures (steel, concrete, masonry)	11	0	2	0	4	5	82%	85.6%
25	Quality of finishing	11	0	1	1	3	6	85%	86%
26	Priority of project	11	0	0	2	5	4	84%	82.4%
27	Duration of the project period	11	0	1	1	5	4	82%	84%
28	Clients financial situation and budget	11	0	0	1	6	4	85%	82.4%

29	Inadequate production of raw material for country	11	0	0	1	5	5	87%	74.8%
30	Cost of materials	11	0	0	0	5	6	91%	84.4%
31	Cost of labors	11	0	0	2	5	4	84%	83.2%
32	Cost of machinery	11	0	0	1	6	4	85%	81.6%
33	Overhead costs	11	0	0	2	5	4	84%	79.6%
34	Governments requirements	11	0	1	2	3	5	82%	74.4%
35	Inflationary pressure	11	0	0	2	8	1	78%	76.4%
36	Price fluctuation	11	0	0	3	5	3	80%	74.4%
37	Weather	11	0	0	5	4	2	75%	68.8%
38	Contract conditions	11	0	0	3	5	3	80%	78.4%
39	Time between project announcement, bid submission and award of contract	11	0	0	1	10	0	78%	74.8%
40	Number of bidders on competitive projects	11	0	0	2	8	1	78%	75.6%
41	Performance Bond/warranty arrangements	11	0	2	1	8	0	71%	72%
42	Availability of other projects for tendering	11	0	1	3	6	1	73%	70.4%

Appendix 9

ANN Model Results And Mathematical Equation Development

This appendix consists of the followings:

Table 1: ANN (5-10-1) Satisfaction prediction performance

Table 2: Calculation of cost estimation satisfaction mathematical model

Table 1: ANN (5-10-1) Satisfaction prediction performance

Respondent number	Inputs					Target satisfaction	ANN_Output	Error
	CP_av	EF_av	EP_av	IQ_av	PC_av			
1	3.5	4	4.25	3.667	2.833	4	4.046	-0.046
2	5	5	4.5	4.333	5	5	3.971	1.029
3	2.5	3.75	3	3.667	3.167	3	2.863	0.137
4	4.5	4.5	3.75	3.667	4.5	4	3.882	0.118
5	3.5	3	3	3	3	3	2.678	0.322
6	3.5	3.25	3.5	2.667	3.5	3	3.272	-0.272
7	3	3.5	3.5	2.333	3.667	2	1.914	0.086
8	5	4	4.25	5	5	4	3.837	0.163
9	4	4	4	4	3.833	2	2.194	-0.194
10	5	3.5	5	5	3.667	1	0.684	0.316
11	4.5	3	4.5	5	4.333	5	3.825	1.175
12	4	4	4.25	3.667	4.333	4	3.789	0.211
13	4.5	5	4.75	5	4.5	3	2.496	0.504
14	4	2.5	3	3.333	3.333	4	3.764	0.236
15	3.5	4	4	4	4	4	3.347	0.653
16	3.5	3.5	4	3.667	4	3	3.505	-0.505
17	4	3.75	4	4	3.833	3	2.639	0.361
18	3.5	4.5	4	4.333	4	3	2.402	0.598
19	3	3.75	4	4.333	3.667	3	3.298	-0.298
20	3.5	5	4.5	5	4	3	2.555	0.445
21	3	4.25	3.75	4	3.833	3	2.895	0.105
22	4.5	4.75	4.5	4.667	4.667	4	4.100	-0.100
23	4	5	4	4.667	4.167	2	1.767	0.233
24	4.5	5	3.75	4.333	4.833	5	3.930	1.070
25	4	4.75	3.75	3.333	4	1	0.900	0.100
26	4.5	5	4.25	4.667	4.667	4	3.816	0.184
27	5	3.5	3.5	4.667	4	3	3.729	-0.729
28	3	4.5	4.25	4	3.833	3	2.908	0.092
29	4.5	4	4.75	5	4.667	5	5.222	-0.222
30	3.5	4	4	4	4	4	3.347	0.653
31	3.5	4	4	4.667	3.833	3	2.483	0.517
32	4	5	5	4.333	5	4	3.937	0.063
33	2.5	2.75	4.75	5	4.667	4	3.993	0.007
34	3.5	4	3.5	3	4.333	4	3.267	0.733
35	3	4	4.25	4	4.167	4	4.025	-0.025
36	3.5	3.5	4.75	4.667	4.833	5	5.271	-0.271
37	3	4.25	4.75	5	4.833	2	2.206	-0.206
Respondent number	Inputs					Target satisfaction	ANN_Output	Error
	CP_av	EF_av	EP_av	IQ_av	PC_av			
38	4	3	4.5	4.333	4.5	3	3.529	-0.529

39	2.5	3.75	3.5	4.667	3.833	4	4.099	-0.099
40	5	4.5	3.75	4	4.167	5	3.859	1.141
41	3.5	4.75	2.75	3	4.5	2	2.131	-0.131
42	1.5	3	4	3.667	2.5	5	5.088	-0.088
43	4	4.25	3.25	3.333	5	3	3.938	-0.938
44	5	4.75	5	4.667	5	4	4.647	-0.647
45	5	4.5	4.25	5	5	3	4.044	-1.044
46	5	5	5	5	5	5	4.797	0.203
47	4.5	4.5	4	4	4.5	4	3.916	0.084
48	4	5	4	4.667	4.667	4	3.768	0.232
49	3.5	3.75	4	4	3.667	3	1.676	1.324
50	4	5	3	3	4.833	4	2.084	1.916

Table 2: Calculation of cost estimation satisfaction mathematical model

Step 1: Tabulation ANN connection weight for input , output and bias								
ANN connection weights								
	input hidden weight					hidden output weight	bias	
Node	CP_av	EF_av	EP_av	IQ_av	PC_av	satisfaction	hidden	output
A	-3.074	1.222	1.854	0.707	-3.114	1.390	-1.692	-0.794
B	-1.733	4.105	1.182	-3.482	-1.552	-0.526	-1.533	
C	4.331	-0.015	-2.972	-0.988	-1.062	-0.438	0.573	
D	0.626	-4.115	-1.153	2.807	-0.689	-0.055	-1.325	
E	-0.516	1.635	2.495	0.247	-4.928	-0.891	-0.124	
F	-1.159	-0.331	-0.445	1.289	-0.285	0.006	-1.234	
G	1.237	1.255	-0.790	-0.607	-0.784	0.982	1.093	
H	-1.155	1.908	-0.899	3.707	2.004	-0.046	1.278	
I	2.429	-1.071	-0.437	-1.623	-0.402	1.858	2.825	
J	1.475	0.844	-1.157	-1.814	-0.681	-1.109	2.358	

Step 3: Enter the input values and standrized them							
	CP_av	EF_av	EP_av	IQ_av	PC_av	target	
input	5	4	4.25	5	5	4	
min	1.5	2.5	2.75	2.333	2.5	1	
max	5	5	5	5	5	5	
standrized input	1.000	0.200	0.333	1.000	1.000	standrized:	=(2*((C3-C4)/(C5-C4)))-1

Step 4: Nodes equations formulation :								
Node	hidden input connection weight*standrized input					bias	sum hidden node	sumh
A	-3.074	0.244	0.618	0.707	-3.114	-1.692	-6.311	=SUM(C10:H10)
B	-1.733	0.821	0.394	-3.482	-1.552	-1.533	-7.085	
C	4.331	-0.003	-0.991	-0.988	-1.062	0.573	1.860	
D	0.626	-0.823	-0.384	2.807	-0.689	-1.325	0.211	
E	-0.516	0.327	0.832	0.247	-4.928	-0.124	-4.162	
F	-1.159	-0.066	-0.148	1.289	-0.285	-1.234	-1.604	
G	1.237	0.251	-0.263	-0.607	-0.784	1.093	0.926	
H	-1.155	0.382	-0.300	3.707	2.004	1.278	5.916	
I	2.429	-0.214	-0.146	-1.623	-0.402	2.825	2.869	
J	1.475	0.169	-0.386	-1.814	-0.681	2.358	1.122	

Step 4: multiplication hidden output* transfer functuon (sum of hidden node)				step 5: calculation of xnor step 7: denormalize xnor		=SUM(E24:E33)+F24		
Node	hidden	tansig(sumh)	multiplication	output bias	xnormalized	sat denor	error	ANN_out
A	1.390	-1.000	-1.390	-0.794	0.418	3.982	0.018	3.837
B	-0.526	-1.000	0.526					
C	-0.438	0.953	-0.418					
D	-0.055	0.208	-0.011					
E	-0.891	-1.000	0.890					
F	0.006	-0.922	-0.006					
G	0.982	0.729	0.716					
H	-0.046	1.000	-0.046					
I	1.858	0.994	1.846					
J	-1.109	0.808	-0.896					

xnormalized

=SUM(E24:E33)+F24

sat denor

=(0.5*(G24+1)*(5-1.5))+1.5

error

=H3-H24

tansig(sumh)=
=(2/(1+EXP(-2*I10)))-1

Appendix 10

Categories Ranking As Per Normalized Connection Weight Method

Table: Categories ranking as per normalized connection weight method

step 1: Tabulation of both input & output connection weights from the ANN network

		hidden nodes									
inputs		1	2	3	4	5	6	7	8	9	10
	CP_av	-3.074	-1.733	4.331	0.626	-0.516	-1.159	1.237	-1.155	2.429	1.475
	EF_av	1.222	4.105	-0.015	-4.115	1.635	-0.331	1.255	1.908	-1.071	0.844
	EP_av	1.854	1.182	-2.972	-1.153	2.495	-0.445	-0.790	-0.899	-0.437	-1.157
	IQ_av	0.707	-3.482	-0.988	2.807	0.247	1.289	-0.607	3.707	-1.623	-1.814
	PC_av	-3.114	-1.552	-1.062	-0.689	-4.928	-0.285	-0.784	2.004	-0.402	-0.681

		hidden nodes									
output		1	2	3	4	5	6	7	8	9	10
	satisfaction	1.390	-0.526	-0.438	-0.055	-0.891	0.006	0.982	-0.046	1.858	-1.109

step 2: Multiplication of input hidden connection weight by output connection

hidden weight =C7*\$C\$15

input	1	2	3	4	5	6	7	8	9	10	sum	influence
CP_av	-4.272	0.912	-1.898	-0.034	0.460	-0.007	1.215	0.053	4.514	-1.636	-0.694	negative
EF_av	1.699	-2.161	0.006	0.225	-1.456	-0.002	1.232	-0.087	-1.990	-0.936	-3.469	negative
EP_av	2.578	-0.622	1.303	0.063	-2.222	-0.003	-0.775	0.041	-0.812	1.283	0.834	positive
IQ_av	0.982	1.833	0.433	-0.154	-0.220	0.008	-0.596	-0.169	-3.017	2.011	1.111	positive
PC_av	-4.329	0.817	0.466	0.038	4.389	-0.002	-0.770	-0.092	-0.747	0.755	0.524	positive

step 3: Compute the importance score by Summating inputs hidden connection

weights =SUM(B21:K21)

input	importance score
CP_av	-0.694
EF_av	-3.469
EP_av	0.834
IQ_av	1.111
PC_av	0.524

step 4: Normalize the importance scores and rank accordingly

=(C31-\$C\$39)/(\$C\$40-\$C\$39)

min	-3.469
max	1.111

input	normilize score	rank
CP_av	0.606	4
EF_av	0.000	5
EP_av	0.940	2
IQ_av	1.000	1
PC_av	0.872	3

ملخص الدراسة

تقدير التكاليف هي واحدة من المهام الحاسمة التي يتم التعامل معها في مرحلة بدء المشروع ، ولها تأثير كبير على أداء المشروع. لوحظ في وزارة شؤون الشباب والرياضة في مملكة البحرين وجود مؤشرات تدعم إمكانية وجود مشاكل في عملية تقدير تكاليف بناء المشاريع الرياضية. المشكلة التي يجب التحقيق فيها في هذا البحث هي عدم دقة تقدير تكاليف مشاريع البناء في وزارة شؤون الشباب والرياضة. الأهداف الرئيسية لهذا البحث تتمثل ب : تقييم مستوى دقة تقدير تكاليف مشاريع البناء في وزارة شؤون الشباب والرياضة ، وتحديد العوامل التي تؤثر على عملية تقدير تكاليف مشاريع البناء الحالية ، وتطوير نموذج لتقدير تكلفة المشاريع مما سيساهم في تعزيز عملية تقدير التكاليف.

تم تحقيق أهداف الدراسة من خلال مراجعة الأدبيات المعنية حول القضايا الرئيسية الثلاث: مستوى دقة تقدير التكاليف ، والعوامل التي تؤثر على دقة تقدير التكاليف ، ووضع نموذج لتقدير التكاليف. البيانات المستخدمة لتحقيق هذه الأهداف هي بيانات مشاريع البناء المنفذة من قبل وزارة شؤون الشباب والرياضة، بالإضافة إلى استخدام الاستبيان لجمع المعلومات من أصحاب المصلحة الذين تم التعامل معهم في تنفيذ مشاريع وزارة شؤون الشباب والرياضة.

تم تطوير نموذج تقدير التكلفة باستخدام نموذج الانحدار الخطي المتعدد (MLR). وكذلك تم استخدام تقنيتين مجتمعين لمعرفة العوامل المؤثرة على عملية تقدير التكاليف: نمذجة المعادلات الهيكلية القائمة على المربعات الصغرى الجزئية (PLS-SEM) والشبكة العصبونية الاصطناعية (ANN). قدم تطبيق المعادلات الهيكلية القائمة على المربعات الصغرى الجزئية على العوامل الحاسمة التي حددها المشاركون الدعم النظري للإطار المفاهيمي للدراسة. في حين حدد نمذجة العوامل باستخدام الشبكة العصبونية الاصطناعية العوامل التي تؤثر على تقدير تكلفة البناء في وزارة شؤون الشباب والرياضة.

يقدم هذا البحث نظرة ثاقبة على مستوى تقدير التكاليف الحالي في وزارة شؤون الشباب والرياضة، والذي يوفر فرصا لتحسين عملية تقدير التكاليف خلال مرحلة تصميم المشاريع. وأيضاً ، تم تأكيد العلاقات بين متغيرات البحث: توافر وحدة تقدير التكاليف ، جودة المعلومات ، أداء المقدر ، خصائص المشروع ، العوامل الخارجية ، والإجراءات التعاقدية على دقة تقدير تكاليف البناء. تشير الدراسة إلى الحاجة إلى التحكم في العوامل الأعلى مرتبة: جودة المعلومات وأداء المقدر وخصائص المشروع لتعزيز كفاءة عملية تقدير التكاليف البناء.



**تطوير نموذج تقدير التكلفة للمشاريع الإنشائية التابعة لوزارة شؤون
الشباب والرياضة
(دراسة حالة من مملكة البحرين)**

أطروحة مقدمة كجزء من متطلبات الحصول على درجة الماجستير في الإدارة الهندسية

إعداد

فاطمة صقر يوسف مال الله

20083399

إشراف

أ.د. سعد سليمان

(أستاذ)

د. أسامة الجمل

(أستاذ مساعد)

مملكة البحرين

أكتوبر 2021