

University of Bahrain

College of Engineering



Department of
Mechanical Engineering

Causes and Effects of Cost Overrun on Construction Projects in Bahrain

Research Proposal submitted in partial fulfillment of the requirements
for the degree of Master of Science

Prepared by:

Hedaya Abdulsamie Mohammed Abu Safiya

20071812

Supervised by:

Prof. Saad Suliman

(Professor)

University of Bahrain

Kingdom of Bahrain

September 2016

University of Bahrain

College of Engineering

Mechanical Engineering Dept.



Causes and Effects of Cost Overrun On Construction Projects in Bahrain

Research Proposal submitted in partial fulfillment of the requirements for
the degree of Master of Science
In Engineering Management

Prepared by:

Hedaya Abdulsamie Mohammed Abu Safiya

20071812

Supervised by:

Prof. Saad Suliman

(Professor)

University of Bahrain

Kingdom of Bahrain

September 2016

Abstract

The main concern of a construction industry is to procure a facility that is able to meet its functional requirements, of the required quality, and delivered within an acceptable budget and timeframe. The cost aspect of these key performance indicators usually ranks highest. In spite of the importance of cost estimation, construction projects have routinely overrun their estimates due to a number of risks that effect final cost of a project. This study investigated the effect of these risk factors affecting cost performance in construction projects. Investigation was carried out using quantitative and semi qualitative approaches of: questionnaire survey, actual projects data and interviews to understand the perception of practitioners involved in construction industry towards various factors causing cost overrun. The targeted respondents were client, contractor, and consultant involved in handling construction projects in Bahrain. A total of 74 completed responses and 40 actual construction project data were collected. Collected data were analyzed with risk zone mapping by creating a risk map for cost overrun factors using severity and frequency indices, then risk factors were ranked according to their importance index. Data were also analyzed with advance multivariate statistical approach of Partial Least Square Structural Equation Modelling (PLS-SEM). It modelled the relationship of various factors and their relative effects to cost overrun.

Based on the final risk map result, frequent design changes, mistakes during construction, and schedule delay were the most prominent factors in exhibiting effect on cost performance in construction projects. Outcome from the model identified that project management and contract administration related factors group has the highest impact on construction cost overrun with path coefficient value of 0.286. The developed model indicated that approximately 60% of variance extraction was achieved, which means that in Bahrain 60% of cost overrun is subjected to these factors. Further, Global Fit Index value of the models was 0.591, which confirms that the developed path models had substantial power in explaining the relationship between the factors and cost overrun in construction projects. It is hoped that these findings will guide efforts to improve the performance of the construction industry in the future.

Keywords: Construction industry, Cost overrun, Risk mapping, PLS-SEM model, Bahrain

Dedication

I dedicate this thesis to my beloved parents,
you have successfully made me the person I am becoming,
thanks for your endless love, support and encouragement.

Acknowledgment

Foremost, I would like to express my sincere gratitude to my advisor Prof. Saad Suliman for the continuous support of my research, for his patience, motivation, and immense knowledge. I would also like to express my appreciation to all organizations and individuals who contributed directly or indirectly to this thesis and provided the necessary materials and support for realization of this thesis. Finally I would like to give my special thanks for my family and friends whose love, support, and encouragement enabled me to complete this work.

Table of Contents

Subject	Page No.
Thesis committee accreditation	ii
Abstract	iii
Dedication	iv
Acknowledgements	v
Table of contents	vi
List of tables	x
List of figures	xiii
List of appendices	xv
Glossary of abbreviations	xvi
Chapter One Introduction to the Study	
1.1 Background	2
1.2 Problem statement	3
1.3 Research Questions	4
1.4 Research Hypothesis	5
1.5 Aim and objectives of the research	6
1.6 Research design and methodology	7
1.7 Thesis outline	9
Chapter Two Review of Literature	
2.1 Introduction	11
2.2 The scale of cost overruns in construction	12
2.3 Impact of risk factors in construction projects	14
2.4 Causes of cost overrun in construction existing studies	18
2.5 Studies conducted to control construction cost overrun	24
2.6 Literature summary	29
Chapter Three Data Collection	

3.1 Data of construction projects	33
3.2 Survey sample size	34
3.3 Questionnaire design and distribution	36
3.4 Actual projects data collection	40
3.5 Interviews Data Collection	42
Chapter Four Data Analysis	
4.1 Introduction	46
4.2 Analysis formulae	46
4.2.1 Survey data	46
4.2.2 Actual projects data	47
4.3 Risk mapping	48
4.3.1 Risk map development	48
4.3.2 Standard risk map	49
4.3.3 Risk maps results	57
4.4 Ranking of cost overrun factors and groups	57
4.5 Analysis of results by data source	59
4.6 Survey based cost overrun factors analysis	61
4.6.1 Factoring method using the principal component analysis (PCA)	62
4.6.2 Sample size	63
4.6.3 Factor analysis tests	63
4.7 Survey based data and actual based data ranking analysis	74
Chapter Five Modelling Conceptual Background	
5.1 Introduction	79
5.2 Different Approaches to SEM	80
5.3 PLS-SEM Theoretical Background	82
5.3.1 The structural model	82
5.3.2 The measurement model	83
5.3.3 The PLS-SEM algorithm	86
5.4 Determination of Sample Size in PLS-SEM	88
5.5 PLS-SEM Model Assessment	89
5.5.1 Assessment of measurement model	90
5.5.2 Assessment of structural model	93

5.6 Model Evaluation	94
5.6.1 Overall model assessment (GoF)	94
5.6.2 Assessing effect size f^2	95
5.6.3 Assessing the predictive relevance Q^2	96
Chapter Six Model Development and Evaluation	
6.1 Introduction	99
6.2 Model Input Data	100
6.2.1 Combined data reliability test	101
6.2.2 Factor analysis	101
6.3 Model Development	102
6.3.1 Building the inner model	102
6.3.2 Building the outer model	103
6.3.3 Rerunning the path-modelling estimation	103
6.3.4 Model assessment	105
6.3.5 Rerunning of the path-modelling estimation (second run)	108
6.3.6 Assessment of inner model	114
6.3.7 Assess the level of R^2 and path coefficient significance	115
6.3.8 Assess structural model relationships significance and relevance	116
6.3.9 Summary	118
6.4 Model Evaluation	119
6.4.1 Goodness of fit	119
6.4.2 Assessing effect size f^2	120
6.4.3 Assessing the predictive relevance Q^2	121
6.5 Model Contribution	122
6.6 Risk Mapping and PLS-SEM Model Results Comparison	123
Chapter Seven Conclusion and Recommendations	
7.1 Conclusion	126
7.1.1 Cost overrun risk mapping	126
7.1.2 Factor analysis	127
7.1.3 Model development using PLS-SEM method	128
7.2 Recommendations	129

7.2.1 Top cost overrun factors recommendations	131
7.3 Research limitations	132
7.3.1 Model limitations	132
7.3.2 Generalization limitations	133
7.3.3 Data limitations	133
7.4 Further Research	134

List of Tables

No.	Subject	Page
1	Table 2.1: Causes of Cost Overrun Groups and Their Impact over Cost Overrun, According to Allahaim and Dr. Liu (2012)	17
2	Table 2.2: Causes of Cost Overrun from Previous Studies Summary	22
3	Table 2.3: Selected Groups and Factors of Cost Overrun	30
4	Table 4.1: Actual Projects Degree of Severity	48
5	Table 4.2: Scale Used to Identify Severity and Frequency Level and Associated Effect on Cost	49
6	Table 4.3: Risk Mapping for Survey Cost Overrun Factors	51
7	Table 4.4: Factors of Cost Overrun Risk Mapping for Actual Data and Combined	53
8	Table 4.5: Ranking of Cost Overrun Factors Based on the Important Index of the Combined Data	57
9	Table 4.6: Groups of Cost Overrun Ranked Using Combined Average Important Index	59
10	Table 4.7: Ranking of Top 10 Cost Overrun Factors	59
11	Table 4.8: KMO and Bartlett's Test	64
12	Table 4.9: Initial and Extraction Communalities for Cost Overrun Variables	66
13	Table 4.10: Total Variance Explained	67
14	Table 4.11: Rotated and Ordered Component Matrix	71
15	Table 4.12: Principal Factor Groupings from Rotated Component Matrix	72
16	Table 4.13: Highest 13 Factors Extracted from PCA (Survey)	74

17	Table 4.14: Cost Overrun in Construction Projects, Factors Ranking Based on Factor Weight	74
18	Table 4.15: Cost Overrun in Construction Projects, Common Factors	75
19	Table 5.1: Suggested Sample Size in a Typical Research	89
20	Table 5.2: Individual Factor (Indicator) Reliability Assessment	90
21	Table 5.3: The Cut-off Value for AVE, CR and Cronbach Alpha	91
22	Table 5.4: GOF Cut-off Values for Global Validation	95
23	Table 5.5: Guidelines for Assessing f^2 Values	96
24	Table 6.1: Combined Data Reliability Test	101
25	Table 6.2: KMO and Bartlett's Test	101
26	Table 6.3: Results Summary for Reflective Outer Models (individual item reliability and convergent validity)	106
27	Table 6.4: Results Summary for Reflective Outer Models (individual item reliability and convergent validity)	109
28	Table 6.5: Checking Discriminant Validity	111
29	Table 6.6: Indicators (Cost Overrun Factors) Outer Weights (Relative Importance)	112
30	Table 6.7: Mean, SD, t-Statistics and P Values of Path Coefficients (Outer Model)	113
31	Table 6.8: Mean, SD, T-Statistics and P Values of Path Coefficients (Inner Model)	116
32	Table 6.9: Summary of f^2 Effect Sizes	120
33	Table 6.10: Groups of Cost Overrun Effects on Total Project Cost Overrun (Model Results)	122
34	Table 6.11: Cost Overrun in Construction Projects in Bahrain: Risk	124

	Mapping and PLS-SEM Model Groups Ranking	
35	Table 6.12: Cost Overrun in Construction Projects in Bahrain: Risk Mapping and PLS-SEM Model Groups Matching	124

List of Figures

No.	Subject	Page
1	Figure 3.1: The Distribution of the Respondents According to their Organization Type	37
2	Figure 3.2(A): The Distribution of the Respondents According to their Companies' Grades	38
3	Figure 3.2(B): The Distribution of the Respondents According to their Companies' Projects Values	38
4	Figure 3.3(A): The Distribution of the Respondents According to their Level of Qualification	39
5	Figure 3.3(B): The Distribution of the Respondents According to their Work Experience	39
6	Figure 3.4: The Frequency of Cost Overrun Experienced by Participants	40
7	Figure 3.5: Projects Types Percentages in Data Sample	41
8	Figure 3.6: Construction Project Contract Type	42
9	Figure 3.7: The Interviewees Numbers and Job Titles	43
10	Figure 3.8: Process Used to Determine Main Causes of Cost Overruns in Actual Construction Projects	44
11	Figure 4.1: Zones of the Risk Map	50
12	Figure 4.2: Scree Plot of Cost Overrun Components	70
13	Figure 5.1: Inner and Outer Model	80
14	Figure 5.2: Reflective Model Path Diagram	85
15	Figure 5.3: Reflective LV Path Diagram	85
16	Figure 5.4: The Diagram Depicts the Flowchart for the PLS Algorithm	86

17	Figure 6.1: Research Process	100
18	Figure 6.2: Inner Model (Structural Model) Framework	102
19	Figure 6.3: Outer Model (Measurement Model) Framework	103
20	Figure 6.4: PLS-SEM Results (First Run)	104
21	Figure 6.5: PLS-SEM Results (Second Run)	108
22	Figure 6.6: Bootstrapping Results, the t-Statistics for Outer Model	114
23	Figure 6.5: Bootstrapping Results - Path Coefficients for Inner Model	117

List of Appendices

No.	Appendix
(1)	Questionnaire Format
(2)	Actual Project Data for the Period 2008-09 to 2012-13
(3)	Survey and Actual Data Reliability Tests
(4)	PLS-SEM Theoretical Background

Glossary of Abbreviations

AMOS	Analysis of Moment Structure
AVE	Average Variance Extracted
BTS	Bartlett Test of Sphericity
CB-SEM	Covariance-Based Structural Equation Modelling
CC	Contractual Claims
CF	Cash Flow
CR	Composite Reliability
CSF	Critical Success Factor
CSM	Contractor's Site Management Related Factors
CSP	Change in the Scope of the Project
DDF	Design and Documentation Related Factors
DDM	Delays in Decisions Making
DPAD	Delay Preparation and Approval of Drawings
DPPO	Delay in Progress Payment by Owner
DPS	Delay Payment to Supplier /Subcontractor
EAF	Equipment Availability and Failure
EI	Economic Instability
EQS	Environmental Quality Standard
EV	Environmental Related Factors
EW	Effects of Weather
FA	Factor Analysis
FDC	Frequent Design Changes
FDO	Financial Difficulties of Owner

FIN	Financial Management Related Factors
FPM	Fluctuation of Prices of Materials
GDP	Gross Domestic Product
GOF	Global Fit Measures Goodness of Fit
GP	Government Policies
GSCA	Generalized Structured Component Analysis
HCL	High Cost of Labour
HR	Human Resources (Workforce) Related Factors
ICT	Information and Communication Related Factors
IDTT	Incomplete Design at the Time of Tender
IMC	Inadequate Monitoring and Control
IPRM	Inadequate Production of Raw Materials by the Country
IPS	Inadequate Planning and Scheduling
IQT	Inaccurate Quantity Take-off
IS	Incompetent Subcontractors
ITCE	Inaccurate Time and Cost Estimates
KMO	Kaiser-Meyer-Olkin
LA	Labour Absenteeism
LCMP	Lack of Communication Between Parties
LCOP	Lack of Coordination Between Parties
LDME	Late Delivery of Materials and Equipment
LISAREL	Linear Structural Relations
LNC	Level and Number of Competitors
LOE	Lack of Experience
LP	Labour Productivity
LVs	Latent Variables

MDC	Mistakes During Construction
MED	Mistakes and Errors in Design
MS	Monopoly by Suppliers
MV	Manifest Variables / Indicators
NEUSREL	Nonlinear Universal Structural Relational Modelling
NHR	Non-Human Resource Related Factors
PCA	Principal Component Analysis
PDDD	Poor Design and Delays in Design
PFCS	Poor Financial Control on Site
PL	Political Situation
PLS	Partial Least Squares
PMCA	Project Management and Contract Administration Related Factors
PPM	Poor Project Management
PS	Number of Projects Going at the Same Time
PSMS	Poor Site Management and Supervision
SD	Schedule Delay
SEM	Structural Equation Modelling
SIFP	Slow Information Flow Between Parties
SM	Shortages of Materials
SPSS	Statistical Package for the Social Sciences
SSW	Shortage of Site Workers
STP	Shortage of Technical Personnel

CHAPTER ONE

INTRODUCTION TO THE STUDY

1.1 Background

1.2 Problem Statement

1.3 Research Questions

1.4 Research Hypothesis

1.5 Aim and Objectives of the Research

1.6 Research Design and Methodology

1.7 Thesis Outline

CHAPTER ONE

INTRODUCTION TO THE STUDY

1.1 Background

The construction industry is truly the engine of national economy through which the total of physical development is achieved. The construction industry has also a significant effect on the efficiency and productivity of other industry sectors. One cannot think of widespread investment in manufacturing, agriculture, or service sectors unless the construction results of facilities are in place. However; nine out of ten construction projects overrun their budget, with an 86% probability of outrunning their set cost targets according to Flyvbjerg et al. (2002, 2003). Love et al. (2012) report overruns of up to 70% more than the initial estimate, while Odeck (2004) reports a cost overrun up to 183% of original cost, and Assaf and Al-Hejji (2006) found that only 30% of construction projects in Saudi Arabia were completed within the specified cost. These statistics might suggest that overruns are the normative, rather than the exception in the construction industry.

Overruns occur irrespective of the type of project, whether its buildings, roads, bridges, rails, or tunnels all suffer the same fate. Is it a small or mega project, size really does not matter either. Interestingly, this endemic phenomenon is not limited to developing countries, ineffective project delivery happens in Europe, Australia, America and everywhere.

In Bahrain the construction sector contributed 10.7% to growth of real GDP in 2014. After social and personal services sector, it was in the second rank of the fastest growth sectors in Bahrain non oil economy, the sector grew by 7.3% in 2014 from 2013, and this growth is mainly attributed to the prioritization of infrastructure and housing projects in the kingdom, Bahrain economic year book (2014).

Hence, why is cost overrun so frequent in the industry? Why has there not been much improvement in the reliability of initial cost estimates over the years? Surely the industry has become a lot better at managing projects. There are more measures now for accountability and cost control for project procurement. Information Technology for construction has also improved significantly; still many construction projects do not achieve their cost objectives.

1.2 Problem Statement

Construction industry defines Best Practice as “an optimal way currently observed by industry to achieve stated goal or objective” which is to achieve timely completion of projects within stipulated budget and required quality. For construction project management, this includes the ability to deliver projects predictably, consistently, and successfully to implement organizational strategies. Oddly though, no learning that would improve cost estimate accuracy seems to take place, according to Flyvbjerg et al. (2002) the size of overruns have not improved over the 70 years that they studied.

The cost estimated in the initial stage of a project is very important. It carries far more economic consequences than the limited decisions which can be

made later. Effective cost estimation is therefore so vital, because it can determine a project's financial fate, an out of control construction cost adds to investment pressure, increases construction cost, affects investment decision-making and wastes the national finance.

That being the case, it is not difficult to appreciate the need for a solution to this problem, for greater estimate reliability at all stages of a project and for greater assurance that initial cost expectations are met. But, as the discussions in this thesis will show, there is a need to understand the sources of these overruns on projects, as identifying and understanding the causes are usually the first step when addressing a problem; and then corrective actions can be taken.

1.3 Research Questions

The purpose of this study is to identify and assess various factors responsible for the cost overrun, which require serious attention to understand and address. Many researches have been devoted to the same purpose, however:

- Most of the previous studies are old, there is a need for a more up to date investigation to reflect any developments in recent years
- Most of the construction related studies were carried outside Bahrain
- Most of the studies focused on identifying factors that have the biggest influence on project costs and time. They did not discuss the relations between the causes, and the effect of each one on the final cost.

Two research questions are posed:

1. What risks are present during the delivery of construction projects in Bahrain that lead to significant project cost overruns?
2. What are the fundamental relationships between factors of cost overrun, and how does each factor correlate with the total amount of cost overrun?

1.4 Research Hypotheses

The idea of research hypotheses is to analyse two different variables: cost overrun factors and cost overruns. The review of literature revealed a list of factors that were critical for the achievement of the budget objective in construction projects. Based on that, a list of cost overrun groups and related factors were proposed. The following eight different hypotheses were formulated to be tested through the research.

Hypothesis 1: The presence of project management and contract administration related risk factors causes cost overruns in construction projects in Bahrain.

Hypothesis 2: The presence of design and documentation related risk factors causes cost overruns in construction projects in Bahrain.

Hypothesis 3: The presence of contractor's site management related risk factors causes cost overruns in construction projects in Bahrain.

Hypothesis 4: The presence of information and communication related risk factors causes cost overruns in construction projects in Bahrain.

Hypothesis 5: The presence of financial management related risk factors causes cost overruns in construction projects in Bahrain.

Hypothesis 6: The presence of human resources (workforce) related risk factors causes cost overruns in construction projects in Bahrain.

Hypothesis 7: The presence of non-human resource related risk factors causes cost overruns in construction projects in Bahrain.

Hypothesis 8: The presence of environmental risk factors causes cost overruns in construction projects in Bahrain.

1.5 Aim and Objectives of the Research

The aim of this research is to develop a model for quantifying and reducing the impact of cost overrun in Bahraini construction projects, through identifying and analyzing the risks associated with cost overrun factors in construction projects. This is expected to assist construction managers in: identifying and managing the critical cost risks associated with construction project activities, and allocating proper contingencies to accommodate these risks.

In order to achieve this aim, the following objectives were set:

- Identify and rank the main causes of cost overrun based on their overall effects on building construction projects in Bahrain.
- Assess the effects of cost overrun factors on project cost in Bahrain, and estimate causal relations among factors of cost overrun and cost performance.

1.6 Research Design and Methodology

This study is based on several research methods, which include literature review, historical records, interviews and survey. These methods acted as supplement to each other, which made the data collection more comprehensive and meaningful. Basically, literature review focused on gaining a better understanding of cost performance and causative factors affecting cost overrun in construction projects. Data was collected through a mixed methodology of quantitative method (Survey), and semi-qualitative methods (interviews and actual projects data). Data was analyzed with statistical tools in order to draw a conclusion and come with satisfactory recommendations.

1.6.1 Quantitative method

A questionnaire is prepared to find:

- frequency and severity of cost overrun experienced by practitioners,
- causes of cost overrun in their opinions, and
- factors that help to effectively control their projects.

1.6.2 Semi-qualitative method

Construction projects historical records are collected from some companies in Bahrain, and the concerned practitioners involved in projects are interviewed to:

- obtain the estimated and the actual cost.
- define the direct and indirect causes of construction projects cost overrun (if any).

1.6.3 Theoretical materials

Literature review uses a combination of sources to document and analyze what has been published on this topic through time. The theoretical materials contains: emails, reports, blue prints, drawings and designs, academic journal articles, conference proceedings, books (monographs and book chapters), and documentaries.

1.6.4 Methods of Analysis

The critical considerations when choosing research analysis methods often revolve around the type of problem under study, and how to maximize the chances of achieving aim and objectives of the study. Based on that the following analysis methods were set:

- Identify each factor risk zone by creating a risk map for cost overrun factors using severity and frequency indices.
- Rank of the factors affecting cost overrun in building construction in Bahrain using the importance index.
- Analyze and consolidate data using Principal Component and Factor Analysis (PCFA).
- Construct a model using the Partial Least Square and the Structural Equation Modelling.
- Validate the model with goodness of fit test.
- Discuss the results of the model development to reach conclusions and recommendations.

1.7 Thesis Outline

The thesis contains six chapters; Chapter 1 begins with a background of the research by providing the problem statement, aim and objectives, purpose of the study, and proposed methodology overview. Chapter 2 is a literature review of construction cost overrun, focuses on the scale of the problem, the causes of cost overrun and their impact on construction projects, and construction cost control measures. Chapter 3 contains the data collection methods; it includes the basic data of the respondents, contents of the questionnaire and actual projects data. Chapter 4 covers the various data analyses of the survey and actual project information. Chapter 5 discusses the development, assessment and validation of the model. Chapter 6 develops conclusions and recommendations that are derived from the research, and how the research objectives align with the findings. Finally, references and appendices including data collected are included.

CHAPTER TWO

REVIEW OF LITERATURE

2.1 Introduction

2.2 The Scale of Cost Overruns in Construction

2.3 Impact of Risk Factors in Construction Projects

2.4 Causes of Cost Overrun in Construction Existing Studies

2.5 Studies Conducted to Control Construction Cost Overrun

2.6 Literature Summary

CHAPTER TWO

REVIEW OF LITERATURE

2.1 Introduction

Construction industries play an important role in economic growth and development of any country. Recently it has been witnessed that a large number of construction projects are facing the problem of cost overrun, which is associated with some form of risk and uncertainty in the project. Such associated risks may be of various kinds which depend on many factors, due to the uniqueness, complexity and dynamic nature of the construction activities.

Risks are unknowns, the probability of the occurrence of which can be assessed by statistical means (risks are usually insurable). Uncertainties are unknowns as well, however the probability of the occurrence cannot be assessed (uncertainties are uninsurable). Uncertainty has two types of origins non-cognitive and cognitive. Non-cognitive uncertainty results from physical randomness. This type of uncertainty is normally dealt with by employing current statistical and probabilistic risk factors; leading to cost overrun in the delivery of construction projects. Cognitive types of uncertainty result from humans expressing subjective judgments. Uncertainty always exists in the modelling and project management of complex construction projects due to the model representing real systems, and is also attributed to humans who express risk in subjective terms.

Many practices in construction are actually driven by risk, such as bonding requirements, insurance, licensing and contract structure (general conditions,

payment terms, delivery method, selection mechanism, etc.). “The inherent uncertainties and risks are generally not only from the unique nature of the project, but also from the diversity of resources and activities” (CII, 1989). Risks are usually dynamic, that is, their characteristic, probability and impact can change during the project process. In addition, external factors can have a very significant effect on projects. Project success is usually measured by its schedule, budget and quality. Broadly, various risks can affect these three basic factors against the success of a project. In many industries including construction industry, risk, if left unmanaged, could have a negative impact on project budget and completion, and can prevent the project from meeting its overall objective. If people intend to use appropriate data to solve problems, make forecasts, develop strategies, and make decisions, then risk analysis is an essential control tool for project management and an important aid in decision-making process. This chapter review literatures concerning the major issues of cost overruns in construction, in order to recognize the risks of the phenomenon.

2.2 The Scale of Cost Overruns in Construction

The overview of the scale of the problem is necessarily as an introduction to the discussion on how and why cost overruns actually occur. Cost overruns occur when “the final cost of the project exceeds the initial estimate or budget” Avotos (1983). Some authors consider that the initial budget is created when the decision of commencing building is made such as Flyvbjerg et al. (2002) and Odeck (2004), some other authors uphold the idea that cost overrun should be found comparing the original contract value with the final cost of the project at the time of completion. The difference between cost overrun definitions; may cause a

difference in the magnitude of cost overruns that have been reported over the years. The magnitude of cost overruns may or may not vary depending on the size of the project, the location of the project and its type. Eden et al. (2004) pointed out that although there is more information about cost overruns in public projects, but this does not mean that there are no cost overruns in the private projects. Koushki et al. (2005) expressed that cost and time overruns are common not only in large projects, but also in small and simple projects. Gkritza and Labi (2008) supported this idea and they also added that cost overruns are also more common in long duration projects.

The statistics on construction cost overruns have been well documented in the literature. Morris (1990) conducted one of the first empirical studies with a narrow focus on cost overruns in large projects. He argues that cost overruns have become a regular feature of public sector projects. The average cost overrun found in this study is 82%. The Auditor General of Sweden (1994) is another study with a narrow focus on cost overruns involving transport projects. It covered 15 road and rail projects. The average capital cost overrun for the eight road projects was 86%, ranging between 2 and 182%, and for the seven rail projects this was 17%, ranging from minus 14% to plus 74%. The Auditor General of Western Australia (2012) assessed the management and performance of 20 capital intensive projects including sports venues, schools and hospitals. The expected cost of all the projects at the time was A\$6.157 billion, around A\$3.275 billion (114%) more than the total original approved budget estimates. The research by Flyvbjerg (1997) covers seven tunnel and bridge projects. They found an average construction cost overrun for the five completed projects of 14%, ranging from -10% to 33%. Nijkamp and Ubbels (1999) concluded that in most projects, cost overruns were common and the extent of cost underestimation varied between 0 and 20%.

The 2012 London Olympics bid was awarded at £2.4 billion in 2005. This was adjusted to about £9.3 billion in 2007 after significant scope changes. The project was eventually completed at £8.9 billion in 2010 (National Audit Office, 2012). The Sydney Opera House was synonymous with delays and overspending, finishing 1357% over budget and ten years late. The Opera House cost \$819M, but it was initially expected to cost about \$56M. The Big Dig (City of Boston's Central Artery project) was to cost US\$2.6 billion but was completed at US\$14.8 billion and 7 years late in 2006 (Gelinas, 2007).

In the Arabian Gulf and other Arab countries cost performance is considered to be one of the most serious and frequent problems in the construction industry (Faridi and Al-Sayegh, 2006), according to Assaf and Al-Hejji (2006) only 30% of construction projects in Saudi Arabia were completed within the specified cost and schedule. A study to investigate variations in the UAE public projects estimates, finds that the variations (positive or negative) between feasibility and contract cost, ranging between -28.5% and +36% (Al-Zarooni et al., 2000). Al-Momani (1996) reported that the actual cost of a construction project in Jordan usually exceeds the original contract price by 30% while changing of orders result in an 8.3% cost overrun.

2.3 Impact of Risk Factors in Construction Projects

One of the challenges facing the construction field is how to assess the risk of cost overruns and deliver projects within budget. These risks can cause losses that lead to increase in costs, time delays and lack of quality of projects. The purpose here is to identify the degree of impact of the risk factors on construction projects within

the construction industry. There are many approaches in literature that has examined the impact of risk in construction industries, notable are the following:

Hlaing et al. (2002) assessed dominating risk factors and identified the following as having the greatest impact on the financial aspects of the project: lack of financial resources of the contractor; financial stability of the client; and cost overruns due to delay.

Zou et al. (2007) used a systematic and holistic approach to identify risks, analyze the likelihood of occurrence, and impacts of risks associated with the development of construction projects from project stakeholder and life cycle perspectives. Postal questionnaire surveys were used to collect data. The research identified twenty major risk factors and found that these risks are mainly related to (in ranking) contractors, clients and designers, with few related to governmental bodies, subcontractors/suppliers and external issues. This research also found that these risks spread through the whole project life cycle and many risks occur at more than one phase, with the construction stage as the most risky phase, followed by the feasibility stage. It is concluded that clients, designers and governmental bodies must work cooperatively from the feasibility phase onwards to address potential risks in time, and contractors and subcontractors with robust construction and management knowledge must be employed early; to make sound preparation for carrying out safe, efficient and quality construction activities.

Chileshe and Fianko (2012) explored elements of risk by using an opinion survey approach to collect data from 103 professionals (clients, consultants and contractors) in the Ghanaian construction industry. Significant differences were found between the perceptions of these sub-groups regarding the likelihood of occurrence of threat risks in five categories: construction method; price inflation;

exceptional weather; ground conditions and site contamination; and poor communication among the project team. The contractors rated construction methods higher than did the clients, and they also rated exceptional weather higher than either the clients or the consultants. On the other hand, consultants rated price inflation little bit higher than the clients. Significant differences between the sub-groups were also found regarding the potential impact of the threat risk of price fluctuation. The consultants rated the price fluctuation threat risk higher than either the contractors or the clients. Nevertheless, there was complete agreement among the three samples (clients, contractors and consultants) regarding the ranking of the financial risk factor “delay in payment” and the economic risk factor “inflation”.

Cantarelli (2011) in Netherlands, discussed the impact of the implementation phase on cost overrun, and concluded that the longer the implementation phase the higher the cost overrun, and the pre-construction phase is significantly shorter than the construction phase but it has the highest influence on cost overruns.

Tipili et al. (2014) identified the likelihood of occurrence and degree of impact of the risk factors on construction projects within the Nigerian construction industry. A total of seventy eight questionnaires were sent to construction industry professionals which comprised of contractors, architects, quantity surveyors and engineers. The results were analyzed using descriptive statistic and analyses of variance (ANOVA) and subsequently exposure rating levels were determined, which enable the categorization of the probability impact score into low, medium and high levels. The study indicates a disparity of the ranking of the degree of occurrence and impact among the group. Based on composite of risk factors, cost related risk and time related risk were found to be the most likely to occur, and have the most impact on project, whereas environmental risk factor was found to be a low risk, since it had the least likelihood to occur and the least impact score.

Allahaim and Liu (2012) identified more than 90 causes of cost overrun; these causes were then grouped into five types based on how the causes impact on cost overrun (Table 2.1).

Table (2.1)
Causes of Cost Overrun Groups and Their Impact over Cost Overrun,
(Allahaim and Liu (2012))

S.N.	Ideal types	Impact over cost overrun
1	Market Volatility	Increase the volatility of input costs and thus chances of overrun.
2	Pressure for distorting estimation	Induces tendency for under estimation of costs.
3	Novelty	Increase the uncertainty of tasks and outcome, thus making planning and estimating difficult.
4	Complexity	Increases the complexity of coordination of parties and tasks, thus making it harder to meet preset targets.
5	Time pressure	Forcing project team to take shortcuts, crashing, concurrent tasks/projects, which are known to cause delays and overrun

Sharma and Goyal (2014) found several factors related to cost overrun and classified them into 11 groups according to the sources of the cost overrun. In the owner group, they found that the slow decision making factor has the highest impact over cost overrun of construction projects. In the contractor and consultant design group, the poor planning and scheduling, contract management, and frequent design changes have the highest frequency of occurrence respectively.

2.4 Causes of Cost Overrun in Construction Studies

Analysis of the reasons for project cost overrun of construction projects, is a necessary step for the improvement of any given cost estimating system. Many studies available in the literature are related to cost overrun in the construction projects worldwide. Some of the important case studies have been discussed here.

Arditi, et al. (1985) conducted a questionnaire survey to identify the factors causing cost overrun in public projects in Turkey, which was distributed to 146 public organizations and 327 contractors. The main idea of the authors was to understand the main reasons for cost overruns in projects that were constructed between the 1970s and the 1980s. The authors attributed the main causes of cost overruns to the increase of materials prices, fast growth of inflation, delays caused by changes in design specification and underestimation of cost at the moment of creating the budget of the project.

Okpala and Aniekwu (1988) wanted to understand the main causes of high costs of construction projects in Nigeria. Their study consisted of 450 questionnaires which were distributed to architects, contractors, engineers and owners in three different cities in Nigeria. The results were analyzed taking into account the different professions and the role of each respondent in the project. The authors concluded the main causes of cost overruns to shortage of materials, finance and payment for completed works, poor contract management, price fluctuations and fraudulent practices. In another study from Nigeria, conducted by Elinwa and Buba (1993), they found that the material cost, fraudulent practices, materials prices increment, high cost of machineries and poor planning were the most important causes of cost overruns. Mansfield et al. (1994) also carried out a questionnaire survey among key stakeholders (contractor, consultant and client organizations) in

Nigeria. It has been found that poor planning, shortage in materials, imported materials, changes in site conditions, design changes were the main causes of construction delays and cost overrun.

Kaming, et al. (1997) interviewed 31 project managers of high-rise construction projects in Indonesia, to find the main factors that influence time and cost overruns in this type of projects. The authors concluded the main causes of cost overruns as; lack of materials due to the inaccurate planning and estimating, increase of materials' cost, complexity of works, poor contractor management and unpredictable weather.

Frimpong, et al. (2002) sent 125 questionnaires to owners, contractors and consultants. The results pointed out that the main factors that caused overruns in Ghana were increase of cost of the materials, poor contractor management, problems with agencies' payments, poor technical performance, and inaccurate estimation of costs.

Koushki, et al. (2004) carried out a study about delays and cost increases in the construction of private residential projects in Kuwait, it aimed to find the reasons for delays and cost increases in the construction projects from the standpoint of the owner. The authors interviewed 450 owners of housing projects, which were chosen in 27 different districts in the country. The results showed that in the projects where owners spent more time on the pre-planning phase, experienced less cost-increases during the implementation phase of their residential projects. The authors concluded the main causes of cost overruns as increase of materials cost, poor contractor management, inaccurate estimating the degree of complexity of works and owners' financial constraints.

Jackson (2002) carried out a study on building construction projects in UK. The most important factors causing cost overrun were identified as poor project planning and management, unexpected ground condition, design development, lack of information, estimating method, time limit, commercial pressure, procurement route and external factors such as changes in pricing conditions, indices, inflation, statutory factors, market trends.

Flyvbjerg, et al. (2002) used a large sample, covers 258 transport infrastructure projects in 20 nations worth approximately US \$90 billion, to analyze the causes of cost escalation in these projects. The main aim was to identify which factors cause the cost escalation, to a lesser degree the reason behind why they cause it. He concluded that the cost escalation in transport infrastructure projects has a significant relation with the length of the implementation phase, the size of the project and the type of ownership. Flyvbjerg also studied cost overruns in megaprojects (Flyvbjerg, et al. (2003)). He defined megaprojects as those construction projects that cost more than one billion dollars. He found that the main reason for cost overruns in construction megaprojects is the underestimation of costs and length of delays during the planning stage. Other factors mentioned by the authors were: the absence of a plan that takes into account not only the geological risks, but also the design and specification changes that the project can have in the construction phase, the lack of knowledge about changes in exchange rates between currencies and safety, and environmental risks during the execution.

Creedy (2006) studied the main causes of cost overrun in Australia Highway project. He identified 37 initial risk factors of 231 highway projects which had reported significant cost overrun, the top 10 factors in order of their magnitude of occurrence were: Design/Project scope change, contract tender price higher than original estimate, drainage design scope change, quality increased measure,

materials design scope change, latent condition, remove and replace unsuitable materials, environmental issue design scope change, constructability under traffic and service relocation costs.

Le-Hoai et al, (2008) adopted field survey methodology to uncover factors influencing delay and cost overruns arising during construction stage. To identify the delay and cost overrun factors in Vietnamese construction market, literature reviews, case analysis published on newspaper, and discussion with practitioners of all parties involved in construction industry were carried out. After that, a pilot questionnaire was prepared. The designed questionnaire (285 questionnaires) was randomly distributed to professionals of major projects of three principal construction parties (owner, consultant and contractor). The authors found the following causes of cost overruns: poor site management and supervision, contractor poor project management assistance, consultant financial difficulties of owner, financial difficulties of contractor, contractor design changes, project unforeseen site conditions, external slow payment of completed works, owner inaccurate estimates and consultant poor contract management.

There were many other studies such as, Kaliba, et al. (2008) who found the major causes of cost overruns in road construction projects in Zambia. The authors based their study on questionnaires, interviews and some case studies of construction projects. Unlike some studies mentioned before, the main reason for cost overruns in Zambia, according to the results, was the inclement weather. However, other reasons were similar to the studies carried out in other developing countries; for example, inflation, changes in the scope of the project, delays in schedule and lack of managerial and technical knowledge. Another recent study was conducted by Al Najaar (2008) in Gaza Strip who wanted to find the main factors that caused cost overruns in construction projects in Gaza. With the use of 120 questionnaires,

which were distributed among owners, consultants and contractors, the authors concluded that the main factors were the increase of cost of materials, shortages in materials and lack of experience and knowledge of workers. It is important to mention that this study pointed out an interesting factor that causes overspending in a project and was not been mentioned by other authors. This factor is the lack of leadership skills in the managerial positions of the project, which not only cause a poor management performance, but can also have an impact on the relationship with the owner or client. A study was conducted in Saudi Arabia to investigate the cost overrun in construction projects in order to identify the causes of cost overrun, and their importance through a questionnaire survey. The study concluded that the top cost overrun causes in construction projects in Saudi Arabia from contractors' perspective are: bid award for lowest price, frequent changes in design, improper planning, long period between design and time of implementation, and payments delay (Alghonamy, 2015). Table 2.2, shows a Summary of cost overrun factors in construction projects found by past studies in different countries.

Table (2.2)

Summary of Causes of Cost Overrun from Previous Studies

S.N.	Author	Location	Year of Publishing	Main Causes of Cost Overruns
1	Arditi, et al.	Turkey	1985	• Increase of materials prices • Fast growth of inflation
2	Okpala and Aniekwu	Nigeria	1988	• Shortage of materials • Finance and payment for completed works • Poor contract management
3	Elinwa and Buba	Nigeria	1993	• Material cost • Fraudulent practices
4	Kaming, et al.	Indonesia	1997	• Lack of materials due to the inaccurate planning and estimating • Increase of materials' cost • Complexity of work

5	Flyvbjerg, et al.	Denmark	2002	• The length and size of the project
6	Jackson	United Kingdom	2002	• Identified as poor project planning and Management • Unexpected ground condition • Design development
7	Flyvbjerg, et al.	Mega Projects from Twenty countries	2003	• lack of knowledge of the project and its complexity • Inaccurate estimation of costs, the absence of a plan that copes with the changes in design and execution requirements • The lack of managerial strategies
8	Frimpong, et al.	Ghana	2003	• Increase of cost of the materials • Poor contractor management • Problems with agencies' payments
9	Koushki, et al.	Kuwait	2004	• Increase of materials prices • Poor contractor management • Inaccurate estimating Degree of complexity of works
10	Creedy	Australia	2006	• Design/Project scope change • Contract tender price higher than original estimate
11	Le-Hoai et al.	Vietnam	2008	• Poor site management and supervision • Contractor Poor project management assistance • Consultant Financial difficulties of owner
12	Kaliba, et al.	Zambia	2008	• Inclement weather
13	Al Najaar	Gaza Strip	2008	• Increase of cost and shortages in materials • lack of experience and knowledge of workers
14	Cantarelli	Netherlands	2011	• Pre-construction phase (Planning) • Implementation phase length
15	Allahaim and Liu	Australia	2012	• Increases the complexity of coordination of Parties and tasks, thus making it harder to meet Preset

				targets
16	Alghonamy	Saudi Arabia	2015	• Frequent changes in design • Bid award for lowest price • Improper planning

2.5 Studies Conducted to Control Construction Cost Overrun

There are many measures which are found from the researchers' studies to control the construction costs or to overcome the problems of cost overruns. The researchers have their own opinions on how to solve the problems. Some are presented below.

Asmi et al. (2012) studied controlling of cost overrun factors in construction projects in Malaysia in order to explore the major contributors of cost overrun, and propose mitigation measures to control cost overrun factors. This study adopted qualitative research method using semi-structured interview session with experienced personnel's involved in handling large construction projects. The respondents were requested to evaluate 8 categories of cost overrun factors, and classify mitigation measures for cost control according to suitable implementation strategies. The categories are contractor's site management, project management and contract administration, design and documentation, labour related factors, material and machinery, financial management, information and communication and external factors. The study resulted in 15 proposals on mitigation measures and these were classified in three implementation strategies: pro-active strategy, organizational strategy and fluid measures which could be adopted as proactive as well as reactive measure.

Khandekar et al. (2015) analyzed construction project cost overrun by statistical method, the analysis was conducted in India to evaluate causes of cost overrun impact on project cost. Questionnaire survey was carried out among the two major participants namely government and private representative from construction projects firms. Total 50 sets of questionnaires were distributed, 37 responses were received. The Relative Important Index was used for ranking of causes of cost overrun in the construction project. It was observed that the factors for cost overrun are the material shortage, shortage of labour, late delivery of materials and equipment, unavailability of competent staff, low productivity level of labours, quality of equipment and raw material. For effective and efficient cost control of construction projects the authors recommended that material management, resource planning and management, and proper financial management could be adopted.

Ibadov and Kulejewski (2014) assessed the construction project risks with the use of fuzzy sets theory. The study objectives were to identify the risk factors affecting construction projects, to rank the risk factors due to their degree of significance and to evaluate the threat to the construction project caused by a given risk. Data from 133 questionnaires was collected, and then analyzed to determine the degree of significance of each risk factor, taking their impact on the execution parameters of the construction project into account. The analysis started by determining the degree of significance for each factor, then using the theory of fuzzy sets to describe and formalize the degree of significance of factors, and after that transforming a fuzzy value into precisely defined real number or what we call (defuzzification), by using the center of gravity method. The authors found that the majority of respondents believe that; the degree of significance given to the increase of prices of construction materials risk factor is only slightly important in

terms of impact on the cost of the project. This could mean that the construction materials market is stable or, during the period of project execution, the probability of price change was low. Use of the fuzzy sets theory for ranking and the modelling of individual risk factors may facilitate the planner/scheduler to determine a degree of influence for each factor against the performance parameters of a construction project.

Han (2005) estimated cost overrun risk in international projects by using fuzzy set theory as well. The author proposed a methodology which can be used by bidders to quantify cost overrun risk so that a realistic risk premium may be determined. The data was collected through conducting of a survey. The data was used to model cost overrun risk by influence diagramming method and determine the membership functions for fuzzy variables. The study resulted in developing of a software to carry out fuzzy number calculations as a decision support tool for decision makers during the bidding stage of international projects, and a risk model which can help construction professionals in the identification of risks in international projects.

Mahamid and Dmaidi (2013) identified the risk map for factors affecting cost overrun in building construction projects from the consultants' perspective. Field survey was conducted to calculate the frequency index and severity index for each factor. And the risk map for cost overrun factors was developed.

Diab and Varma (2012) used risk assessment to improve highway construction project performance. The researchers objective was to analyze and evaluate the different risk drivers in highway construction projects in the US, and to determine the impact of using the risk assessment on construction cost and schedule performance. The data were collected from previous studies and a survey

questionnaire which was developed to collect quantitative and qualitative data from completed transportation construction projects in the US. Eight hypotheses were developed. Each hypothesis tests the null hypothesis (H_0) that there is no dependency correlation, while the alternative hypothesis (H_a) tests that there is dependency correlation. To test the Hypothesis the data was examined and analyzed using frequency tables and chi-square dependency correlations. The results showed that majority of respondents (80%) indicated that risk assessment was important for highway construction projects. The use of risk assessment improves project and construction management practices by helping transportation professionals to focus on the potential significant risk drivers in the planning phase. This practice would also improve cost and schedule performance through better management of potential risks.

Chileshe and Fianko (2011) investigated the perceptions of threat risk frequency and impact on construction projects in Ghana. The study aimed to investigate perceptions of the likelihood of occurrence and severity of risk in construction projects by conducting a survey, which was divided into five sections: general demographics of the respondents, risk assessment and management processes, awareness of risk assessment and management processes, barriers to the use of risk management techniques and critical success factors. The results of the survey were then analyzed to examine the frequency with which the risk factors were likely to occur, the severity of their impact on construction projects, and any differences in the perceptions of the respondent sub-groups (contractors, clients and consultants) regarding the likelihood of occurrence and severity of risk impact on construction projects. They concluded that the ranking and importance of different factors vary between developed and developing economies. The different risk variable rankings between developing countries and developed countries illustrate how the country

context sheds new theoretical light on the interpretation and importance of risk variables.

Arcila (2012) conducted a study to identify the main causes of cost overruns in UK construction projects, and also to identify the Critical Success Factors (CSFs) that can help to improve the possibility of avoiding or preventing cost overruns in construction projects. The data was collected through interview questions and two hypothesis were established; one about the lack of certain CSFs causes cost overruns in construction projects; the second one about the presence of certain CSFs prevents cost overruns in construction projects. The initial idea of this research was to analyze six different projects from the same company, three projects with good cost performance and three projects with poor cost performance. Four different aspects were tested in the case studies design to assess if the design represents a logical flow or not, those were construct validity, internal validity, external validity and reliability. The results showed that when the poor cost performance projects were compared against each other, Hypothesis 1 could be confirmed based on the fact that all of those projects lacked some factors that caused cost overruns. On the other hand, a comparison of the good cost performance projects helped in confirming Hypothesis 2; due to the fact that these projects had some CSFs that helped prevent cost overruns. The comparison between good and poor cost performance projects also helped in confirming Hypothesis 1 and 2, because it showed that the presence of some factors in good cost performance projects, which were not present in poor cost performance projects, helped avoid cost overruns on good cost performance projects and caused overruns on poor cost performance projects.

2.6 Literature Summary

This chapter has presented an overview of the scale of the problem of construction cost overruns as a global phenomenon that is indifferent to the size of the project under construction, the geographical location, duration or type of project. The literature as well aimed to identify the causes of construction cost overrun, a total of 45 factors of cost overrun were selected from the literature and will be investigated in the following chapters, These factors were classified into 8 groups, named as Contractor's Site Management Related Factors, Design and Documentation Related Factors, Financial Management Related Factors, Information and Communication Related Factors, Human Resource Related Factors, Non-human Resource Related Factors, Project Management and Contract Administration, and Environmental Related Factors (Table 2.3).

Table (2.3)

Selected Groups and Factors of Cost Overrun

a) Project Management and Contract Administration Related Factors:	
1	Poor project management
2	Change in the scope of the project
3	Delays in decisions making
4	Inaccurate quantity take-off
b) Design and Documentation Related Factors:	
5	Frequent design changes
6	Mistakes and errors in design
7	Incomplete design at the time of tender
8	Poor design and delays in design
9	Delay preparation and approval of drawings
c) Contractor's Site Management Related Factors:	
10	Poor site management and supervision
11	Incompetent subcontractors
12	Schedule delay
13	Inadequate planning and scheduling
14	Lack of experience
15	Inaccurate time and cost estimates
16	Mistakes during construction
17	Inadequate monitoring and control
d) Financial Management Related Factors:	
18	Cash flow
19	Poor financial control on site
20	Financial difficulties of owner
21	Delay in progress payment by owner
22	Delay payment to supplier /subcontractor
23	Contractual claims
e) Information and Communication Related Factors:	
24	Lack of coordination between parties
25	Slow information flow between parties
26	Lack of communication between parties
f) Human Resources (Workforce) Related Factors:	
27	Labour productivity
28	Shortage of site workers

29	Shortage of technical personnel
30	High cost of labour
31	Labour absenteeism
g) Non-Human Resource Related Factors:	
32	Fluctuation of prices of materials
33	Shortages of materials
34	Equipment availability and failure
35	Late delivery of materials and equipment
h) Environmental Related Factors:	
36	Level and number of competitors
37	economic instability
38	effects of weather
39	government policies
40	inadequate production of raw materials by the country
41	monopoly by suppliers
42	number of projects going at the same time
43	political situation
44	project location
45	Social and cultural impacts

CHAPTER THREE

DATA COLLECTION

3.1 Data of Construction Projects

3.2 Survey Sample Size

3.3 Questionnaire Design and Distribution

3.4 Actual Projects Data Collection

3.5 Interviews Data Collection

CHAPTER THREE

DATA COLLECTION

3.1 Data of Construction Projects

The research required a sample of construction projects that was appropriate for this area of research. The data sample was required to be large enough to allow statistical analyses of cost overrun factors and project costs. Data on actual costs in construction projects were relatively difficult to obtain. One reason is that such data is time consuming to generate. For private projects, such data are often classified as commercial-in-confidence and thus kept secret from the hands of competitors. Unfortunately, this also tends to keep private project data out of the reach of researchers. For both public and private projects, data on actual costs have been held confidential by project owners, because actual costs reveal substantial cost escalation that was often considered an embarrassment to project owners. For public sector projects, funding and accounting procedures often made it difficult to keep track of the multiple and complex changes that occur in total project costs over time. For large public projects, reconstructing their actual total costs would typically entail long and difficult archival work and complex accounting. Therefore a survey was considered to support and enhance the limited amount of actual projects data that has been obtained.

3.2 Survey Sample Size

Sampling is the process of selecting representative units of a population for the study in research investigation (Wood and Haber, 1998). A sample is a small proportion of a population selected for observation and analysis. The sample was selected randomly from contracting companies, consultants and clients. Statistical equation was used in order to calculate the sample size. Equation 3.1 was used to determine the sample size of the unlimited population. (Creative Research System, 2001)

$$SS = \frac{Z^2 * P * (1-P)}{C^2} \quad (3.1)$$

Where: SS is the sample size for infinite population, Z is the normal standard variate (taken as 1.96 for 95% confidence level), P is the percentage of picking a choice, expressed as a decimal (0.50 is used for sample size needed), C is the margin of error (9%) according to the Creative Research System (2001). Thus,

$$SS = \frac{1.96^2 * 0.5 * (1-0.5)}{0.09^2} = 118.57 \approx 119$$

Correction for finite population is given by:

$$SS_{\text{finite}} = \frac{SS}{1 + \frac{SS - 1}{POP}} \quad (3.2)$$

Where POP is the population, it was taken equal to 35 representing the proposed grades of contracting companies in Bahrain (grade A, B and C):

$$SS_{finite} = \frac{119}{1 + \frac{119-1}{35}} = 27.2 \approx 27$$

The total number of questionnaires distributed to engineers working in contracting companies was 53 and the total number returned was 42.

The population of consultants is 23 registered offices in Bahrain, thus the sample size is:

$$SS_{finite2} = \frac{119}{1 + \frac{119-1}{23}} = 19.4 \approx 19$$

The total number of questionnaires distributed to engineers working in consultant offices was 37 and the total number returned was 21. Since the construction project client in Bahrain could be any one, calculation for sample size is difficult. However, thirteen questionnaires were distributed, but only eleven were returned.

3.3 Questionnaire Design and Distribution

For the study purpose primary data was used. The primary data was collected by conducting a survey to obtain statistically useful information about cost overrun in construction projects in Bahrain. Survey material was first examined and identified through a relevant literature review and sought advice from experienced construction practitioners. Survey was performed by questionnaire method, where a list of different types of closed ended and open ended questions, were answered by professional practitioners engineers, in public and private sectors in Bahrain. Questionnaire was distributed in printed and electronic form, through mail, fax and by hand.

The questionnaire contains three sections. The first section aims to collect the background information of the respondents, e.g. their age, gender, position, education, work experience and professional background. The second section includes the respondents' opinion on the risk factor in terms of its likelihood and impact to overall construction project success, A five-point scale of 1 to 5 was considered for evaluating each factor, these numerical impact and likelihood values are assigned to the respondents' rating: (1: Not significant; 2: Slightly significant; 3: Moderate; 4: Very significant; 5: Extremely significant). The third section includes the respondent's additions and comments on the risk of cost overrun for construction projects in Bahrain. In the structured part of the questionnaire, 45 causes of cost overrun are listed in 8 respective categories (Appendix 1).

The questionnaire was sent out to a total of 103 engineers of different levels of work experiences, asking for their contribution in this study; a total of 74 engineers were willing to participate and have filled the questionnaire. The

response rate was 71.8%, including 42 contractors, 21 consultants and 11 clients (Figure 3.1).

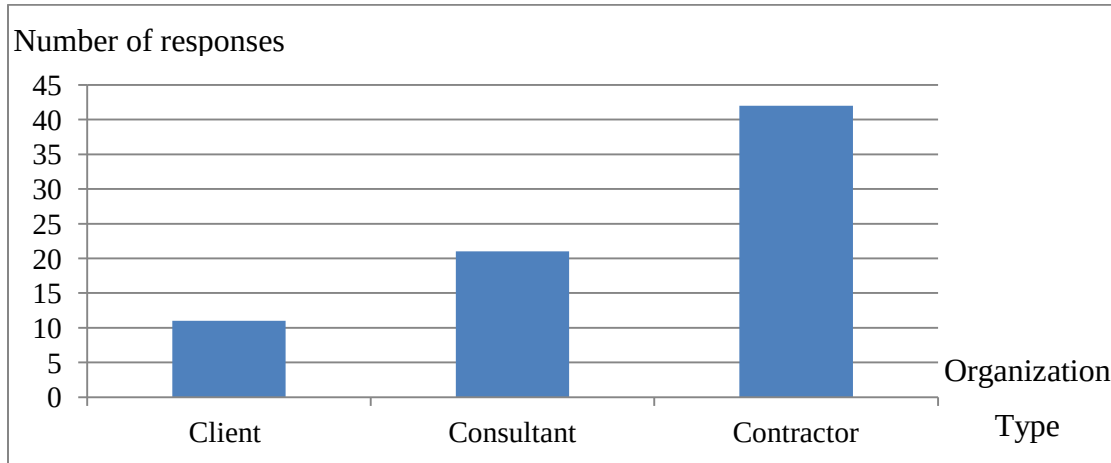
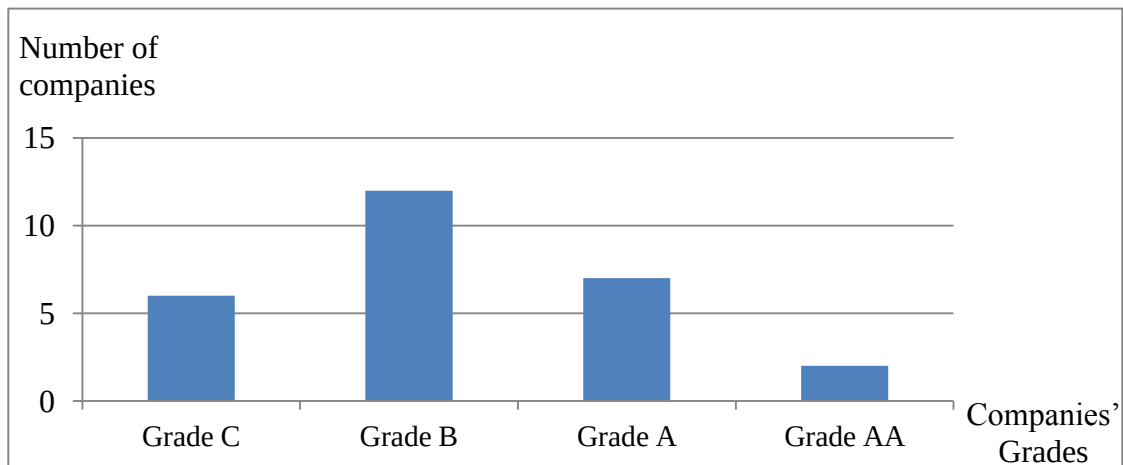


Figure (3.1)

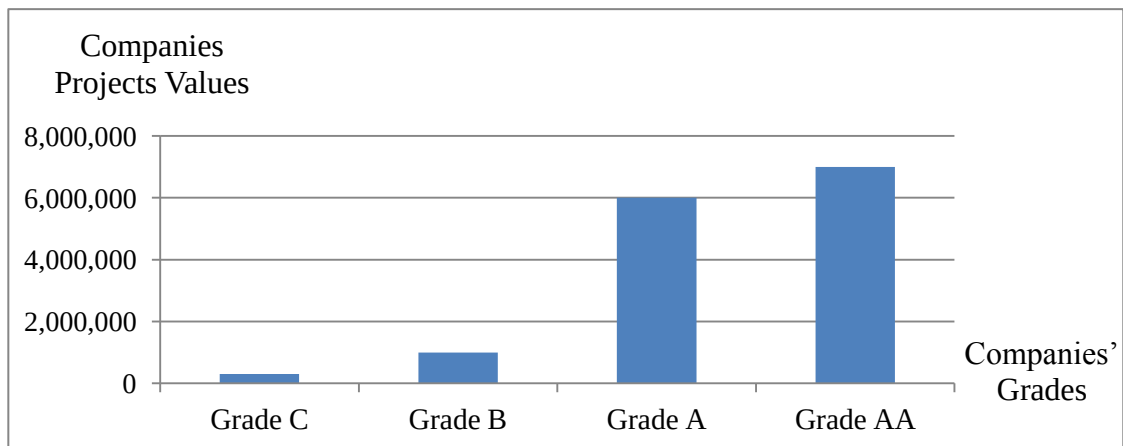
The Distribution of the Respondents According to their Organization Type

According to Bahrain Ministry of Works, each construction company is graded based on the undertaken projects' value, if projects value fall below BHD 300,000 it is considered a grade C, between BHD300,000 – BHD1,000,000 is grade B, up to BHD 6,000,000 is grade A and above that is considered grade AA. The 74 engineers came from 27 different companies, 6 were grade C, 12 companies from grade B, 7 companies were grade A and only 2 were grade AA. (Figure 3.2)



(Figure 3.2) (A)

The Distribution of the Respondents According to their Companies' Grades



(Figure 3.2) (B)

The Distribution of the Respondents According to their Companies' Projects Values

Figure 3.3, shows the distribution of the respondents according to their level of qualification and experience in building construction. Most of respondents have the Bachelor's degree, and experience of more than 10 years in building construction. It would have been favoured if the percentage

of respondents whose experiences were 10 years or more could be increased. However it has to be mentioned that a large number of young practitioners have been graduated in recent years to meet the increasing demand in construction industry, and they have got high positions in their organizations. This is the most important factor which makes the results of the questionnaire reliable

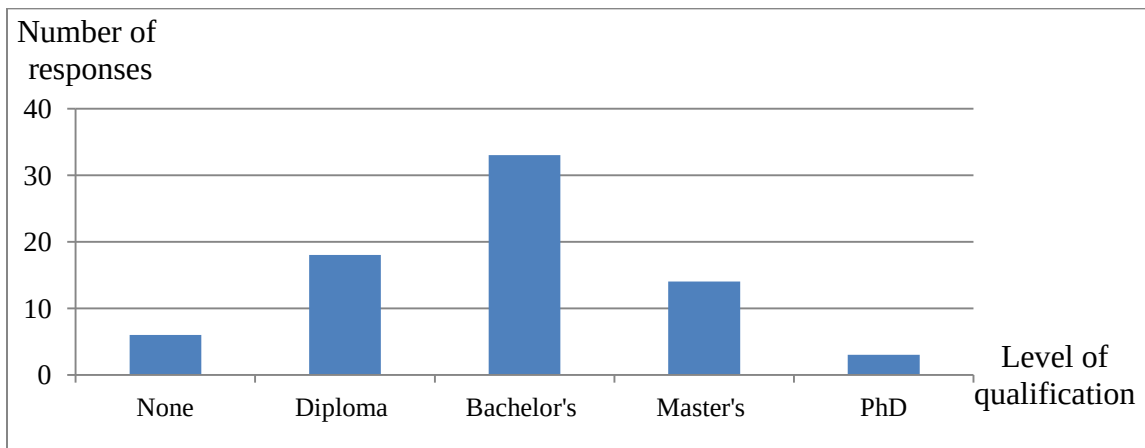


Figure (3.3) (A)

The Distribution of the Respondents According to their Level of Qualification

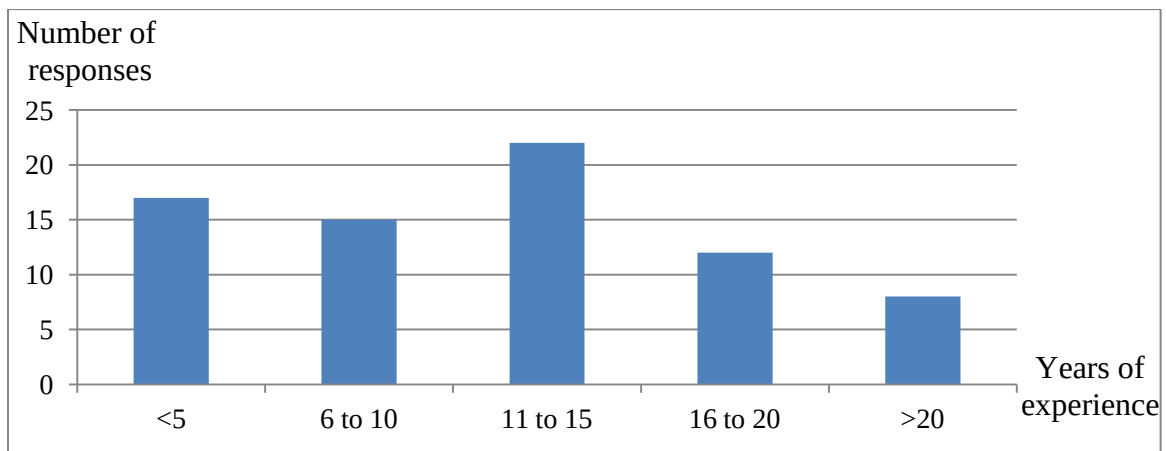


Figure (3.3) (B)

The Distribution of the Respondents According to their Work Experience

The analysis of the participants' responses regarding the cost overrun in construction projects reveals that most of respondents indicated that the average cost overrun in building construction projects that they have experienced fall between 15% and 35% of the original estimated cost of a project. More illustration is shown in Figure 3.4.

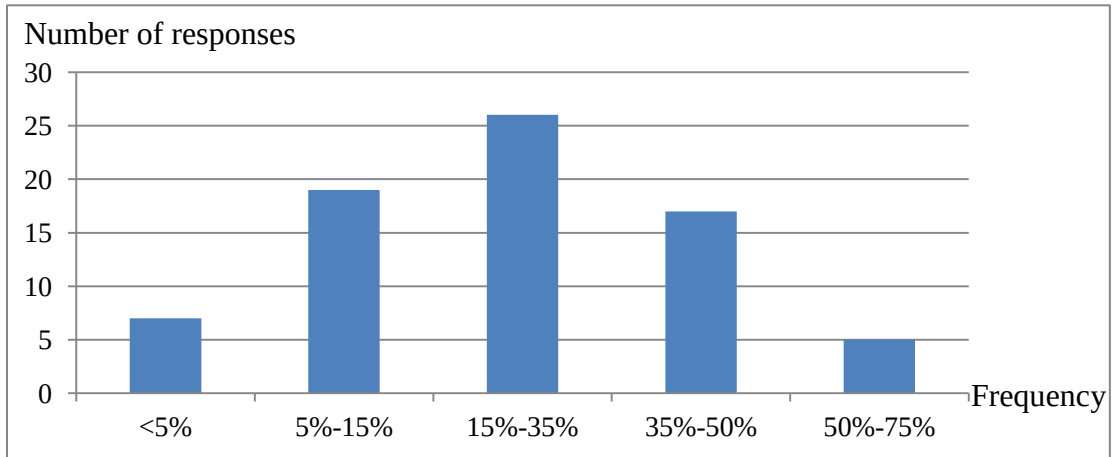


Figure (3.4)

The Frequency of Cost Overrun Experienced by Participants

3.4 Actual Projects Data Collection

Cost overrun occurs when the actual cost for the project is more than the estimated cost. In this part, forty building construction projects were selected to evaluate. The projects were selected based on the types of the projects and availability of estimated and actual cost data. Selected project types are new construction, reconstruction and improvement of existing facilities.

Summary of each project estimated cost, actual cost and causes of cost overrun is shown in Appendix 2.

Figure 3.5: shows the different types of projects that were analyzed in the sample of actual data. Forty seven and half percent of the projects were mainly houses, and residential buildings, other projects including general (utility buildings... etc.) 20%, health (hospitals and clinics) 10%, educational (schools, kindergartens and institutes, etc.) 17.5%, and sports and youth projects 5%.

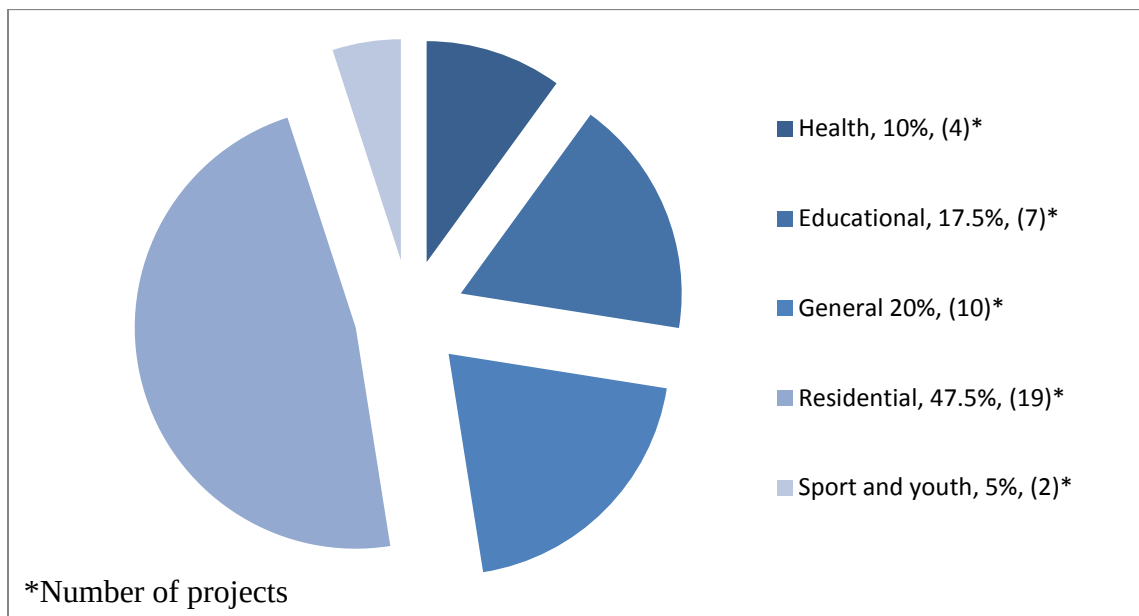


Figure (3.5)
Projects Types' Percentages in Data Sample

The data was collected from 40 construction projects. In these projects there were 31 projects categorized as open tender contracts; and this constituted 77.5% of the total projects analyzed, and 9 (22.5%) were negotiated price contracts. (Figure 3.6)

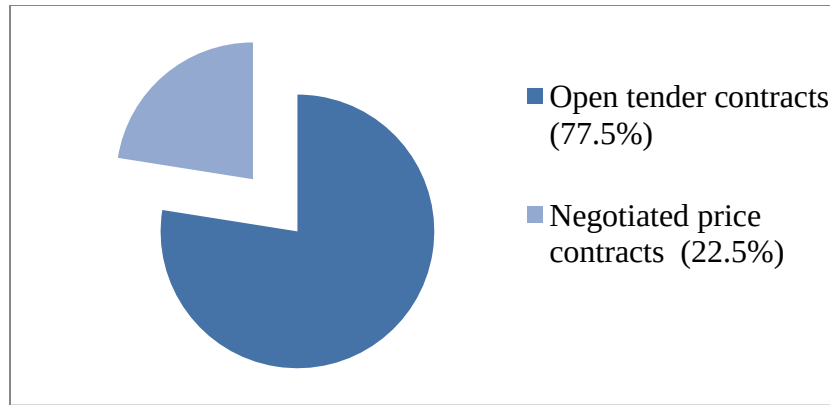


Figure (3.6)
Construction Project Contract Type

3.5 Interviews Data Collection

Some interviews were carried out during the second stage of the actual data collection, to determine the causes that led to a difference in the actual and estimated costs. Selection of interviewees was based on candidate years of experience and amount of participation in the concerned projects. The questions were designed to deal with background details about the project, and to get enough information to determine the exact causes responsible for the cost overrun. In total, thirteen practitioners with different job positions (Figure 3.6) accepted to have a one hour semi-structured interview. The interviewees were mostly contractors and four of them were working in consultancy companies. All of the interviewees had more than 10 years of experience.

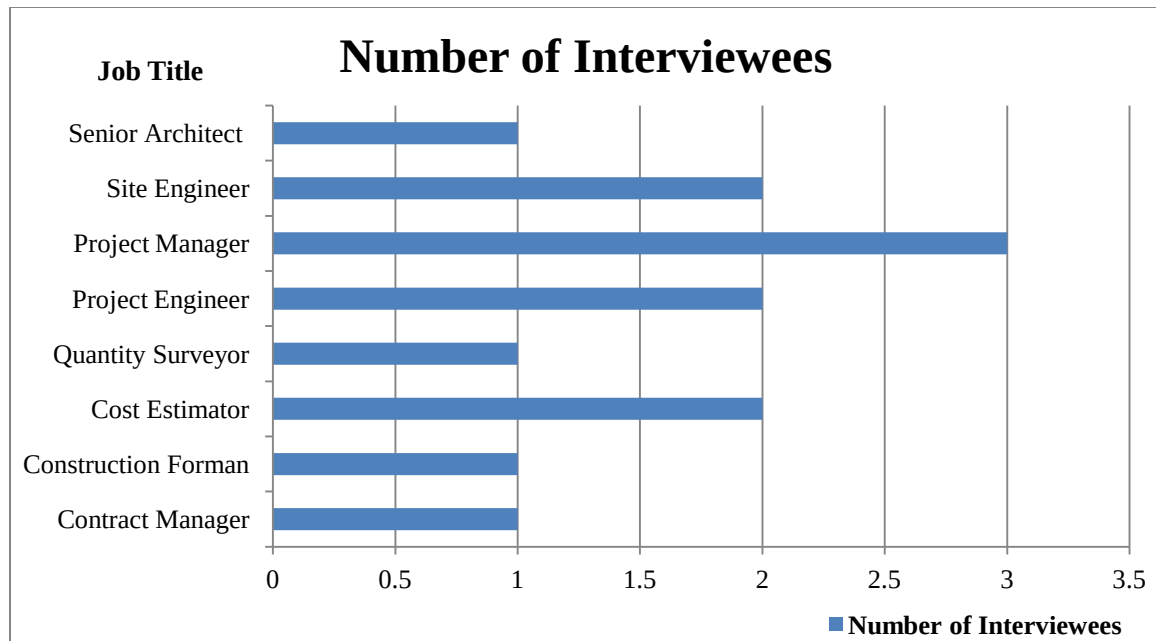


Figure (3.7)

The Interviewees Numbers and Job Titles

The interview had two different sections that helped in data collection process. Firstly, the interviewees were asked some background information about the project. In this section, they talked about the main features of the project and explained the problems they encountered during the implementation phase of the project. The next section of the interview had some questions about the cost performance of the project; where interviewees were asked to talk about the main causes of poor cost performance, the important issues in the project execution, and what are the main problems encountered in the project that caused cost overrun. This information was used to determine “main causes of cost overruns in each actual project” (Figure 3.8).

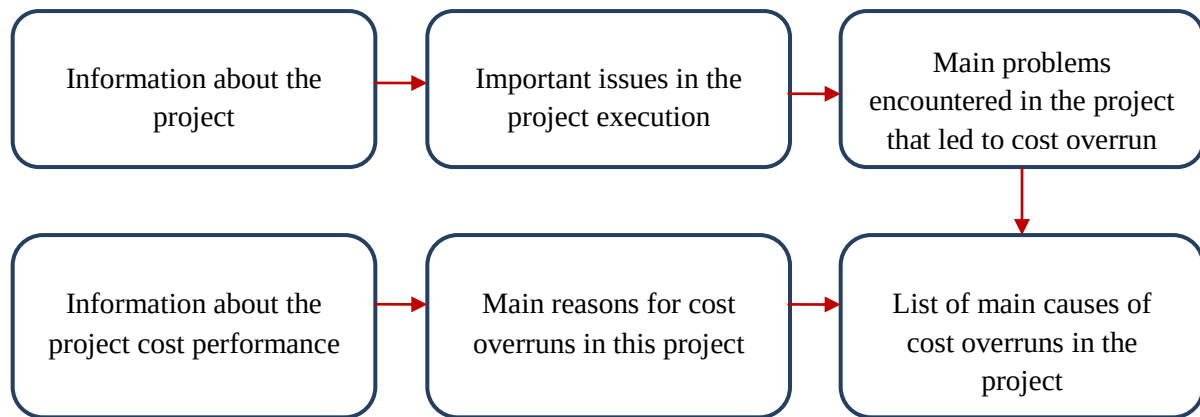


Figure (3.8)

Process Used to Determine Main Causes of Cost Overruns in Actual Construction Projects

CHAPTER FOUR

DATA ANALYSIS

4.1 Introduction

4.2 Analysis Formulae

4.3 Risk Mapping

4.4 Ranking of Cost Overrun Factors and Groups

4.5 Analysis of Results by Data Source

4.6 Survey Based Cost Overrun Factors Analysis

4.7 Survey and Actual Based Data Rankings Analysis

CHAPTER FOUR

DATA ANALYSIS

4.1 Introduction

This part of the research is oriented to investigate the causes of cost overrun and their effects. Cost overrun in building construction projects are caused by many factors. Each factor of cost overrun has a different rate of occurrences, and impact on the final cost. To assess these risks efficiently, it is important to rank them based on their effects on building construction projects.

4.2 Analysis Formulae

The procedure used in analyzing the data aimed at establishing the frequency and severity indices of the various factors responsible for project cost overruns, which are calculated according to the following formulae:

4.2.1 Survey data

$$\text{Survey based frequency index (FI (i))} = \frac{\sum_{w_i=1}^5 w_i \times f_i}{5 \times N} \quad (4.1)$$

$$\text{Survey based severity index (SI (i))} = \frac{\sum_{w_i=1}^5 w_i \times S_i}{5 \times N} \quad (4.2)$$

Where:

W_i is a constant representing weight assigned to factor i (ranges from 5 for Extremely Significant/Continual to 1 for Not Significant).

FI(i)/SI(i) is the frequency/severity of each response for each factor(i).

N is the total number of responses.

4.2.2 Actual projects data

Frequency index and severity index are calculated for each factor according to the following formulae:

Actual Data Based Frequency Index (FI'(i)) = (Factor incidences across projects / Total number of projects) (4.3)

Actual Data Based Severity Index (SI'(i)): The severity of cost overrun factor on an actual project cost was based on projects contingency and the percentage of cost overrun in each project. According to the project contingency guidelines, contingency for \$5 million and greater worth of project is 3% to 7% of project total value, for projects less than \$5 millions, contingency is ranging from 5% to 10% (ABOR-CCG, 2002). In Bahrain most of construction projects contingency falls between 5% to 10%; therefore severity was defined based on cost overrun as follow:

$$\text{Percentage of cost overrun} = \frac{\text{Estimated project cost} - \text{Actual project cost}}{\text{Actual project cost}} \quad (4.4)$$

Table 4.1, shows the degree of severity for the actual projects data, based on the percentage of cost overrun established using equation 4.4. The degree of severity was organized in the form of a priority scaling using likert scale (1 = not significant, 2 = slightly significant, 3 = moderately significant, 4 = very significant, and 5 = extremely significant).

Table (4.1)
Actual Projects Degree of Severity

Percentage of Cost Overrun (%)	Degree of Severity
Less than 2%	1
2-5%	2
5-7%	3
7-10%	4
Greater than 10%	5
*Bahrain Construction Project Contingency 5-10%	

Actual Data Based Severity Index (SI(i)) = (Sum of factor (i) severities/number of projects experienced this factor) (4.5)

4.3 Risk Mapping

Develop a risk map for factors affecting cost overrun in building construction projects in Bahrain. The objectives of this part include identify the risk map for cost overrun factors and rank the factors affecting cost overrun in Bahrain.

4.3.1 Risk map development

Frequency and severity are the two primary characteristics used to screen risks and separate them into minor risks that do not require further management attention and significant risks that require management attention and possibly quantitative analysis. Various methods have been developed to help classify risks according to their seriousness. One common method is to develop a two-dimensioned matrix that classifies risks into three categories based on the combined effects of their frequency and

severity. It helps in classifying risks into one of five states of likelihood and into five states of consequence (minimal through unacceptable). Table 4.2 shows the scale used to determine the severity and frequency levels for cost overrun factors. (The U.S. Federal Highway Administration Office of International Programs, 2007).

Table (4.2)
Scale Used to Identify Severity and Frequency Level and Associated Effect on Cost (FHWA, 2007)

Level	Index value (Scale)	Severity	Frequency	Effect on Cost
A	$\leq 20\%$	very low (VL)	very low (VL)	Minimal or no impact
B	20% - 40%	low (L)	low (L)	<5%
C	40% - 60%	moderate (M)	moderate (M)	5-7%
D	60% - 80%	high (H)	high (H)	7-10%
E	80% - 100%	very high (VH)	very high (VH)	>10%

4.3.2 Standard risk map

Standard risk map is used to determine the risk zone for each cost overrun factor. The map is a five by five matrix with severity ranging from very low (VL) to very high (VH) on the horizontal axis and frequency with the same range on the vertical axis. Three zones are presented in the map, that classifies a risk as either “high” (red), “moderate” (yellow), or “low” (green), as shown in Figure 4.1 The zones have the following characteristics:

1. **Green zone:** Minimum impact. Minimum oversight needed to ensure risk remains low.

2. **Yellow zone:** It can cause some disruption. Different approach may be required to reduce it and additional management attention may be needed to control it.
3. **Red zone:** Unacceptable, major disruption likely. Different approach required with priority management attention.

Frequency	VH					
	H					
	M					
	L					
	VL					
		VL	L	M	H	VH
Severity						

Figure (4.1)
Zones of the Risk Map (FHWA, 2007)

The risk map used a 45 factors causing cost overrun in construction projects, which were defined through a detailed literature review (Chapter 2). These causes were classified into 8 groups according to the sources of cost overrun. Frequency and severity of each factor were categorized according to Table 4.3.

Table (4.3)
Risk Mapping for Survey Cost Overrun Factors

Factors		Severity Index	Severity level	Frequency index	Frequency level	Survey Map zone
a) Project Management and Contract Administration Related Factors:						
1	Poor project management	72.43	H	73.24	H	Red
2	Change in the scope of the project	59.19	M	57.84	M	Yellow
3	Delays in decisions making	51.62	M	51.43	M	Yellow
4	Inaccurate quantity take-off	71.08	H	74.59	H	Red
b) Design and documentation Related Factors:						
5	Frequent design changes	88.38	VH	86.49	VH	Red
6	Mistakes and errors in design	78.92	H	44.59	M	Yellow
7	Incomplete design at the time of tender	67.03	H	67.03	H	Red
8	Poor design and delays in design	75.95	H	45.95	M	Yellow
9	Delay preparation and approval of drawings	63.51	H	73.78	H	Red
c) Contractor's site Management Related Factors:						
10	Poor site management and supervision	69.73	H	74.32	H	Red
11	Incompetent subcontractors	80.54	VH	59.46	M	Red
12	Schedule delay	87.57	VH	77.03	H	Red
13	Inadequate planning and scheduling	78.65	H	62.43	H	Red
14	Lack of experience	79.73	H	55.14	M	Yellow
15	Inaccurate time and cost estimates	92.97	VH	60.27	H	Red
16	Mistakes during construction	90.27	VH	64.32	H	Red
17	Inadequate monitoring and control	77.03	H	62.70	H	Red
d) Financial Management Related Factors:						
18	Cash flow	60.54	H	48.92	M	Yellow

19	Poor financial control on site	59.46	M	44.86	M	Yellow
20	Financial difficulties of owner	49.19	M	59.73	M	Yellow
21	Delay in progress payment by owner	55.41	M	58.11	M	Yellow
22	Delay payment to supplier /subcontractor	65.14	H	64.86	H	Red
23	Contractual claims	57.30	M	45.14	M	Yellow
e) Information and Communication Related Factors:						
24	Lack of coordination between parties	58.92	M	57.57	M	Yellow
25	Slow information flow between parties	55.41	M	58.38	M	Yellow
26	Lack of communication between parties	59.46	M	54.59	M	Yellow
f) Human Resources (Workforce) Related Factors:						
27	Labour productivity	80.00	VH	70.54	H	Red
28	Shortage of site workers	78.38	H	69.46	H	Red
29	Shortage of technical personnel	76.49	H	47.03	M	Yellow
30	High cost of labour	69.73	H	48.92	M	Yellow
31	Labour absenteeism	70.00	H	42.70	M	Yellow
g) Non-Human Resource Related Factors:						
32	Fluctuation of prices of materials	83.51	VH	45.95	M	Red
33	Shortages of materials	82.43	VH	41.89	M	Red
34	Equipment availability and failure	76.76	H	39.73	L	Yellow
35	Late delivery of materials and equipment	81.08	VH	53.24	M	Red
h) Environmental factors						
36	Level and number of competitors	69.73	H	64.86	H	Red
37	economic instability	63.24	H	51.08	M	Yellow
38	effects of weather	60.81	H	38.38	L	Yellow
39	government policies	46.76	M	40.81	M	Yellow
40	inadequate production of raw materials by the country	55.14	M	45.95	M	Yellow

41	monopoly by suppliers	64.86	H	57.57	M	Yellow
42	number of projects going at the same time	70.00	H	64.86	H	Red
43	political situation	64.32	H	34.05	L	Yellow
44	project location	28.92	L	21.89	L	Green
45	social and cultural impacts	23.24	L	21.35	L	Green

The risk mapping of the actual data was based on the same 45 factors of cost overrun. However; in case the factor didn't occur in any one of the projects; a very low level of severity and frequency were considered, and the factor was immediately located in the green zone, as shown in Table 4.4.

Table (4.4)
Factors of Cost Overrun Risk Mapping for Actual Data and Combined

	Factors	Severity index	Severity level	Frequency Index	Frequency level	Actual Data Map zone	Combined Map zone
a) Project Management and Contract Administration Related Factors:							
1	Poor project management	60.00	H	60.00	H	Red	Red
2	Change in the scope of the project	100.00	VH	40.00	M	Red	Yellow
3	Delays in decisions making	65.00	H	100.00	VH	Red	Yellow
4	Inaccurate quantity take-off	65.00	H	100.00	VH	Red	Red
b) Design and documentation Related Factors:							
5	Frequent design changes	60.00	H	100.00	VH	Red	Red
6	Mistakes and errors in design	66.67	H	80.00	VH	Red	Red
7	Incomplete design at the time of tender	50.00	M	100.00	VH	Red	Red
8	Poor design and delays in design	70.00	H	40.00	M	Yellow	Yellow
9	Delay preparation and approval of drawings	100.00	VH	40.00	M	Red	Red

c) Contractor's site Management Related Factors:							
10	Poor site management and supervision	66.67	H	100.00	VH	Red	Red
11	Incompetent subcontractors	0.00	VL	0.00	VL	Green	Green
12	Schedule delay	51.43	M	100.00	VH	Red	Red
13	Inadequate planning and scheduling	25.00	L	100.00	VH	Yellow	Red
14	Lack of experience	45.00	M	100.00	VH	Red	Red
15	Inaccurate time and cost estimates	60.00	H	100.00	VH	Red	Red
16	Mistakes during construction	70.00	H	100.00	VH	Red	Red
17	Inadequate monitoring and control	30.00	L	60.00	H	Yellow	Yellow
d) Financial Management Related Factors:							
18	Cash flow	55.00	M	100.00	VH	Red	Yellow
19	Poor financial control on site	40.00	M	40.00	M	Yellow	Yellow
20	Financial difficulties of owner	0.00	VL	0.00	VL	Green	Green
21	Delay in progress payment by owner	66.67	H	80.00	VH	Red	Red
22	Delay payment to supplier /subcontractor	50.00	M	60.00	H	Yellow	Yellow
23	Contractual claims	80.00	VH	40.00	M	Red	Yellow
e) Information and Communication Related Factors:							
24	Lack of coordination between parties	70.00	H	60.00	H	Red	Yellow
25	Slow information flow between parties	30.00	L	60.00	H	Yellow	Yellow
26	Lack of communication between parties	80.00	VH	40.00	M	Red	Yellow
f) Human Resources (Workforce) Related Factors:							
27	Labour productivity	60.00	H	60.00	H	Red	Red
28	Shortage of site workers	52.00	M	100.00	VH	Red	Red
29	Shortage of technical personnel	20.00	L	40.00	M	Green	Yellow
30	High cost of labour	60.00	H	60.00	H	Red	Yellow
31	Labour absenteeism	60.00	H	40.00	M	Yellow	Yellow

g) Non-Human Resource Related Factors:							
32	Fluctuation of prices of materials	0.00	VL	0.00	VL	Green	Green
33	Shortages of materials	30.00	L	100.00	VH	Yellow	Yellow
34	Equipment availability and failure	20.00	L	40.00	M	Green	Green
35	Late delivery of materials and equipment	53.33	M	80.00	VH	Red	Red
h) Environmental factors							
36	Level and number of competitors	20.00	L	40.00	M	Green	Yellow
37	economic instability	0.00	VL	0.00	VL	Green	Green
38	effects of weather	40.00	M	40.00	M	Yellow	Green
39	government policies	0.00	VL	0.00	VL	Green	Green
40	inadequate production of raw materials by the country	0.00	VL	0.00	VL	Green	Green
41	monopoly by suppliers	20.00	L	40.00	M	Green	Yellow
42	number of projects going at the same time	30.00	L	60.00	H	Yellow	Yellow
43	political situation	0.00	VL	0.00	VL	Green	Green
44	project location	0.00	VL	0.00	VL	Green	Green
45	social and cultural impacts	0.00	VL	0.00	VL	Green	Green

Both survey and actual data risk mapping results were established, some factors had zone agreement; others did not, therefore the average results (survey and actual data) were calculated and new risk map was introduced, column eight in table 4.4, shows the combined results risk mapping.

The first group Project Management and Contract Administration related factors, combined risk mapping had four factors of cost overrun, half of them were located in the red zone, while the other half in the yellow zone, this result go hand in hand with both survey and actual risk mapping. The

risk map for factors related to Design and Documentation; indicates that 4 factors were located in the red zone and 1 factor was located in the yellow zone, there was no factor under this group in the green zone, these results goes well also with both survey and actual data mappings.

In the combined risk mapping for Contractor's Site Management related factors, eight factors were identified under this group. The results indicate that 6 factors were located in the red zone, 1 factor was located in the yellow zone and 1 in the green zone. In the risk mapping for Financial Management related factors, six factors were considered under this group, the combined map indicates that only 1 factor under this group was located in the red zone, 4 in the yellow zone and 1 in the green zone, half of these factors matched the survey mapping; while the other half followed the actual data map results.

In the Information and Communication combined risk mapping; three factors were considered under this group, the results indicate that all factors were located in the yellow zone, and that was the exact results, the survey mapping showed, the same agreement with survey results; was seen in the Human resources related factors risk mapping, two of the factors were located in the red zone, while the other three in the yellow zone. However the Non-Human Resources related factors combined risk mapping followed the actual data map, with only 1 factor in both red and yellow zone, while two of the factors were located in the green zone. The Environmental related factors risk mapping indicates that most factors were located in the green zone, three factors in the yellow zone and none in the red zone, however the factor Level and number of competitors; was a considered red zone in the survey map only.

4.3.3 Risk maps results

The analysis of the identified 45 factors of the survey risk mapping, indicates that only 2 factors were located in the green zone, 23 factors were located in the yellow zone, and 20 factors were located in the red zone of the risk map. The actual data risk mapping, had 13 factors in the green zone, 10 in the yellow zone and 22 were located in the red zone. The last risk mapping was for the combined results and it showed 10 factors in the green zone, 19 in the yellow zone; and 16 factors were located in the red zone of the risk map.

4.4 Ranking of Cost Overrun Factors and Groups

To calculate the weighted indexes for severity and frequency of cost overrun factors, the research employed the methodology used by Abd El-Razek et al. (2008) and Le-Hoai et al. (2008). The overrun factors were ranked according to their Importance Index (the product of Severity Index and Frequency Index), as shown in Table 4.5. The eight groups of cost overrun were ranked in Table 4.6.

$$\text{Importance Index (II)} = \text{Severity Index} \times \text{Frequency Index} \quad (4.6)$$

Table (4.5)

Ranking of Cost Overrun Factors Based on the Important Index of the Combined Data

S.N.	Code	Factors	I.I% Survey	I.I% Projects	I.I% Combined
1	FDC	Frequent design changes	76.44	60.00	68.22
2	MDC	Mistakes during construction	58.07	70.00	64.03
3	DS	Schedule delay	67.45	51.43	59.44
4	PSMS	Poor site management and supervision	51.83	66.67	59.25
5	IQT	Inaccurate quantity take-off	53.02	65.00	59.01

6	ITCE	Inaccurate time and cost estimates	56.04	60.00	58.02
7	SSW	Shortage of site workers	54.44	52.00	53.22
8	DPAD	Delay preparation and approval of drawings	46.85	40.00	50.81
9	IDTT	Incomplete design at the time of tender	44.93	50.00	47.46
10	LP	Labour productivity	56.43	36.00	46.22
11	DDM	Delays in decisions making	26.65	65.00	45.82
12	PPM	Poor project management	53.05	36.00	44.53
13	LOE	Lack of experience	43.96	45.00	44.48
14	MED	Mistakes and errors in design	35.19	53.33	44.26
15	LDME	Late delivery of materials and equipment	43.17	42.67	42.92
16	DPPO	Delay in progress payment by owner	32.20	53.33	42.76
17	CF	Cash flow	29.62	55.00	42.31
18	LCOP	Lack of coordination between parties	33.92	42.00	37.96
19	CSP	Change in the scope of the project	34.23	40.00	37.12
20	IPS	Inadequate planning and scheduling	49.10	25.00	37.05
21	DPS	Delay payment to supplier /subcontractor	42.25	30.00	36.12
22	HCL	High cost of labour	34.11	36.00	35.06
23	IMC	Inadequate monitoring and control	48.30	18.00	33.15
24	SM	Shortages of materials	34.53	30.00	32.27
25	LCMP	Lack of communication between parties	32.46	32.00	32.23
26	NP	number of projects going at the same time	45.41	18.00	31.70
27	PDDD	Poor design and delays in design	34.89	28.00	31.45
28	CC	Contractual claims	25.86	32.00	28.93
29	LA	Labour absenteeism	29.89	24.00	26.95
30	LNC	Level and number of competitors	45.23	8.00	26.62
31	SIFP	Slow information flow between parties	32.34	18.00	25.17
32	IS	Incompetent subcontractors	47.89	0.00	23.94
33	MS	monopoly by suppliers	37.34	8.00	22.67
34	STP	Shortage of technical personnel	35.97	8.00	21.98
35	PFCS	Poor financial control on site	26.68	16.00	21.34
36	EW	effects of weather	23.34	16.00	19.67
37	EAF	Equipment availability and failure	30.50	8.00	19.25
38	FPM	Fluctuation of prices of materials	38.37	0.00	19.19
39	EI	Economic instability	32.31	0.00	16.15
40	FDO	Financial difficulties of owner	29.38	0.00	14.69
41	IPRM	Inadequate production of raw materials by the country	25.33	0.00	12.67
42	PS	Political situation	21.91	0.00	10.95
43	GP	Government policies	19.08	0.00	9.54
44	PL	Project location	6.33	0.00	3.17
45	SCI	Social and cultural impacts	4.96	0.00	2.48

Table (4.6)
Groups of Cost Overrun Ranked Using Combined Average Important Index

S.N.	Group	Average IMP. I.
1	Design and documentation Related Factors	48.44
2	Contractor's site Management Related Factors	47.42
3	Project Management and Contract Administration Related Factors	46.62
4	Human Resources (Workforce) Related Factors	36.68
5	Information and Communication Related Factors	31.78
6	Financial Management Related Factors	31.02
7	Non-Human Resource Related Factors	28.41
8	Environmental factors	15.56

4.5 Analysis of Results by Data Source

In order to assess the cost overrun causes by each data source independently, data were separated and analyzed separately. This helped in identifying the degree of agreement between participant's opinions and actual projects data. The frequency index, severity (impact) index and importance index were calculated and causes were ranked accordingly as shown in Table 4.7.

Table (4.7)
Ranking of Top 10 Cost Overrun Factors

S.N.	Survey data based ranking	Actual data based ranking	Combined data ranking
1	Frequent design changes	Mistakes during construction	Frequent design changes
2	Schedule delay	Poor site management and supervision	Mistakes during construction
3	Mistakes during construction	Delays in decisions making	Schedule delay
4	Labour productivity	Inaccurate quantity take-off	Poor site management and supervision
5	Inaccurate time and cost estimates	Frequent design changes	Inaccurate quantity take-off
6	Shortage of site workers	Inaccurate time and cost estimates	Inaccurate time and cost estimates

7	Poor project management	Cash flow	Shortage of site workers
8	Inaccurate quantity take-off	Mistakes and errors in design	Delay preparation and approval of drawings
9	Poor site management and supervision	Delay in progress payment by owner	Incomplete design at the time of tender
10	Inadequate planning and scheduling	Shortage of site workers	Labour productivity

“Frequent design changes” was considered as the most important factor of cost overrun causes regarding the survey and combined results. Frequent design changes can affect the project’s estimated budget and its consequences involve not only the work package for which the change is directed, but other work packages and overhead functions as well, and for sure, it introduces a delaying event into the schedule.

“Mistakes during construction” came first in actual data and third in respondent’s opinions, which can be explained by the amount of reworks, delays and subsequent claims that can cause.

“Inaccurate time and cost estimates” falls in the middle of the most important factors of cost overrun causes according to the overall results. Many construction companies suffer from lack of engineers understanding of cost and value and the subsequent cost implications of what included in the design, and estimators who are inexperienced with estimating process or Bahrain market. But even with experienced estimators some companies tend to put in unrealistically low allowances, in order to keep their bids attractive. Another factor falls in the middle of the most important factors is the “Shortage of site workers“, the labour shortage has a high price for construction businesses, it leads to long lead time, and causes loss of works, therefore in Bahrain; over time is very familiar in all construction firms, and

if this wasn't enough; they head to rent labour with higher expenses; and that leads to huge cost escalation.

The Causes “Inaccurate quantity take-off” and “Poor site management and supervision” recorded a high importance in all data as well, however the rest of causes weren't agreed on.

The result raises a question about those differences generated between opinions of respondents (Survey) and actual project data, therefore a different method of analysis was used to reduce and reclassify cost overrun factors in survey. Factors which had fallen in the green zone were excluded from the factor analysis, because they are characterized as low risk factors, and can usually be disregarded and eliminated from further assessment. Ten such factors had a low risk event, therefore the 45 identified cost overrun factors were reduced to 35 for the factor analysis process.

4.6 Survey Based Cost Overrun Factors Analysis

There are different methods of extracting the factors from a set of data, the method chosen matters more when the sample is small, the variables are few, and/or the communality estimates of the variables differ. The most relevant classical risk research on qualitative aspects of risk is that of Slovic et al. (1980) in which they utilized the factor analysis (principal component analysis) to reduce many qualitative risk characteristics or attributes to a much smaller number of higher order factors. In discussing this original research, Gregory and Mendelson (1992) noted that the original factor analysis conducted by Slovic et al. (1980) has been granted classic status in the field of risk perception.

4.6.1 Factoring method using the Principal Component Analysis (PCA)

PCA seeks a linear combination of variables such that the maximum variance is extracted from the variables. Then it removes this variance and seeks a second linear combination which explains the maximum proportion of the remaining variance, and so on. This is called the principal axis method and results in orthogonal (uncorrelated) factors.

The PCA technique of factor analysis was considered appropriate, because of the limited a priori knowledge available about the number of different cluster relationships; that could be expected for the data sample. The PCA provided a deterministic method to group elements into meaningful subdivisions, in order to overcome multicollinearity problems in the project data. As well, it was a statistical procedure that could uncover relationships among many variables, in the context of this research the variables were the specific cost overrun reasons in construction projects.

In the factor analysis technique, correlations and interactions among variables were summarized into a small number of underlying factors. The method aimed at identifying key variables or groups of variables; that controlled the cost overrun system under study (survey data), in order to compare it with the previous results of the actual data. PCA is a reduction technique used to classify and reduce the number of variables. It was applied on the top 35 factors according to risk map ranking. (Ten factors in green zone were excluded). This method was performed using the SPSS statistics software, which is a software package used for statistical analysis, the software name originally stood for **Statistical Package for the Social Sciences (SPSS)**, **version:** IBM SPSS statistics 23.

4.6.2 Sample size

Guidelines for the minimum sample size needed to conduct factor analysis suggested a minimum sample size of 100 to 200 observations, or ratio of sample size to number of variables above 2.00 (Guadagnoli and Velicer, 1988). PCA requires a large sample size. It is based on the correlation of the variables involved, and correlations usually need a large sample size before they stabilize. Tabachnick and Fidell (2001) suggested the following regarding sample size: 50 observations is poor, 100 - 200 is fair, 300 is good, 500 is very good and 1000 or more is excellent.

For the data set in this project, the sample size was **74** observations (survey responses) and the number of factors to be tested was 35. This gave an acceptable ratio of 2.11 according to Guadagnoli and Velicer, (1988).

4.6.3 Factor analysis tests

Several tests were conducted to test the data as follow:

a) Bartlett test of sphericity

It compares the observed correlation matrix to the identity matrix. In other words, it checks if there is a certain redundancy between the factors of cost overrun that can be summarized with a few number of variables. If the factors are perfectly correlated, only one variable is sufficient. If they are orthogonal, the number of variables must be equal to the number of factors. In this last case, the correlation matrix is the same as the identity matrix. A simple strategy is to visualize the correlation matrix. If the values outside the main diagonal are often high (in absolute value), some factors are correlated; if most of these values are near to zero, the PCA is not really useful.

The Bartlett test of sphericity was good with an approximate chi-square of 906.391 and 595 degrees of freedom. The associated significance level was small at $p = 0.000$ for all the factors as shown in (Table 4.8). It can be said that there is some sort of correlation between factors of cost overrun in the survey results; and therefore it's possible to reduce the number of factors.

Table (4.8)
KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.444
Bartlett's Test of	Approx. Chi-Square	906.391
Sphericity	df	595
	Sig.	0.000

b) Kaiser-Meyer-Olkin measure

The Kaiser-Meyer-Olkin (KMO) index, has the same goal of testing factors correlation; but it's based on another idea. The correlation matrix is always the starting point. It's known that the factors are more or less correlated, but the correlation between two factors can be influenced by the others. Therefore, the partial correlation is used in order to measure the relation between two factors, by removing the effect of the remaining factors. The KMO index compares the values of correlations between factors and those of the partial correlations. If the KMO index is high (almost 1), the PCA can act efficiently; if KMO is low (almost 0), the PCA is not relevant. The KMO gave a sampling adequacy of 0.444. This was much less than 0.5 recommended by (Kaiser, 1974; Hutcheson and Sofroniou, 1999).

The low figure threw initial doubt on the adequacy of the planned factor analysis technique; in overcoming multicollinearity in the data. KMO results

indicated that any factor analysis process carried out on the data would produce marginal results, therefore KMO with values less than 0.5 require remedial action, either by deleting the offending variables or by including other variables related to the offenders. After considering the marginal sample characteristic test results, the principal component/factor analysis method was still used as the preliminary method for reducing the large number of cost overrun factors down to a more significant number.

c) Principle component analysis

An initial analysis of the 35 factors ascertained the proportion of each variable's variance that could be explained by principal components. Columns 2 and 5 of Table 4.9 show that initial communalities of the 35 cost overrun variables, which were 1.000 because, by definition, the full orthogonal principal component analysis showed all of the variances in the variables to be explained by all the factors. The extraction communality shown in Columns 3 and 6 of Table 4.9 gave the percent of variance in a given variable. Variables with high values were represented in the common factor space, while variables with low values were moderately represented.

Table (4.9)**Initial and Extraction Communalities for Cost Overrun Variables**

Communalities					
Cost overrun variable	Initial communality	Extraction communality	Cost overrun variable	Initial communality	Extraction communality
PPM	1.000	0.704	DPPO	1.000	0.763
CSP	1.000	0.675	DPS	1.000	0.710
DDM	1.000	0.690	CC	1.000	0.762
IQT	1.000	0.657	LCOP	1.000	0.666
FDC	1.000	0.649	SIFP	1.000	0.756
MED	1.000	0.635	LCMP	1.000	0.684
IDTT	1.000	0.709	LP	1.000	0.701
PDDD	1.000	0.647	SSW	1.000	0.785
DPAD	1.000	0.755	STP	1.000	0.774
PSMS	1.000	0.724	HCL	1.000	0.806
IS	1.000	0.688	LA	1.000	0.637
SD	1.000	0.730	SM	1.000	0.682
IPS	1.000	0.715	LDME	1.000	0.749
LOE	1.000	0.707	LNC	1.000	0.771
ITCE	1.000	0.702	MS	1.000	0.754
MDC	1.000	0.720	NP	1.000	0.695
IMC	1.000	0.749			
CF	1.000	0.696			
PFCS	1.000	0.751			

The principal component analysis was undertaken with the view of reducing the number of initial cost overrun factors by the redistribution of the variance to the first components extracted in a correlation matrix. The method of eigenvalue decomposition was used. Table 4.10, shows the total variance derived from the principal component analysis before rotation.

Table (4.10)
Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total Variance	% of Variance	Cumulative %	Total Variance	% of Variance	Cumulative %
1	4.954	14.155	14.155	4.954	14.155	14.155
2	2.680	7.657	21.812	2.680	7.657	21.812
3	2.424	6.926	28.738	2.424	6.926	28.738
4	2.122	6.063	34.801	2.122	6.063	34.801
5	1.887	5.390	40.191	1.887	5.390	40.191
6	1.752	5.007	45.198	1.752	5.007	45.198
7	1.657	4.735	49.933	1.657	4.735	49.933
8	1.481	4.232	54.165	1.481	4.232	54.165
9	1.352	3.863	58.028	1.352	3.863	58.028
10	1.314	3.755	61.783	1.314	3.755	61.783
11	1.192	3.405	65.188	1.192	3.405	65.188
12	1.102	3.147	68.335	1.102	3.147	68.335
13	1.081	3.088	71.423	1.081	3.088	71.423
14	.948	2.708	74.132			
15	.875	2.500	76.631			
16	.808	2.309	78.940			
17	.799	2.284	81.224			
18	.767	2.192	83.416			
19	.684	1.954	85.371			
20	.651	1.861	87.232			
21	.615	1.756	88.988			
22	.511	1.460	90.448			
23	.454	1.296	91.744			
24	.423	1.208	92.952			
25	.375	1.071	94.024			
26	.356	1.017	95.041			
27	.308	.879	95.919			

28	.302	.864	96.784			
29	.258	.736	97.520			
30	.215	.615	98.134			
31	.188	.537	98.671			
32	.163	.465	99.136			
33	.131	.375	99.511			
34	.088	.251	99.762			
35	.083	.238	100.000			

Column one of Table 4.10 shows the 35 components analyzed. The total % variance explained by each cost overrun factor was listed in the "Initial eigenvalue Column" of Table 4.10. Column two shows the total eigenvalues in descending order of magnitude with component one accounting for the most variance in the analysis and hence the highest eigenvalue. Component two accounted for as much of the leftover variance as it could, and so on. Column three contained the % of variance that was accounted for by each principal component. Column four of Table 4.10 contained the cumulative percentage of variance accounted for by the current and all preceding principal components. The factors were listed in descending order of variance explained. The third row showed a value of 28.738 which meant that; the first three components together account for 28.738% of the total variance.

The three columns of the right half of Table 4.10 represented the extraction sums of squared loadings and reproduced the values given on the same row on the left side of the table. The number of rows reproduced on the right side of the table was determined by the number of principal components whose eigenvalues were 1.00 or greater. There were 13 principal factors that had eigenvalues greater than one (1.000). Column seven displayed the computed

variance of each variable explained by the 13 factor model. Since the factors were uncorrelated, the total proportion of the variance explained was the sum of the variance proportions of each factor. The figures in Column seven show that variation was relatively even across almost all components and this indicated a low significance of the principal components in reducing the dimensionality of the original set of cost overrun variables. It showed that the first 13 components collectively accounted for only 71.423% of the total variability, while the other 15 components together accounted for the remaining 28.577%.

d) Scree plot

A further test using a scree plot was carried out on the project data. The scree plot graphed the eigenvalue against the number of initial factors (component numbers); so as to support a particular eigenvalue cut-off level. The plot in Figure 4.2 shows the plotted values that have been derived from the first two columns of Table 4.10 (component number and eigenvalue). There is a big drop in eigenvalue between component 1 and component 2; however the slope of the rest of line flattened slightly; indicating that each successive component was accounting for smaller and smaller amounts of the total variance out to component 35.

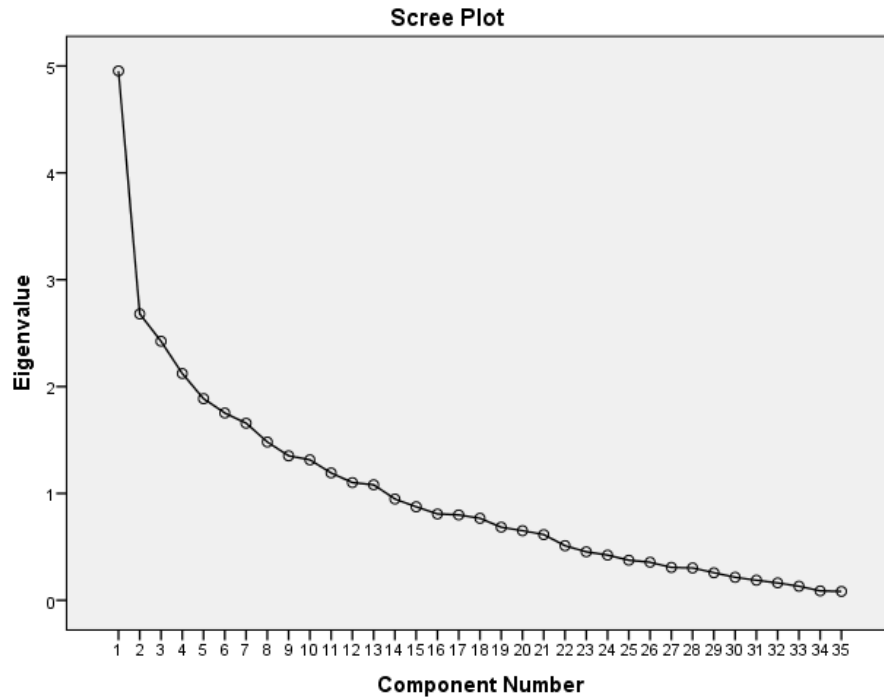


Figure (4.2)
Scree Plot of Cost Overrun Components

e) Factor rotation

In order to facilitate the interpretation of factors, factor analysis requires the rotation of axes. The rotation procedure does not affect the goodness-of-fit of the factor solutions; but serves to make the output more understandable. Three rotation techniques are in general use: varimax, equimax and quartermax. Of these, the most popular is Kaiser's varimax algorithm, which is known to provide the best parsimonious analytical solution (Harman, 1967). This minimizes the number of variables with high loadings on factors, thus causing the factor loadings of each variable to be more clearly differentiated.

Factor rotation was performed based on the varimax method with 9 iterations. It was found that a rotated and ordered factor loading matrix

established 13 component groups for the 35 cost overrun variables as supported by the scree plot. The shaded portions of Table 4.11 show clearly the 13 groupings of the 35 cost-overrun factors. Table 4.12 shows the cost overrun factors grouping from the rotated component matrix.

Table (4.11)
Rotated and Ordered Component Matrix

	Component (Groups)												
	1	2	3	4	5	6	7	8	9	10	11	12	13
MED	.687	.144	.097	.068	.164	.102	.034	.154	.065	.022	.224	.056	.091
SM	.653	.043	.109	.269	.146	.043	.137	.074	.185	.150	.242	.081	.008
SD	.649	.276	.099	.073	.178	.103	.059	.107	.150	.311	.103	.030	.173
CC	.628	.148	.045	.245	.188	.439	.180	.088	.096	.008	.021	.057	.058
PDDD	.530	.174	.370	.103	.136	.133	.055	.194	.026	.178	.053	.223	.162
IMC	.526	.156	.184	.094	.050	.104	.150	.189	.097	.332	.043	.342	.309
IDTT	.506	.222	.144	.001	.141	.229	.078	.354	.009	.143	.308	.062	.244
DPPO	.496	.001	.369	.044	.032	.118	.394	.090	.157	.291	.264	.093	.116
FDC	.494	.317	.275	.293	.128	.063	.036	.047	.045	.300	.146	.074	.005
CSP	.491	.058	.361	.001	.339	.078	.022	.021	.245	.008	.043	.149	.306
MS	.464	.151	.136	.144	.134	.289	.385	.044	.074	.021	.045	.466	.006
LOE	.459	.251	.048	.005	.243	.316	.023	.164	.347	.048	.157	.075	.302
HCL	.060	.650	.125	.077	.365	.182	.142	.044	.101	.219	.327	.039	.057
MDC	.042	.560	.230	.278	.095	.068	.221	.047	.039	.126	.266	.293	.188
ITCE	.150	.532	.003	.091	.071	.509	.027	.082	.080	.192	.246	.061	.099
DDM	.357	.490	.234	.014	.011	.232	.245	.031	.289	.192	.146	.019	.105
PSMS	.371	.418	.317	.105	.088	.238	.134	.323	.274	.191	.042	.027	.018
NP	.156	.310	.532	.026	.055	.261	.201	.233	.252	.030	.130	.192	.085
DPS	.404	.289	.528	.102	.081	.256	.089	.006	.041	.254	.135	.011	.096
IS	.176	.207	.412	.187	.407	.091	.124	.216	.071	.094	.181	.095	.345
PFCS	.244	.259	.095	.544	.255	.241	.283	.182	.261	.019	.087	.059	.058
STP	.121	.208	.356	.496	.149	.021	.120	.208	.147	.192	.074	.400	.198
LCOP	.394	.117	.025	.481	.292	.123	.108	.087	.265	.131	.178	.115	.118
LP	.216	.118	.245	.450	.008	.145	.412	.139	.108	.078	.101	.350	.136

SIFP	.003	.418	.357	.436	.040	.159	.024	.272	.327	.000	.102	.155	.143
LA	.106	.166	.310	.423	.344	.118	.260	.271	.113	.085	.036	.066	.154
DPAD	.337	.017	.267	.393	.342	.004	.176	.220	.102	.087	.267	.172	.316
PPM	.027	.248	.146	.054	.401	.265	.297	.282	.156	.275	.295	.150	.098
LNC	.240	.322	.271	.156	.269	.450	.140	.020	.113	.020	.405	.199	.035
LCMP	.161	.163	-.01	.314	.193	.375	.366	.247	.343	.103	.021	.107	.145
IQT	.274	.145	.226	.125	.351	.157	.554	.011	.039	.014	.053	.154	.107
SSW	.260	.020	.273	.094	.275	.078	.010	.604	.010	.407	.075	.047	.113
CF	.357	.019	.400	.133	.281	.146	.078	.409	.067	.130	.007	.027	.307
LDME	.023	.230	.037	.022	.406	.315	.123	.071	.562	.043	.069	.018	.294
IPS	.004	.055	.090	.188	.264	.032	.330	.004	.183	.510	.347	.255	.097

Table (4.12)

Principal Factor Groupings from Rotated Component Matrix

Component group	Code	Description	Highest factor in component group
Component group 1	MED, SM, CC, PDDD, MC, IDTT, DPPO, CSP, LOE, DDM	Mistakes and errors in design Shortages of materials Contractual claims Poor design and delays in design Incomplete design at the time of tender Delay in progress payment by owner Change in the scope of the project Lack of experience Delays in decisions making	Mistakes and errors in design
Component group 2	HCL, MDC, ITCE, PSMS, SIFP	High cost of labour Mistakes during construction Inaccurate time and cost estimates Poor site management and supervision Slow information flow between parties	High cost of labor
Component group 3	NP, DPS	Delay payment to supplier /subcontractor number of projects going at the same time	number of projects going at the same time
Component group 4	STP, LCOP, LA, DPAD	Shortage of technical personnel Labour absenteeism Lack of coordination between parties Delay preparation and approval of drawings	Shortage of technical personnel
Component group 5	IQT	Inaccurate quantity take-off	Inaccurate quantity take-off

Component group 6	LNC, LCMP	Level and number of competitors Lack of communication between parties	Level and number of competitors
Component group 7	PFCS, LP, PPM	Poor project management Labour productivity Poor financial control on site	Poor financial control on site
Component group 8	SSW, CF	Shortage of site workers Cash flow	Shortage of site workers
Component group 9	LDME	Late delivery of materials and equipment	Late delivery of materials and equipment
Component group 10	SD, FDC	Schedule delay Frequent design changes	Schedule delay
Component group 11	IPS	Inadequate planning and scheduling	Inadequate planning and scheduling
Component group 12	MS	monopoly by suppliers	monopoly by suppliers
Component group 13	IS	Incompetent subcontractors	Incompetent subcontractors

A review of the 13 groupings lead to the conclusion that; no common factor component names were developed in the principal component analysis, that could be allocated to the individual component groups so as to allow consistency in purpose of the group of factors. There appeared to be no evidence of strong correlations within any of the 13 groupings, even after rotation. This was particularly evidenced by the diversity in the construction processes associated with each of the principal component groups. These findings, supported by the weak scree plot demonstrated the invalidity of the factor analysis process in this research. This finding was further supported by; the weak sampling adequacy of the initial cost overrun data available as determined previously by the KMO test. For further analysis the principle component analysis of the survey data results is compared with the actual data.

4.7 Survey and Actual Based Data Rankings Analysis

From Table 4.12, and based on the conducted PCA analysis, 13 factors have been extracted, according to factor weights (Table 4.13). The top 13 factors of cost overrun in construction projects (survey and actual data) are shown in Table 4.14.

Table (4.13)

Highest 13 Factors Extracted from PCA (Survey)

S.N.	Factor	weight
1	Mistakes and errors in design	0.687
2	High cost of labour	0.650
3	number of projects going at the same time	0.532
4	Shortage of technical personnel	0.496
5	Inaccurate quantity take-off	0.351
6	Level and number of competitors	0.450
7	Poor financial control on site	0.283
8	Shortage of site workers	0.604
9	Late delivery of materials and equipment	0.562
10	Schedule delay	0.311
11	Inadequate planning and scheduling	0.347
12	monopoly by suppliers	0.466
13	Incompetent subcontractors	0.345

(Table 4.14)

Cost Overrun in Construction Projects: Factors Ranking Based on Factor Weight

#	Practitioners Opinions (Survey)	Actual Data (Actual Projects)
1	Mistakes and errors in design	Mistakes during construction
2	High cost of labour	Poor site management and supervision
3	Shortage of site workers	Delays in decisions making
4	Late delivery of materials and equipment	Inaccurate quantity take-off
5	number of projects going at the same time	Frequent design changes
6	Shortage of technical personnel	Inaccurate time and cost estimates
7	monopoly by suppliers	Cash flow

8	Level and number of competitors	Mistakes and errors in design
9	Inaccurate quantity take-off	Delay in progress payment by owner
10	Inadequate planning and scheduling	Shortage of site workers
11	Incompetent subcontractors	Schedule delay
12	Schedule delay	Incomplete design at the time of tender
13	Poor financial control on site	Lack of experience

Table (4.15)
Cost Overrun in Construction Projects: Common Factors

#	Factors	Rank		Matching%
		Survey	Actual	
1	Schedule delay	12	11	91.6%
2	Inaccurate quantity take-off	9	4	45%
3	Shortage of site workers	3	10	30%
4	Mistakes and errors in design	1	8	12.5%

After conducting the analysis on both the survey and actual data, results showed agreement only on four causes (Table 4.15), these causes were: mistakes and errors in design, schedule delay, inaccurate quantity take-off, and shortage of site workers. While disagreement was found on other factors such as high cost of labour and inaccurate time and cost estimates.

The differences in the results can be referred to several reasons; survey respondents for example may not feel encouraged to provide accurate detailed answers, or not be fully aware of their reasons for any given answer because of lack of memory on the subject, or even boredom. Engineers usually have a busy work and survey was a time consuming; lack of adequate time to complete the instrument may result in the return of superficial data. The actual data on the other hand contained forty construction projects only; which is considered a very low number in order to draw a sufficient conclusion, also some of the common causes of cost

overrun probably did not occur in any one of the forty projects despite its commonness.

The findings show consensus of ranking among the different analysis methods (PCA, Survey, Actual, and Combined Risk Mapping) that schedule delay, inaccurate quantity take-off, shortage of site workers, and mistakes and errors in design served as main factors contributing to construction projects cost overrun. Most of these causes affect each other, for example shortage of site workers, and mistakes and errors in design may lead to schedule delay. Schedule delay is the construction schedule being extended, meaning that the contractor remains on site longer than allowed for. It incurs additional costs for the construction site facilities, staff salaries plus their associated costs, and the inefficient and unproductive use of personnel and equipment. Schedule delay will also affect the construction materials which have been ordered, resulting in storage costs, double handling, and the associated risks of damage and theft. It also reflects on the disruption of cash flow because the project's end date is extended, deferring the release of retention monies and securities, and subcontractors claiming delay costs.

Quantity take-off is the basis of the cost estimation process; any estimating errors can be very expensive and embarrassing. The worst type of error in construction cost estimating is leaving out an item of work. On a lump sum bid and contract, the contractor is obligated to do all of the work, even if costs for part of the project are omitted in the estimate bid and contract to build. Ensuring that all items have been included in the takeoff portion of the estimate is perhaps the most difficult part of the estimate thus it requires careful attention. It can leads to shortage of site workers and other cost overrun factors. However, most estimators in hurry to complete the cost

estimate, what causes lack of attention to details, and failure to use a reliable checklist. An important safeguard against overlooking items is to have another person review and double check the cost estimate before sending it out to the owner or general contractor. Also not allowing for realistic contingencies in the estimate may result in not getting the contract or losing money if awarded the contract. Justifications for a realistic contingency could include anticipated labour troubles, material shortages, or even extremely hot and humid weathers.

CHAPTER FIVE

MODELING CONCEPTUAL BACKGROUND

5.1 Introduction

5.2 Different Approaches to SEM

5.3 PLS-SEM Theoretical Background

5.4 Determination of Sample Size for PLS-SEM

5.5 PLS-SEM Model Assessment

5.6 Model Evaluation

CHAPTER FIVE

MODELING CONCEPTUAL BACKGROUND

5.1 Introduction

Structural Equation Modelling (SEM) is a second-generation multivariate data analysis method that can test theoretically supported linear and additive causal models. With SEM, it's possible to visually examine the relationships that exist among variables of interest (cost overrun factors) in order to prioritize resources. The unobservable, hard-to-measure latent variables, which are underlying variables that cannot be observed directly, can be measured using SEM and that makes it ideal for tackling research problems.

There are two sub-models in a structural equation model; the inner model (structural model) specifies the relationships between the independent and dependent latent variables (LVs), whereas the outer model (measurement model) specifies the relationships between the latent variables and their observed indicators which act like indicators for an underlying latent variable.

In SEM, a variable is either exogenous or endogenous. An exogenous variable has path arrows pointing outwards and none leading to it. The endogenous variable has at least one path leading to it and represents the effects of other variable(s).

Depending on the SEM design, a variable can technically act as an independent variable or a dependent variable for different parts of the model. As long as a variable has path leading to it (i.e., arrows pointing to it from another variable), it

is categorized as endogenous. In first measurement model the constructs are endogenous variables and the factors are exogenous variables, however in the structural model the constructs turn in to exogenous variables, while the cost overrun becomes the endogenous variable (Figure 5.1).

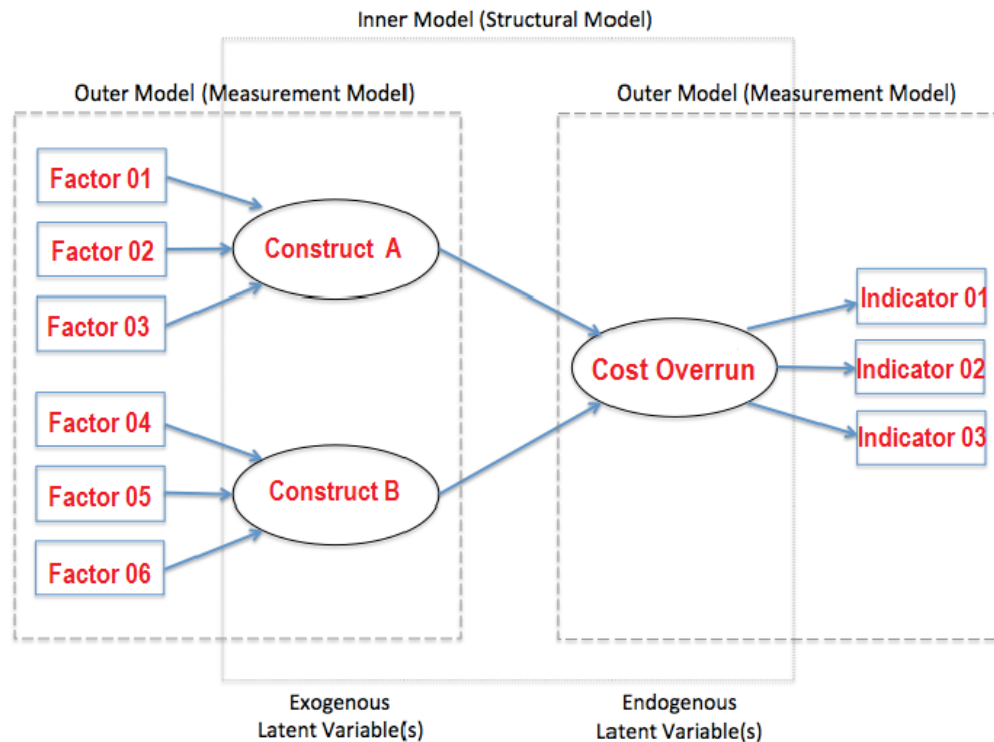


Figure (5.1)

Inner and Outer Model

5.2 Different Approaches to SEM

There are several distinct approaches to SEM: The first approach is the widely applied Covariance-based SEM (CB-SEM), using software packages such as analysis of moment structure (AMOS), environmental quality standard (EQS), and linear structural relations (LISREL). The second approach is Partial Least Squares (PLS), which focuses on the analysis of variance and can be carried out using PLS-

Graph, VisualPLS, SmartPLS, and WarpPLS. It can also be employed using the PLS module in the “r” statistical software package. The third approach is a component-based SEM known as Generalized Structured Component Analysis (GSCA); it is implemented through VisualGSCA or a web-based application. Another way to perform SEM is called Nonlinear Universal Structural Relational Modeling (NEUSREL), using NEUSREL’s Causal Analytics software.

In this research the Partial Least Squares (PLS) was applied, it focuses on the analysis of variance and can be carried out using SmartPLS software. PLS have advantages and disadvantages. PLS is a soft modelling approach to SEM with no assumptions about data distribution (Vinzi et al., 2010). Thus, PLS-SEM becomes a good alternative to CB-SEM when the following situations are encountered (Bacon, 1999; Hwang et al., 2010; Wong, 2010):

1. Sample size is small.
2. Applications have little available theory.
3. Predictive accuracy is paramount.
4. Correct model specification cannot be ensured.

It is important to note that PLS-SEM is not appropriate for all kinds of statistical analysis, some weaknesses of PLS-SEM, including:

1. High-valued structural path coefficients are needed if the sample size is small.
2. Problem of multicollinearity if not handled well.
3. Since arrows are always single headed, it cannot model undirected correlation.
4. A potential lack of complete consistency in scores on latent variables may result in biased component estimation, loadings and path coefficients.

5. It may create large mean square errors in the estimation of path coefficient loading.

In spite of these limitations, PLS is useful for structural equation modelling in applied research projects especially when there are limited participants and that the data distribution is skewed, PLS-SEM has been deployed in many fields, such as behavioural sciences (Bass et al., 2003), marketing (Henseler et al., 2009), organization (Sosik et al., 2009), management information systems (Chin et al., 2003), and business strategy (Hulland, 1999).

5.3 PLS-SEM Theoretical Background

Every PLS Path Model is formed by two sub models: the structural or inner model, and the measurement or outer model. The structural model is the part of the model that has to do with the relationships between the latent variables (cost overrun and cost overrun groups). In turn, the measurement model is the part of the model that has to do with the relationships of the latent variable with its block of manifest variables (cost overrun groups and cost overrun factors). PLS-SEM path models only permit recursive relationships and can be expressed as simple connected digraphs. A digraph is called simple if it has no loops and at most one arc between any pair of nodes. A digraph is connected if an undirected path between any two nodes exists; consequently no node is isolated from the rest.

5.3.1 The structural model

In the structural model the latent variables (LVs) are related with each other according to substantive theory. As mentioned before, LVs are divided into two classes, exogenous and endogenous. Exogenous LVs do not have any predecessor in the structural model, all others are endogenous.

In the PLS-SEM path models $\mathbf{K}$ is the number of variables measured on $\mathbf{N}$ observations, the variables can be divided into $\mathbf{G}$ blocks. $\mathbf{X}$ is the data set containing the $\mathbf{N}$ observations and $\mathbf{K}$ variables; $\mathbf{X}$ is a matrix of dimension $\mathbf{N} \times \mathbf{K}$; $\mathbf{X}$ can be divided into $\mathbf{G}$ (mutually exclusive) blocks $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_g$. Each block $\mathbf{X}_g$ is assumed to be associated with a latent variable $\mathbf{LV}_g$; $\mathbf{LV}_g$ is just an abstract representation (i.e. unobserved). The estimation of a latent variable, also known as score, is denoted by $\hat{\mathbf{LV}}_g = \mathbf{Y}_g$. For simplicity of the notation the structural model dismisses the difference, between exogenous and endogenous variables, with the compact form (Monecke and Leisch, 2012):

$$\mathbf{Y} = \mathbf{Y} \mathbf{B} + \mathbf{Z} \quad (5.1)$$

Where, $\mathbf{Y}$: denotes the matrix for the latent variables, both exogenous and endogenous.

$\mathbf{Z}$: the error terms, $\mathbf{Z}$ are assumed to be centred, i.e., $E[\mathbf{Z}] = 0$.

$\mathbf{B}$: elements of the coefficients matrix, $\mathbf{B}$ are restricted to zero where the elements of the adjacency matrix $\mathbf{D}$ are zero.

5.3.2 The measurement model

The measurement model or outer model relates observed variables to their latent variables (LVs). Often observed variables are referred to as manifest variables (MVs) or indicators which are causes or factors of cost overrun, latent variables are the groups of cost overrun. Within the PLS framework one manifest variable can only be related to one LV. All manifest variables related to one LV are called a block. So each LV has its own block of observed variables. A block must contain at least one MV. The way a block can be related to an LV can be either reflective

(as in this model) or formative. Since the model in this research is reflective, the following assumptions can be made (Monecke and Leisch, 2012):

1. All MVs contained in the data matrix X are scaled to have zero mean and unit variance.
2. Each block of MVs X_g is already transformed to be positively correlated for all LVs y_g , $g = 1, \dots, G$.

Reflective measurement: In the reflective way each block of MVs reflects its LV and can be written as the multivariate regression:

$$X_g = y_g w_g^T + F_g, \quad E[F_g | y_g] = 0 \quad (5.2)$$

So the weight of the MV block w_g^T can be estimated by least squares as

$$\begin{aligned} W_g^T &= (y_g^T y_g)^{-1} y_g^T X_g \\ &= \text{VAR}(y_g)^{-1} \text{COV}(y_g, X_g) \\ &= \text{COV}(y_g, X_g) \\ &= \text{COR}(y_g, X_g). \end{aligned} \quad (5.3)$$

Where;

X_g , $g = 1, \dots, G$: Block of observed variables (MV) belonging to latent y_g

Y : $N \times G$ matrix of for the latent variables (LVs)

W_g^T : the weight of the observed manifest variable block

F_g : $N \times K$ matrix measurement error for reflective block X_g

T : $G \times G$ matrix of the total effects

E : $G \times G$ matrix of inner weights

PLS algorithm estimates all the LVs y_g , $g = 1, \dots, G$, as linear combination of their MVs under the constraint to have unit variance. All MVs are scaled to zero mean and unit variance. Consequently, the equality above is valid. The latent variable y_g is measured by the block X_g consisting of observed variables, $x_1, x_2, x_3, \dots$, in a reflective way (Figure 5.2).

It was assumed that all the MVs to be scaled to zero mean and unit variance. Consequently, the equality above is valid. Figure 5.3 depicts a path diagram for a reflectively measured LV.

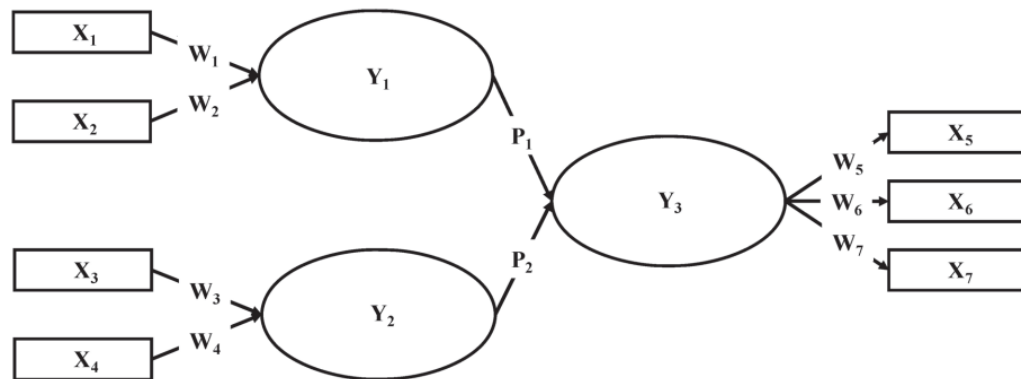


Figure (5.2)

Reflective Model Path Diagram

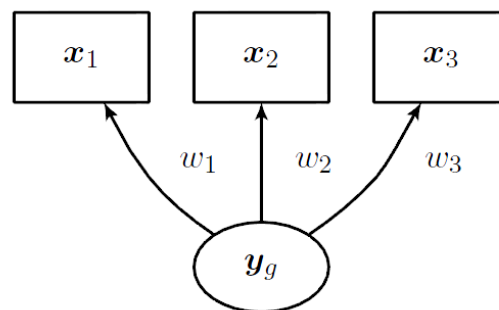


Figure (5.3)

Reflective LV Path Diagram

The latent variable y_g is measured by the block X_g consisting of three observed variables, $x_1, \dots, x_3$, in a reflective way.

5.3.3 The PLS-SEM algorithm

The PLS algorithm aims at estimating the values for LVs (factor scores) by an iterative procedure. Figure 5.4 depicts the flowchart of the algorithm. The idea is first to construct each LV by the sum of its MVs. Then, in the inner approximation, each LV is reconstructed by the means of its neighbouring LVs. In the outer approximation the best linear combination must be found to express each LV by means of its MVs; the coefficients are referred to as outer weights. Finally, each LV is constructed as weighted sum or linear combination of its MVs. After each step the LVs are scaled to have zero mean and unit variance. The algorithm stops if the relative change for all the outer weights is smaller than a predefined tolerance (Wold 1982; Lohmoller 1989). (Equations of each step are available in Appendix 4).

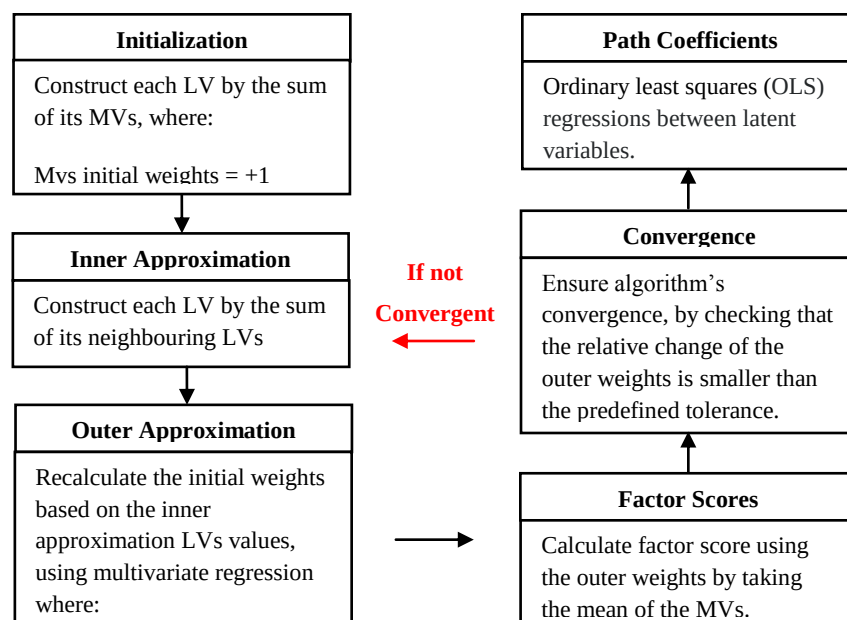


Figure (5.4)

The Diagram Depicts the Flowchart for the PLS Algorithm

Stage one has four steps. In Step 1 of stage one, the outer proxies of the latent variables scores' are computed as linear combinations of the values of all (standardized) indicators associated with a particular latent variable. In Figure 5.2, values of X_1 and X_2 are used to compute the proxy score for the latent variable Y_1 . Later iterations use the estimated coefficients of the paths (e.g., W_1 and W_2 for Y_1), whereby W represents the outer weight or loading coefficient. For the initial iteration, any combination of indicators can serve as a proxy for the latent variable. PLS-SEM software programs, such as SmartPLS use a uniform value of 1 as an initial value for each of the outer weights (W_1 to W_7). In Step 2, the PLS-SEM algorithm computes proxies for the structural model relationships (P_1 and P_2). Several different weighting schemes are available to estimate these proxies. It's recommended to apply the path weighting scheme that uses combinations of regression analyses and bivariate correlations based on latent variable scores, as proxies for structural model relationships. This method develops latent construct scores that maximize the final R^2 value estimations of the endogenous latent constructs (Lohmoller, 1989). In Step 3, the inner proxies of the latent construct scores (Y_1 , Y_2 , and Y_3) are calculated as linear combinations of their respective adjacent latent construct outer proxies (from Step 1) using the previously determined inner weights (Step 2). Finally, in Step 4, the outer weights (W_1 to W_7) are calculated in two different ways, depending on the type of measurement model represented by each construct. Since the variable is measured reflectively, then the correlations between the inner proxy of each latent variable and its indicators are applied (outer loadings).

The four steps in Stage one are repeated until the sum of the outer weights' changes between two iterations is sufficiently low. That is, the sum of the changes drops below a predetermined limit. The recommended predetermined limit is a

threshold value of 10–5 to ensure the PLS-SEM algorithm’s convergence while simultaneously minimizing computational requirements. If the algorithm converges in Step 4 of Stage one, then the final outer weights are used to compute the final latent variable scores in Stage two. The final latent variable scores are used to run the ordinary least squares regressions for each construct to determine the structural model relationships’ estimates (path coefficients). If the algorithm not convergent Step 2 must be recalculated.

5.4 Determination of Sample Size in PLS-SEM

In order to determine the appropriate sample size the following has to be considered: the background of the model, the distributional characteristics of the data, the psychometric properties of variables, and the magnitude of their relationships. Hair et al. (2013) suggested that sample size can be driven by the following factors in a SEM design:

1. The significance level
2. The statistical power
3. The minimum coefficient of determination (R^2 values) used in the model
4. The maximum number of arrows pointing at a latent variable

In practice, a typical marketing research study would have a significance level of 5%, a statistical power of 80%, and R^2 values of at least 0.25. Using such parameters, the minimum sample size required can be looked up from the guidelines suggested by Marcoulides and Saunders (2006), depending on the maximum number of arrows pointing at a latent variable as specified in the SEM (Table 5.1).

Table (5.1)
Suggested Sample Size in a Typical Research According to Marcoulides and
Saunders, (2006)

Minimum sample size required	Maximum number of arrows pointing at a latent variable in the model
52	2
59	3
65	4
70	5
75	6
80	7
84	8
88	9
91	10

Because there are 8 groups of cost overrun factors, one of them consist of 10 variables, and based on Table 5.1, minimum sample size required is 91. The sample size used in developing this model will consist of 114 samples which is acceptable according to Marcoulides and Saunders (2006).

5.5 PLS-SEM Model Assessment

A two-step process will be adopted to calculate PLS model criteria (Henseler et al., 2009). The PLS path model evaluation steps are:

1. Outer model (measurement model) assessment

- a. individual item reliability and convergent validity
- b. discriminant validity
- c. assess significance and relevance of the measurement model relationships

2. Inner model (structural model) assessment

- a. assess the level of R^2 and path coefficient significance
- b. assess significance and relevance of the structural model relationships

5.5.1 Assessment of measurement model

a) Assessing individual factor reliability and convergent validity of measurement model:

Individual factor (indicator) reliability is the extent to which measurements of the latent variables (Constructs) measured with multiple-item scale reflects mostly the true score of the latent variables relative to the error (Hulland, 1999). It is the correlations of the items with their respective latent variables. To evaluate individual item reliability, the standardized loadings (or simple correlations) were assessed (Table 5.2).

Table (5.2)

Individual Factor (Indicator) Reliability Assessment

Year	Author	Individual item reliability
1998	Chin	• factor with loading below than 0.5 should be dropped
1999	Hulland	• factors with loadings of less than 0.4 should be dropped • factors with outer loading higher than 0.7 should be considered highly satisfactory • factors with loading between 0.4 to 0.7 practical potential significance should be assessed prior to elimination
2010	Aibinu and Al-Lawati	• factors with low loadings should be reviewed, and perhaps dropped since they would add very little explanatory power to the model and therefore biasing the estimates of the parameters linking the latent variables
2011	Aibinu et al.	• If an indicator's reliability is low and eliminating this indicator goes along with a substantial increase of composite reliability, it makes sense to discard this indicator

Convergent validity is the measure of the internal consistency which ensures that the items assumed to measure a particular construct actually measure it and not another construct (Hulland, 1999). Composite Reliability scores (CR), Cronbach's Alpha and Average Variance Extracted (AVE) tests were used to determine the convergent validity of measured constructs. The cut-off value for AVE, CR and Cronbach Alpha were 0.5, 0.7 and 0.6, respectively (Table 5.3).

Table (5.3)
The Cut-off Value for AVE, CR and Cronbach Alpha

Code	Measure	Utilization	Cut-off Value	Source
CR	Composite Reliability	check how well a construct is measured by its assigned indicators	can vary between 0 and 1. Researchers argue that composite reliability value for a good model should be more than 0.7	Akter et al., 2011
Alpha	Cronbach's alpha	reliability coefficient used to generalize measure of a uni-dimensional scale's internal consistency.	can vary between 0 to 1. A common threshold for sufficient values of Cronbach's alpha is 0.6 and if the value is more than 0.7, data is considered as highly acceptable	Yang and Ou, 2008; Wong and Cheung, 2005; Akter et al., 2011
AVE	Average Variance Extracted	measures the amount of variance that a latent variable captures from its measurement items relative to the amount of variance due to measurement errors.	should be higher than 0.5. This means that at least 50% of measurement variance is captured by the latent variables	Fornell and Larcker 1981; Akter et al., 2011; Aibinu et al., 2011

b) Discriminant validity of constructs:

Discriminant validity indicates the extent to which a given construct is different from other constructs (Hulland, 1999). The discriminant validity of the measurement will be evaluated using analysis of the average variance extracted by using the criteria that “a construct should share more variance with its measures than it shares with other constructs in the model (Fornell and Larcker, 1981; Aibinu et al., 2011). This can be examined by comparing the AVE of construct shared on itself and other constructs. For valid discriminant of construct, AVE shared on itself should be greater than shared with other constructs.

- Square root of the AVE of each construct should be larger than the correlation of two constructs. This is done by replacing the diagonal of correlation matrix with the value of square root of the AVE.
- For adequate discriminant validity, the diagonal elements need to be greater than the off-diagonal elements in the corresponding rows and columns.

c) Significance and relevance of the indicators:

Examine the **statistical significance** of the **outer weights** using the bootstrapping. If any indicator weights are not statistically significant, then the **size and significance** of the **indicator loadings must be examined using the t values** ($t > 0.01$).

Examine each indicator's outer weight (relative importance) and outer loading (absolute importance) by using bootstrapping to assess their significance. Three rules applied:

- When an indicator's weight is significant there is empirical support to retain the indicator

- When an indicator's weight is not significant; but the corresponding factor loading is relatively high (i.e. >0.5), the indicator should generally be retained.
- If both the outer weight and outer loading are nonsignificant, there is no empirical support to retain the indicator and it should be removed from the model.

5.5.2 Assessment of structural model

Once the construct measures have been confirmed as reliable and valid, the next step is to assess the structural model results. This involves examining the model's predictive capabilities and the relationships between the constructs.

a) Assessing the level of R^2 :

The R^2 of the endogenous latent variable is used to assess the explained variance. According to Cohen (1992) R^2 of endogenous can be assessed as substantial = 0.26, moderate = 0.13 and weak = 0.02. (Cut-off values).

b) Assessing significance and relevance of the structural model relationships:

After applying the PLS-SEM algorithm, estimates are obtained for the structural model relationships (the path coefficients), which represent the hypothesized relationships between the constructs. The significance of the path coefficient is tested by calculating t-value using non-parametric bootstrap procedure to provide confidence intervals for all parameter estimates, building the basis for statistical inference. In general, the bootstrap technique provides an estimate of the shape, spread, and bias of the sampling distribution of a specific statistic. Bootstrapping treats the observed sample as if it represents the population. The procedure creates

a large, pre-specified number of bootstrap samples (e.g.,5,000). Each bootstrap sample should have the same number of cases as the original sample. Bootstrap samples are created by randomly drawing cases with replacement from the original sample. The PLS results for all bootstrap samples provide the mean value and standard error for each path model coefficient. This information permits a t-test to be performed for the significance of path model relationships (Henseler et al., 2009). All the paths retrieved t-value must be higher than minimum cut-off value i.e., 2.58 at significance level = 1% (Hair et al., 2013). This will implies that the constructs have significant effect on cost overrun.

5.6 Model Evaluation

5.6.1 Overall model assessment (GoF)

Global fit measures goodness of fit (GoF) for PLS path modelling, defined as the geometric mean of the average communality and average R^2 (for endogenous constructs), will be used to assess over all model fitness and to explain power of the model. The GoF index is bounded between 0 and 1. The cut-off values for global validation of PLS model are shown in Table 5.4, and the equation used to calculate GoF as adopted by Akter et al. (2011) is:

$$\text{GoF} = \sqrt{\overline{\text{AVE}} \times \overline{R^2}} \quad (5.4)$$

Table (5.4)

GoF Cut-off Values for Global Validation

GoF	Value
small	0.1
medium	0.25
large	0.36

5.6.2 Assessing effect size – f^2

The f^2 effect size is a measure of the impact of a specific predictor construct on an endogenous construct. In addition to evaluating the size of the R^2 values of all endogenous constructs, the **f^2 effect size** can be calculated. The f^2 effect size measures the change in the R^2 value when a specified exogenous construct is omitted from the model. It is used to evaluate whether the omitted predictor construct has a substantive impact on the R^2 values of the endogenous construct(s). The formula for calculating the f^2 effect size is:

$$f^2 = \frac{R^2_{\text{included}} - R^2_{\text{excluded}}}{1 - R^2_{\text{included}}} \quad (5.5)$$

In the formula R^2_{included} and R^2_{excluded} are the R^2 values of the endogenous latent variable when a selected exogenous latent variable is included or excluded from the model. The change in the R^2 values is calculated by estimating the PLS path model twice. It is estimated the first time with the exogenous latent variable included (yielding R^2_{included}) and the second time with the exogenous latent variable excluded (yielding R^2_{excluded}).

Guidelines for assessing f^2 values for the exogenous latent constructs in predicting the endogenous constructs (Cohen, 1992), are shown in Table 5.5.

Table (5.5)
Guidelines for Assessing f^2 Values

<u>Value</u>		<u>Effect Size</u>
0.02	=	small
0.15	=	medium
0.35	=	large

5.6.3 Assessing the predictive relevance Q^2

Q^2 value is an indicator of the model's **predictive relevance**. The **Q^2** measure applies a sample re-use technique that omits part of the data matrix and uses the model estimates to predict the omitted part. Specifically, when a PLS-SEM model exhibits predictive relevance, it accurately predicts the data points of the indicators in reflective measurement models of multi-item as well as single-item endogenous constructs (the procedure does not apply to formative endogenous constructs). For SEM models, Q^2 values larger than zero for a specific reflective endogenous latent variable indicate the path model's predictive relevance for a particular construct. Q^2 values of zero or below indicate a lack of predictive relevance. As a relative measure of predictive relevance, values of 0.02, 0.15, and 0.35 indicate that an exogenous construct has a small, medium, or large predictive relevance for a selected endogenous construct.

The Q^2 value can be calculated by using two different approaches.

- The **cross-validated redundancy** approach uses the path model estimates of both the structural model (scores of the antecedent constructs) and the measurement model (target endogenous constructs).
- The **cross-validated communality** approach. This method uses only the construct scores estimated for the target endogenous construct (without

including the structural model information) to predict the omitted data points.

In this research the cross-validated redundancy will be used as a measure of Q^2 since it includes the key element of the path model, the structural model, to predict eliminated data points. The same guidelines for assessing the effect size f^2 apply on the predictive relevance (Q^2) which were shown in Table 5.6.

CHAPTER SIX

MODEL DEVELOPMENT AND EVALUATION

6.1 Introduction

6.2 Model Input Data

6.3 Model Development

6.4 Model Evaluation

6.5 Model Contribution

6.6 Risk Mapping and PLS-SEM Model Results Comparison

CHAPTER SIX

MODEL DEVELOPMENT AND EVALUATION

6.1 Introduction

The research aims to investigate the causes of cost overrun and their effects in building construction projects in Bahrain. Cost overrun is caused by many factors; each factor has a different rate of occurrences, and severity on project cost, to assess these risks efficiently each data source was analyzed independently, three risk maps were developed (survey, actual, and combined data) and factors of each map were ranked according to their importance indices. The degree of agreement between participant's opinions and actual projects data was then evaluated, but the results of analysis raised a question about those differences generated between opinions of respondents (Survey) and actual project data. Therefore a different method of analysis was used to reduce and reclassify cost overrun factors in survey data. Survey data was reanalyzed using the PCA technique of factor analysis, because of the limited prior knowledge available about the number of different cluster relationships that could be expected for the data sample. The PCA results were rejected due to the conclusion that; no evidence of strong correlations within any of the 13 groupings has appeared, even after rotation process. Therefore, an alternative approach is required, precisely constructing a model to investigate the effects of different cost factors on project cost overrun. Before carrying on with the model development, a reliability test was performed on data sets to choose the most reliable one; Cronbach's Alpha reliability method was applied to check the

reliability of all factors in the survey and actual data. The reliability coefficient value was highly significant in the combined data set, therefore a model is developed using a combination of survey and actual data (Figure 6.1).

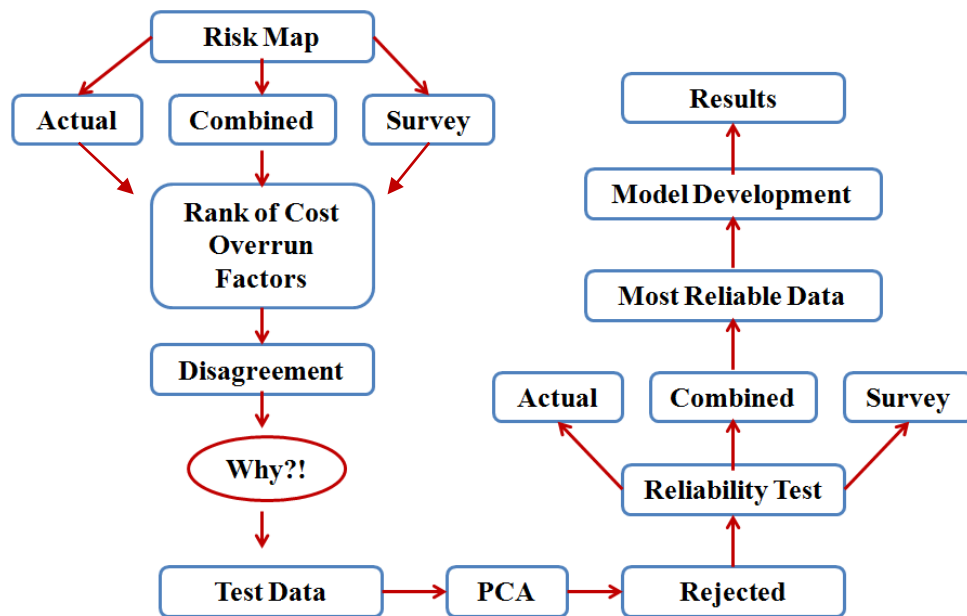


Figure (6.1)

Research Process

6.2 Model Input Data

The study was causal in nature and was aimed at evaluating the impact of different factors on cost overrun in construction projects. The study was conducted using actual and survey data. Cronbach's Alpha reliability was computed using SPSS software to evaluate the reliability of the survey, actual and combined data. The combined sample size of 74 respondents and 40 actual construction projects was selected to conduct the study, since it had the most reliable results. (Appendix 3 contains the reliability tests of survey data and actual data).

6.2.1 Combined data reliability test

Cronbach's Alpha reliability method was applied to check the reliability of all factors in the survey and actual data. The reliability coefficient value was highly significant; it is 0.826 for the combined data as given in Table 6.1.

Table (6.1)
Combined Data Reliability Test

Cronbach's Alpha	No. of Factors
0.826	45

6.2.2 Factor analysis

The Kaiser - Meyer - Olkin Measure of Sampling Adequacy value was 0.734 indicating that the sample was adequate to consider the data as normally distributed. The Bartlett's Test of Sphericity tests the null hypothesis that the factor-to-factor correlation matrix was an identity matrix. The hypothesis was tested through Chi-Square test; the value of Chi-square was found to be 4417.802, which is significant at 0% level of significance. Therefore, null hypothesis was rejected; indicating that the factor-to-factor correlation matrix was not an identity matrix and was therefore suitable for factor analysis. Table 6.2 shows the KMO and Bartlett's test results.

Table (6.2)
KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.734
Bartlett's Test of Sphericity	Approx. Chi-Square	4417.802
	df	990
	Sig.	0.000

6.3 Model Development

Using a combination of survey and actual data, and based on the conceptual framework that has been designed and discussed in Chapter 5, the model can be built.

6.3.1 Building the inner model

Each group of cost overrun related factors (exogenous latent variable) is represented by a red circle, cost overrun (endogenous latent variable) is represented by a blue circle and the groups are the cost overrun indicators (Figure 6.2).

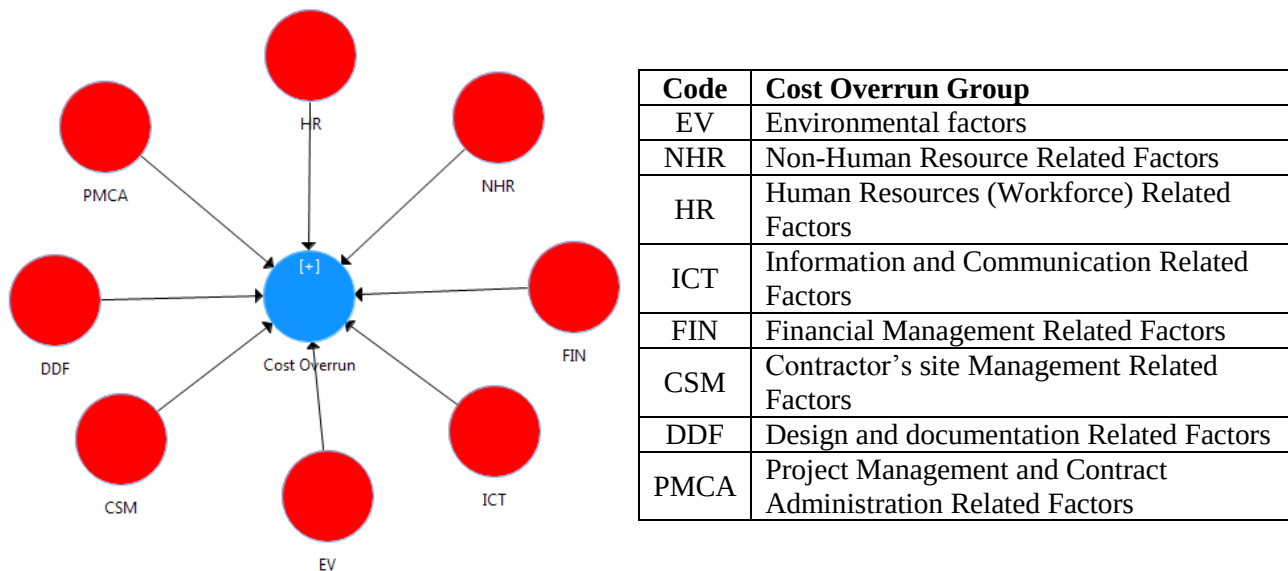


Figure (6.2)
Inner Model (Structural Model) Framework

6.3.2 Building the outer model

The next step is to build the outer model; by linking the indicators (cost overrun factors) to the latent variable (group). Each indicator is represented by a yellow rectangle and the colour of the latent variable will be changed from red to blue when the linkage is established (Figure 6.3).

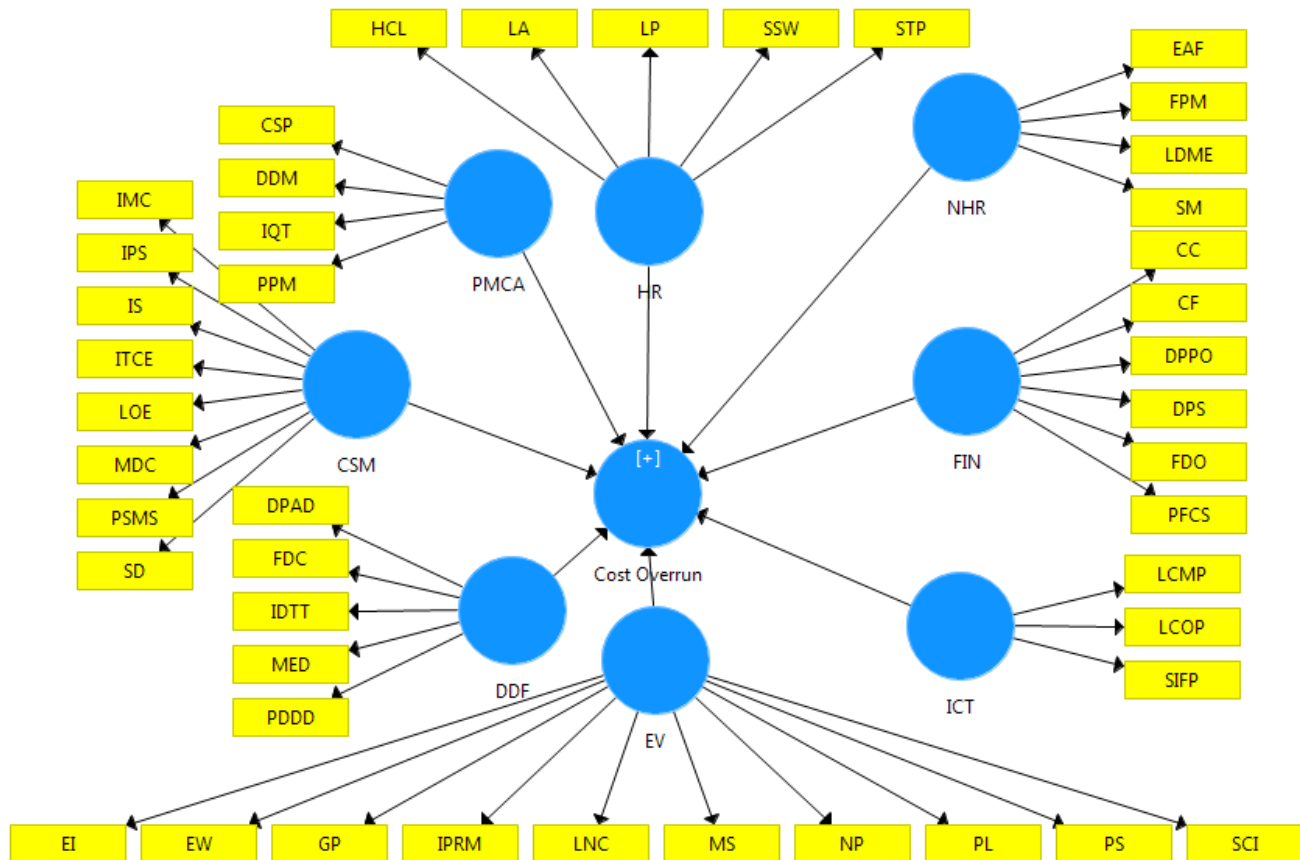


Figure (6.3)

Outer Model (Measurement Model) Framework

6.3.3 Running the path-modelling estimation

Once the indicators and latent variables are linked together successfully; the path modelling procedure can be carried out using the “PLS Algorithm” (Software: smartPLS 3).

In the PLS-SEM diagram (Figure 6.4), there are two types of numbers:

- 1) Numbers in the circle: These show how much the variance of the latent variable is being explained by the other latent variables.
- 2) Numbers on the arrow: These are called the path coefficients. They explain how strong the effect of one variable is on another variable. The weights of different path coefficients enable the ranking of their relative statistical importance (Appendix 4).

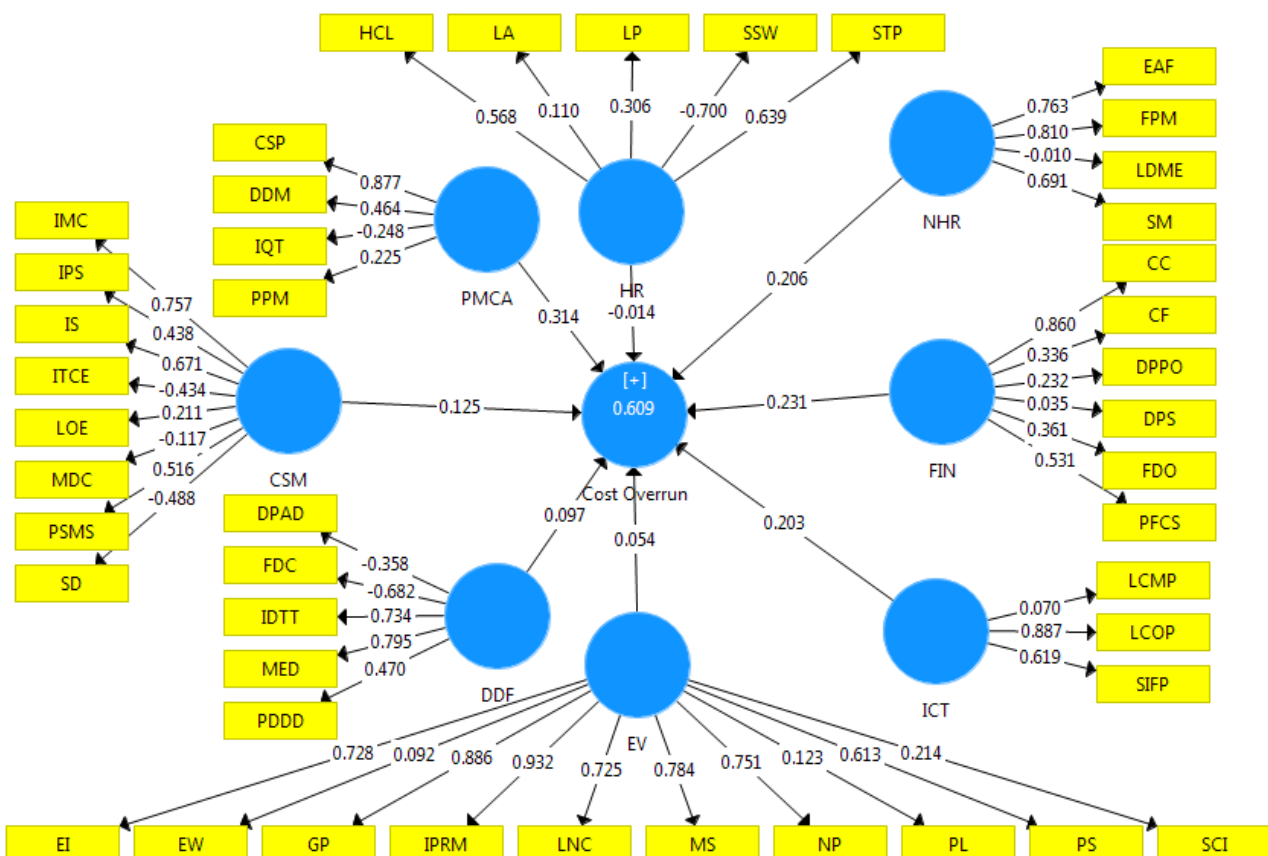


Figure (6.4)
PLS-SEM Results (First Run)

6.3.4 Model assessment

The indicators of cost overrun in construction are correlated and interchangeable, all arrows are pointing outward; which means they are reflective. In a reflective measurement scale the indicators reliability and validity should be thoroughly examined (Haenlein and Kaplan 2004, Hair et al. 2013, Petter et al., 2007).

There are two parts in a PLS path model:

- 1) A measurement model relating the observable variables to their own latent variables, and
- 2) A structural model relating some endogenous latent variables to other latent variables

1) Assessment of outer model (measurement model)

As explained before in Chapter 5, the assessment of the outer model consists of three parts; the individual factor reliability and convergent validity; the discriminant validity; and the assess of significance and relevance of the measurement model relationships

a) Assessing factor reliability and convergent validity of reflective measurement model (first run)

Table 6.3, shows the results of the reflective outer model (measurement model), the first column represents the cost overrun groups (exogenous latent variable), second column shows the related factors of cost overrun in each group, and each factor loading is represented in the third column, the green colour represents a satisfied loading with a value above or equal to 0.400, the yellow colour indicates that loading of the factor is below the extinguished value 0.400. Last three columns represent the average variance extracted, the composite reliability and Cronbach's Alpha.

Table (6.3)**Results Summary for First Reflective Outer Models (individual factor reliability and convergent validity)**

Construct	Factor	Loading	AVE	CR	Alpha
Project Management and Contract Administration Related Factors	PPM	0.225	0.274	0.374	-0.642
	CSP	0.877			
	DDM	0.464			
	IQT	-0.248			
Design and documentation Related Factors	FDC	-0.682	0.397	0.234	-0.316
	MED	0.795			
	IDTT	0.734			
	PDDD	0.470			
	DPAD	-0.358			
Contractor's site Management Related Factors	PSMS	0.516	0.246	0.286	0.109
	IPS	0.438			
	PSMS	-0.488			
	IMC	0.757			
	LOE	0.211			
	ITCE	-0.434			
	MDC	-0.117			
	IS	0.671			
Financial Management Related Factors	CF	0.336	0.220	0.542	0.113
	PFCS	0.531			
	FDO	0.361			
	DPPO	0.232			
	DPS	0.035			
	CC	0.860			
Information and Communication Related Factors	LCOP	0.887	0.391	0.576	0.159
	SIFP	0.619			
	LCMP	0.070			
Human Resources Related Factors	LP	0.306	0.265	0.188	0.049
	SSW	-0.700			
	HCL	0.568			
	LA	0.110			
	STP	0.639			
Non-Human Resource Related Factors	FPM	0.810	0.429	0.690	0.454
	SM	0.691			
	EAF	0.763			
	LDME	-0.010			
Environmental factors	LNC	0.725			
	EI	0.728			
	EW	0.092			
	GP	0.086			

	IPRM	0.932	0.433	0.858	0.851
	MS	0.784			
	PS	0.613			
	PL	0.123			
	NP	0.751			
	SCI	0.214			

b) Individual factor reliability

It can be seen (Table 6.3) that not all of the indicators have individual indicator reliability values that are larger than the minimum acceptable level of 0.4 (Hulland, 1999); therefore these indicators have to be eliminated.

c) Convergent validity

Several tests were used to determine the convergent validity:

- Composite Reliability scores (CR), for a good model is 0.7; however the minimum acceptable value falls down to 0.5, from Table 6.3, half of the scores are not acceptable.
- A good Cronbach's alpha value is 0.6 or higher; all values except for environmental factors groups are not acceptable.
- Average Variance Extracted (AVE), must be 0.5 at least; therefore all values are not acceptable.

The current model is not reliable or valid; therefore another model needs to be established after eliminating of unreliable indicators with the yellow colour in Table 6.3, these indicators are: PPM, IQT, FDC, DPAD, PSMS, LOE, ITCE, MDC, CF, FDO, DPPO, DPS, LCMP, LP, SSW, LA, LDME, EW, PL, and SCI .

6.3.5 Rerunning of the path-modelling estimation (second run)

Table 6.3 shows loadings for all reflective indicators (factors of cost overrun) in the model, some factors were eliminated after assessment of their practical potential significance prior to elimination. If an indicator's reliability is low and eliminating this indicator goes along with a substantial increase of composite reliability, it makes sense to discard this indicator. A new model with the remaining indicators needs to be introduced by rerunning of the path model estimation (Figure 6.5).

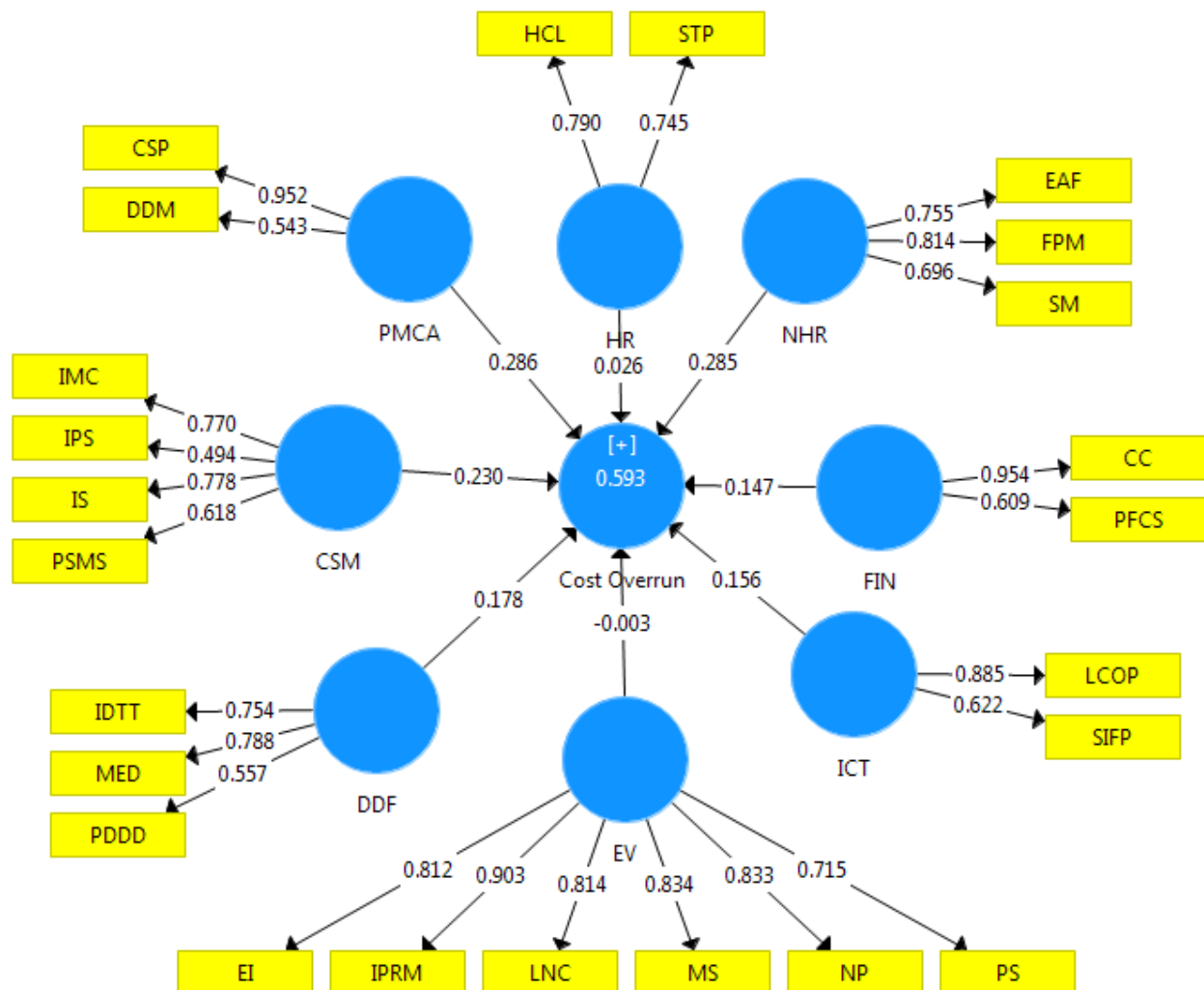


Figure (6.5)

PLS-SEM Results (Second Run)

Table 6.4 shows the results of the second iteration of the individual factor reliability and convergent validity measures.

Table (6.4)
Results Summary for the Second Reflective Outer Models (individual factor reliability and convergent validity)

Construct	Factor	Loading	AVE	CR	Alpha
Project Management and Contract Administration Related Factors	CSP	0.952	0.601	0.737	0.512
	DDM	0.543			
Design and documentation Related Factors	MED	0.788	0.500	0.746	0.487
	IDTT	0.754			
	PDDD	0.557			
Contractor's site Management Related Factors	PSMS	0.618	0.556	0.765	0.616
	IPS	0.494			
	IMC	0.770			
	IS	0.778			
Financial Management Related Factors	PFCS	0.885	0.641	0.773	0.512
	CC	0.622			
Information and Communication Related Factors	LCOP	0.885	0.585	0.732	0.414
	SIFP	0.622			
Human Resources Related Factors	HCL	0.790	0.590	0.742	0.404
	STP	0.745			
Non-Human Resource Related Factors	FPM	0.814	0.572	0.800	0.636
	SM	0.696			
	EAF	0.755			
Environmental factors	LNC	0.814	0.673	0.925	0.916
	EI	0.812			
	IPRM	0.903			
	MS	0.834			
	PS	0.715			
	NP	0.833			

a) Assessing factor reliability and convergent validity of reflective measurement model

It can be seen (Table 6.4) that all indicators have individual indicator reliability values that are much larger than the minimum acceptable level of 0.4 and close to the preferred level of 0.7. From Table 6.4, Cronbach's alpha values are shown to be larger than 0.4, so moderate to high levels of internal consistency reliability have been demonstrated among all eight reflective latent variables.

To check convergent validity, each latent variable's Average Variance Extracted (AVE) is evaluated. Again from Table 5.3, it is found that all of the AVE values are greater than the acceptable threshold of 0.5, so convergent validity is confirmed. The values of the Average Variance Extracted values of latent variable in the tested model vary from 0.500 to 0.673; which exceeded the cut off value of 0.500. Thus, the constructs are considered satisfactory with the evidence of adequate reliability, convergent validity.

b) Discriminant validity

As a rule for valid discriminant of construct is that the square root of the AVE of each construct should be larger than the correlation of two constructs (Chin, 1998). This is done by replacing the diagonal of correlation matrix with the value of square root of the AVE and the diagonal elements must be greater than the off-diagonal elements in the corresponding rows and columns (Hulland, 1999). Latent variable correlations for this study are calculated with SmartPLS software and are shown in Table 6.5. The values of square root of AVE in the diagonals of matrix are higher than off-diagonal values in the model. This confirms that all the variables represent their constructs and the discriminant validity is well established.

Table (6.5)
Checking Discriminant Validity

	CSM	DDF	EV	FIN	HR	ICT	NHR	PMCA
CSM	0.745							
DDF	-0.020	0.707						
EV	0.729	-0.136	0.820					
FIN	0.231	0.342	0.347	0.800				
HR	0.595	-0.334	0.540	0.095	0.768			
ICT	0.333	0.171	0.292	0.005	0.240	0.765		
NHR	0.588	0.170	0.726	0.370	0.294	0.338	0.756	
PMCA	-0.110	0.475	-0.247	0.312	-0.205	-0.026	0.082	0.775

The construct measures have been confirmed as reliable and valid, the next step is to assess the structural model results.

c) Assessing significance and relevance of the measurement model relationships using bootstrapping

PLS-SEM does not assume that the data is normally distributed, which implies that parametric significance tests (e.g., as used in regression analyses) cannot be applied to test whether coefficients such as outer weights, outer loadings and path coefficients are significant. Instead, PLS-SEM relies on a nonparametric bootstrap procedure (Efron and Tibshirani, 1986; Davison and Hinkley, 1997) to test the significance of estimated coefficients in PLS-SEM. In bootstrapping, subsamples are created with randomly drawn observations from the original set of data (with replacement). The subsample is then used to estimate the PLS path model. This process is repeated until a large number of random subsamples has been created.

The parameter values (e.g., outer weights, outer loadings and path coefficients) estimated from the subsamples are used to derive standard errors for the estimates. With this information, t-values are calculated to assess each estimate's significance. From Table 6.6, outer weights for two factors from the environmental group (PS

and NP) are very low; therefore their significance must be tested using bootstrapping. The t-statistics from Table 6.7 show that all formative indicators are significant, all t values are above 2.58, which indicate the significance of their outer weights ($p < .001$). Therefore, the model is significant, and there is no need to eliminate any more factors. Bootstrapping results for the outer model are shown in Figure 6.5.

Table (6.6)

Indicators (Cost Overrun Factors) Outer Weights (Relative Importance)

	CSM	DDF	EV	FIN	HR	ICT	NHR	PMCA
CC				0.845				
CSP								0.869
DDM								0.318
EAF							0.382	
EI			0.165					
FPM							0.404	
HCL					0.678			
IDTT		0.540						
IMC	0.463							
IPRM			0.443					
IPS	0.234							
IS	0.262							
LCOP						0.797		
LNC			0.208					
MED		0.464						
MS			0.247					
NP			0.071					
PDDD		0.409						
PFCS				0.318				
PS			0.044					
PSMS	0.525							
SIFP						0.474		
SM							0.55	
STP					0.623			

Table (6.7)**Mean, SD, t-Statistics and P Values of Path Coefficients (Outer Model)**

	Original Sample	Sample Mean	Standard Deviation	t-Statistics	P Values
CC --> FIN	0.954	0.941	0.078	12.279	0.000
CSP --> PMCA	0.952	0.945	0.044	21.606	0.000
DDM --> PMCA	0.543	0.526	0.15	3.62	0.000
EAF --> NHR	0.755	0.745	0.079	9.6	0.000
EI --> EV	0.812	0.8	0.056	14.493	0.000
FPM --> NHR	0.814	0.804	0.067	12.151	0.000
HCL --> HR	0.79	0.746	0.2	3.954	0.000
IDTT --> DDF	0.754	0.736	0.103	7.329	0.000
IMC --> CSM	0.770	0.762	0.071	10.926	0.000
IPRM --> EV	0.903	0.905	0.024	36.867	0.000
IPS --> CSM	0.494	0.489	0.0101	4.867	0.000
IS --> CSM	0.778	0.768	0.067	11.675	0.000
LCOP --> ICT	0.885	0.877	0.07	12.57	0.000
LNC --> EV	0.814	0.807	0.052	15.803	0.000
MED --> DDF	0.788	0.770	0.078	10.053	0.000
MS --> EV	0.834	0.823	0.054	15.564	0.000
NP --> EV	0.833	0.819	0.057	14.525	0.000
PDDD --> DDF	0.557	0.557	0.184	3.024	0.003
PFCS -->FIN	0.609	0.582	0.221	2.759	0.006
PS -->EV	0.715	0.698	0.097	7.377	0.000
PSMS--> CSM	0.618	0.619	0.099	6.224	0.000
SIFP--> ICT	0.622	0.611	0.153	4.078	0.000
SM -->NHR	0.696	0.698	0.069	10.098	0.000
STP--> HR	0.745	0.711	0.213	3.504	0.000

***Critical t-values for a two-tailed test are: <1.96 (p >.05), 1.96 (p = .05), and 2.58 (p = .001).**

Figure (6.6) shows the t-statistics for outer model, in bootstrapping, subsamples are created with observations randomly drawn from the original set of data (with replacement). To ensure stability of results, the number of subsamples should be large, and it was equal to 5,000 subsamples.

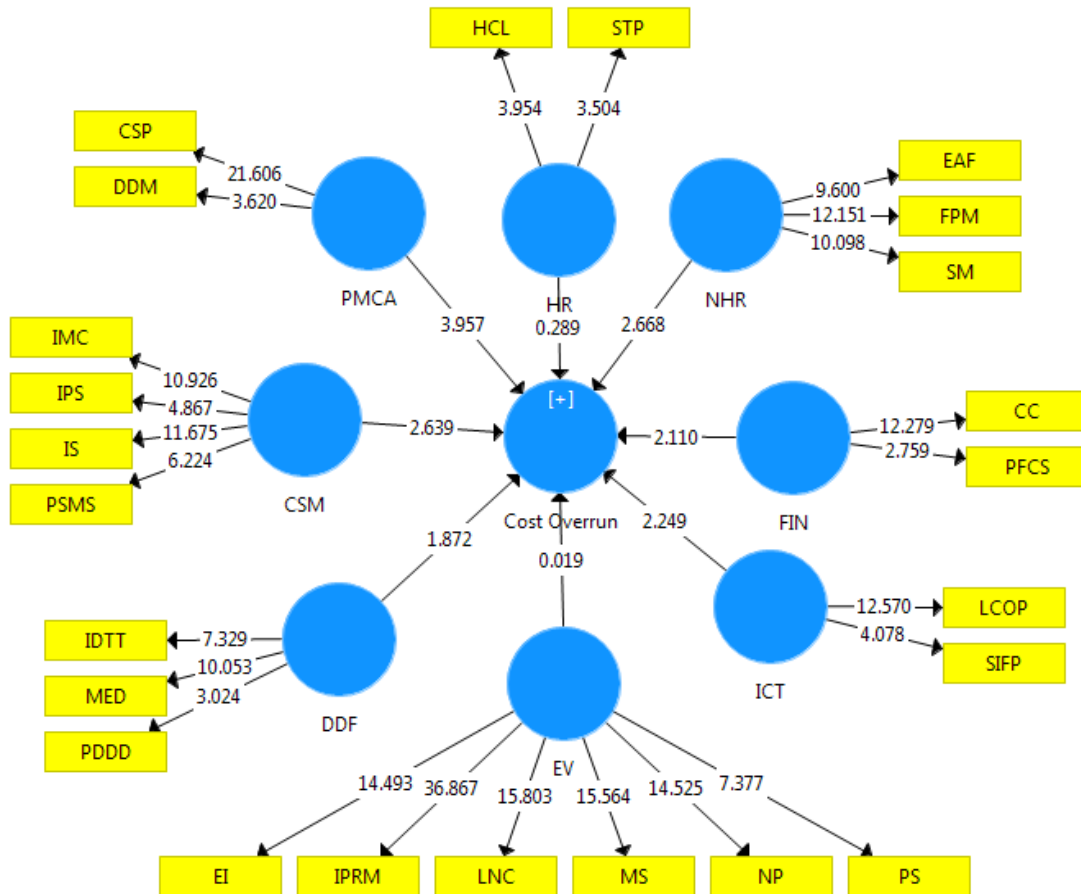


Figure (6.6)

Bootstrapping Results, the t-Statistics for Outer Model

6.3.6 Assessment of inner model (structural model)

For the reason that PLS does not require a normally distributed data it is evaluated with R-squared calculation for dependent latent variables (Cohen, 1992). The first characteristic that PLS provides to determine how well the model fits the hypothesized relationship is the squared multiple correlations (R^2) for each dependent construct in the model. The R^2 measures a construct's percent variation that is explained by the model (Wixom and Watson, 2001). The quality of the structural model for each endogenous block can be assessed by the Redundancy

index. Redundancy is the “capacity of the model to predict its manifested variables from the indirectly connected latent variables” (Chantelin, et.al. 2002). Since the objective of PLS is to maximize variance explained rather than fit, therefore prediction-oriented measures such as R^2 are used to evaluate PLS models (Chin, 1998).

6.3.7 Assessing the level of R^2 and path coefficient significance

The PLS path modelling estimation is shown in Figure 6.4 By looking at the diagram, we can make the following preliminary observations:

- a) Explanation of target endogenous variable (cost overrun) variance:

The coefficient of determination, R^2 is 0.593 for the cost overrun endogenous latent variable. This means that the 8 latent variables (groups) moderately explain 59.3% of the variance in cost overrun. This proves that the developed model has substantial explaining power to represent the relationship of construction groups of cost overrun towards cost overrun

- b) Inner model path coefficient sizes and significance:

The inner model suggests that project management and contract administration has the strongest effect on cost overrun (0.286), followed by Non-Human Resource (0.285), Contractor’s site Management (0.230), Design and Documentation (0.178), Information and Communication (0.156), Fin (0.147), Human Resources (0.026) and Environmental (-0.003). The hypothesized path relationship between (PMCA, NHR, CSM, DDF, ICT, and Fin) and cost overrun is statistically significant. The hypothesized path relationship between (HR and EV) and cost overrun is not statistically significant. This is because their standardized path coefficient (0.026 and -0.003) is lower than 0.1. Thus, it can be concluded that: (PMCA, NHR, CSM, DDF, ICT, and Fin) are

moderately strong predictors of cost overrun, but (HR and EV) do not predict cost overrun directly.

6.3.8 Assessing structural model relationships significance and relevance

According to Chin's (1998) recommendations, a bootstrapping procedure using 5000 sub samples was performed to evaluate the statistical significance of each path coefficient. Table 6.8 shows hypothesized path coefficients along with their bootstrap values and t-values. Using a two-tailed t-test with a significance level of 5%, the path coefficient will be significant if the t-statistics is larger than 1.96. From Table 6.8 sixth column "t-Statistics", it can be seen except for EV -Cost Overrun and HR - Cost Overrun linkages all "t-Statistics" values are significant; therefore the path coefficients of the inner model are significant.

Table (6.8)

Mean, SD, T-Statistics and P Values of Path Coefficients (Inner Model)

Hypothesis	Hypothesis Path	Original Sample (β)	Sample Mean	Standard Deviation	t-Statistics	Accept/Reject significance
H1	CSM → Cost Overrun	0.230	0.233	0.087	2.639	Accept
H2	DDF → Cost Overrun	0.178	0.178	0.095	1.872	Accept
H3	EV → Cost Overrun	-0.003	0.012	0.138	0.019	Reject
H4	FIN → Cost Overrun	0.147	0.144	0.07	2.11	Accept
H5	HR → Cost Overrun	0.026	0.018	0.091	0.289	Reject
H6	ICT → Cost Overrun	0.156	0.155	0.07	2.249	Accept

H7	PMCA → Cost Overrun	0.286	0.279	0.068	4.221	Accept
H8	NHR → Cost Overrun	0.285	0.278	0.107	2.668	Accept

Figure (6.7) shows the t-statistics for inner model, the bootstrapping was used to test the significance of estimated path coefficients, and t-values were calculated to assess each estimate's significance.

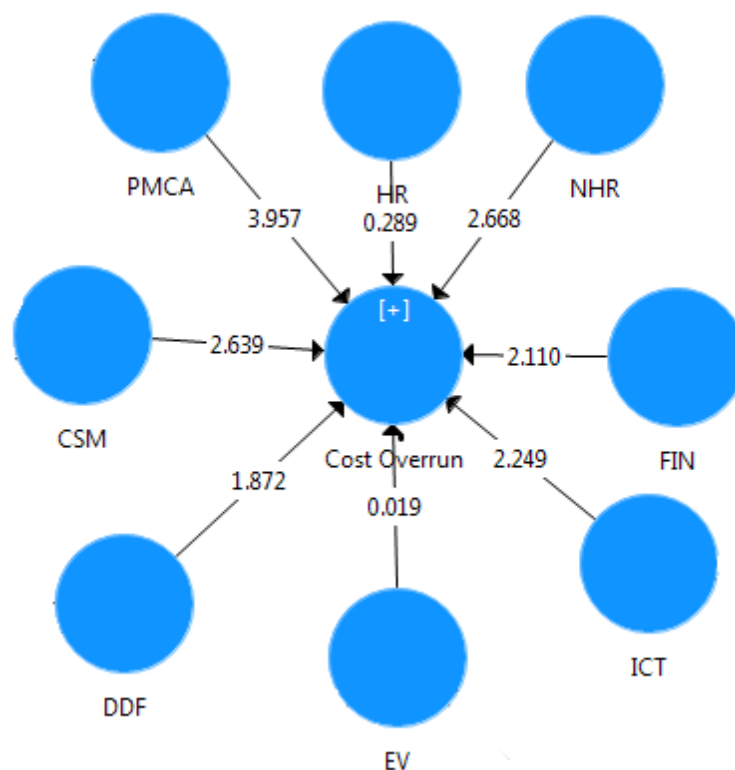


Figure (6.7)

Bootstrapping Results - Path Coefficients for Inner Model

The relationship between human resources related factors and cost overrun was insignificant with a proportionality coefficient of 0.026 and t value of 0.289; indicating that the human resources factors has direct positive insignificant influence on the cost overrun. The cost overrun changes in direct proportion to human resources factors with a coefficient of 0.026. This clearly indicates that a

100 points change in human resource will bring only 2.6 points change in the cost overrun in a construction project.

The relationship between environmental related factors and cost overrun in construction projects was insignificant with a proportionality coefficient of -0.003 and t value of 0.019; indicating that the environmental related factors has direct negative insignificant influence on cost overrun. The cost overrun changes in direct proportion to environmental factors with a coefficient of -0.003. This clearly indicates that a 100 points change in environmental factors will bring 0 points change in the total cost overrun in a construction project. It means that the environmental factors don't have the strong impact on construction cost overrun in Bahrain.

The relationship between project management and contract administration related factors was significant with a proportionality coefficient of 0.286 and t value of 3.957; indicating that the project management and contract administration related factors has direct positive influence on cost overrun. The cost overrun changes in direct proportion to project management and contract administration related factors with a coefficient of 0.286. This clearly indicates that a 100 points change in project management and contract administration related factors will bring 28.6 points change in the construction cost overrun in Bahrain.

6.3.9 Summary

Out of the eight paths used to connect the measures in the structural model, six paths were supporting the hypothesis that (PMCA, NHR, CSM, DDF, ICT, and FIN) have a direct positive influence on the cost overrun in construction projects.

However two path coefficients between environmental related factors, human resources related factors and construction cost overrun don't support the hypothesis; because human resources related factors has a very little impact on cost overrun and environmental related factors does not have an impact on cost overrun at all. It can be concluded that around 28.60%, 28.50%, 23.00%, 17.80%, 15.60%, 14.70%, 2.60%, -0.03% of change in cost overrun can be explained by Project Management and Contract Administration, Non-Human Resource, Contractor's site Management, Design and documentation, Information and Communication, Financial Management, Human Resources, and Environmental factors, respectively.

6.4 Model Evaluation

Since PLS makes no distributional assumptions for parameter estimation. The evaluation of PLS model is therefore, based on prediction-oriented measures that are non-parametric (Chin, 1998).

6.4.1 Goodness of fit

The PLS model is mainly evaluated by Goodness-of-Fit (GoF) (Tenenhaus et al., 2005). GoF was employed to judge the overall fit of the model, it is the geometric mean of the average communality and the average R^2 . It represents an index for validating the PLS model globally, as looking for a compromise between the performance of the measurement and the structural model, respectively. Thus, GoF is given by:

$$GoF = \sqrt{\overline{AVE} \times \overline{R^2}} \quad (5.4)$$

And the calculations are:

$$\overline{AVE} (0.601+0.500+0.556+0.641+0.585+0.590+0.572+0.673)/8= 0.5897$$

$$\bar{R}^2 = 0.593$$

$$\begin{aligned} GoF &= \sqrt{0.5897 \times 0.593} \\ &= \sqrt{0.3497} = 0.591 \end{aligned}$$

For this model the GoF index was 0.591, which exceeds the cut-off value in comparison of baseline value as $GoF_{small} = 0.1$, $GoF_{medium} = 0.25$, $GoF_{large} = 0.36$. This shows that the model has substantial explaining power.

6.4.2 Assessing effect size – f^2

It measures the impact of a specific group on cost overrun. It is used to evaluate whether the omitted group has a substantive impact on the R^2 values of the cost overrun or not. Table 6.9 shows an effect size value of 0.124 for PMCA group, which indicates a close to medium effect in producing the R^2 for cost overrun. However, no effect at all is represented by the f^2 effect size for the predictive value of EV group on cost overrun and this can be explained by the insignificant effect of the EV group related factors on cost overrun. All other values indicate a small effect in producing the R^2 for cost overrun.

Table (6.9)
Summary of f^2 Effect Sizes

Group	f^2 values	Effect Size
CSM	0.049	Small
DDF	0.042	Small
EV	0.000	No effect
FIN	0.033	Small
HR	0.001	Small

ICT	0.047	Small
NHR	0.072	Small
PMCA	0.124	Medium

6.4.3 Assessing the predictive relevance Q^2

Further the quality of path model can also be evaluated by calculating the Q - square statistic. The Q square statistic measures the predictive relevance of the model by reproducing the observed values by the model itself. A Q square greater than zero means the model has predictive relevance; whereas Q-square statistic less than zero means that the model lacks predictive relevance (Fornell and Larcker, 1981). The cross-validated redundancy approach was used; the Q- square statistics are estimated by using Blindfold method of calculations.

$$Q^2 = 1 - (SSE/SSO) \quad (6.1)$$

Where:

SSO: Sum of the squared observation

SSE: Sum of the squared prediction error

	SSE	SSO	Q^2
Cost Overrun	47.768	100.000	0.522

In this model, cost overrun is predicted by 8 latent variables, in which all of them are reflective constructs, the study obtains a Q^2 of 0.522 which is an indicative of a highly predictive model.

6.5 Model Contribution

The contributions of this research to the field of construction project management are closely aligned with the research aims described previously. The model can be used in construction management as a key business tool to improve construction performance, through determining the risks which are present during the delivery of construction projects in Bahrain that lead to significant project cost overruns, and explain how each risk correlates with the total amount of cost overrun. This could essentially help construction firms to transform their data into cutting edge decision support systems for business improvement and gain competitive advantage. The results can also be applied to the problem of final cost estimation of construction projects, through allocation of contingency percentages, based on the effect of each risk on the total cost overrun.

Table 6.10, not only can be used to prioritize grouping of risks, but also aimed at simply providing a better understanding of the contingency management of a project, thus validating better risk contingency weightings in budget estimates.

Table (6.10)
Groups of Cost Overrun Effects on Total Project Cost Overrun (Model Results)

S.N.	Group of cost overrun	Effects on cost overrun %	Contingency %
1	Project Management and Contract Administration Related Factors	28.60	21.87
2	Non-Human Resource Related Factors	28.50	21.79
3	Contractor's site Management Related Factors	23.00	17.58
4	Design and documentation Related Factors	17.80	13.61
5	Information and Communication Related Factors	15.60	11.93
6	Financial Management Related Factors	14.70	11.24
7	Human Resources (Workforce) Related Factors	2.60	1.98
8	Environmental factors	-0.03	0.00

6.6 Risk Mapping and PLS-SEM Model Results Comparison

After conducting the analysis using both risk mapping and PLS_SEM modelling methods, results in Table (6.11) and Table (6.12) showed agreement on most of the groups, while noticeable disagreement was found in the group of Non- Human resources related factors. The findings show 100% consensus among three groups; information and communication, financial management, and environmental related factors. Semi consensus among contractor's site management, project management and contract administration, human resources, and design and documentation related factors, while the last group non-human resources demonstrated remarkable difference which can be explained by the following:

1. Some factors of cost overrun were present in the risk mapping and absent in the model. Factors with low loadings were reviewed and dropped, since they would add very little explanatory power to the model, and therefore biasing the estimates of the parameters linking the latent variables.
2. The model considers the consistency of indicators for each group, which is how they are consistent with each other, and how well they coordinate within the group. Then it calculates the effect of each group on the total cost overrun, which makes the results more accurate and reliable. On the other hand the risk mapping considers the influence of each factor on cost overrun directly, and then it takes the average influence of each group related factors to find the group effect on the total cost overrun.

Table (6.11)

**Cost Overrun in Construction Projects in Bahrain: Risk Mapping and
PLS-SEM Model Groups Ranking**

#	Risk Map Groups Ranking	Model Groups Ranking
1	Design and documentation Related Factors	Project Management and Contract Administration Related Factors
2	Contractor's site Management Related Factors	Non-Human Resource Related Factors
3	Project Management and Contract Administration Related Factors	Contractor's site Management Related Factors
4	Human Resources (Workforce) Related Factors	Design and documentation Related Factors
5	Information and Communication Related Factors	Information and Communication Related Factors
6	Financial Management Related Factors	Financial Management Related Factors
7	Non-Human Resource Related Factors	Human Resources (Workforce) Related Factors
8	Environmental factors	Environmental factors

Table (6.12)

**Cost Overrun in Construction Projects in Bahrain: Risk Mapping and
PLS-SEM Model Groups Matching**

#	Group Name	Matching
1	Design and documentation Related Factors	62.5%
2	Contractor's site Management Related Factors	87.5%
3	Project Management and Contract Administration Related Factors	75%
4	Human Resources (Workforce) Related Factors	75%
5	Information and Communication Related Factors	100%
6	Financial Management Related Factors	100%
7	Non-Human Resource Related Factors	37.5%
8	Environmental factors	100%

CHAPTER SEVEN

CONCLUSION AND RECOMMENDATIONS

7.1 Conclusion

7.2 Recommendations

7.3 Research Limitations

7.4 Further Research

CHAPTER SEVEN

CONCLUSION AND RECOMMENDATIONS

The study investigated factors of cost overruns in construction projects to improve risk assessment and contingency allocation process. This chapter provides some recommendations for the construction project management regarding cost overrun, and present some suggestions for further research.

7.1 Conclusion

7.1.1 Cost overrun risk mapping

The research identified 45 initial risk factors of cost overrun from the literature review. Frequency and severity indices of each factor were calculated, and the factor risk zone was categorized accordingly. The survey risk mapping indicates that only 2 factors were located in the green zone, 23 factors were located in the yellow zone, and 20 factors were located in the red zone of the risk map. The actual data risk mapping, had 13 factors in the green zone, 10 in the yellow zone and 22 were located in the red zone. The last risk mapping was for the combined results and it showed 10 factors in the green zone, 19 in the yellow zone and 16 factors were located in the red zone of the risk map. The results indicated that the top five affecting factors were: frequent design changes, mistakes during construction, schedule delay, poor site management and supervision, and inaccurate quantity take-off.

Furthermore, in comparison to other similar studies conducted in Arabian Gulf countries, results of the current study showed agreement with Alghonamy (2015) from Saudi Arabia regarding the cost overrun cause “frequent design changes ”. While another study in Oman showed agreement in some factors such as schedule delays (Amoudi and Al Brashdi, 2016). It is vitally important to note that major factors causing cost overruns differ greatly among countries. Each country possesses its own divine characteristics and its own construction projects environment which indicates that the impact of each of the previously listed factors differ greatly between countries. Although most Arabian Gulf countries possess similar characteristics in their construction environment, they still hold opposing views regarding factors affecting the cost overruns.

7.1.2 Factor analysis

Factor analysis was initially adopted to statistically isolate the large number of project risks that had relationships to cost overrun in the construction projects. The factor analysis reduced the number of independent cost overrun factors to 13.

KMO sample testing found that the data sample used in the research was less than that recommended by Kaiser (1974). However; the Bartlett Test was good, with higher value than the one recommended in the research case. Factor variations were relatively small (except for one factor) across all factors. This indicated low success in reducing the dimensionality of the original set of cost overrun variables using this method, therefore a rotation of factors was necessary.

The rotation of axes was done with the Varimax method of rotation. The Varimax rotation differentiated the factor loadings of each variable, by minimizing the number of variables that had high loadings on the factors. There was no evidence of strong correlations within any of the 13 groupings, even after rotation. A weak scree plot also indicated that the derived model of 13 factors was not a reliable representation of the total factors to be used in the factors ranking or the model analysis.

7.1.3 Model development using PLS-SEM method

The study adopted structural equation modelling to examine the effects of different risks on project cost, and to provide a causal relationship model for construction cost overrun related factors. With the 45 factors identified from literature, which were then reduced to 24 factors a theoretical model demonstrating how these factors affect cost overrun was developed. The model was constructed using SEM-PLS. Results showed that approximately 60% of cost overrun was influenced with the identified factors. The Global Fit Index (GoF) value of the developed model was 0.591, indicating that the model is adequate to generalize the phenomenon of Bahrain construction industry. “Project management and contract administration” related factors were found as most dominant factors causing cost overrun, followed by the “non human resources” related factors; such as materials and equipments.

This study undertakes the analysis of risks of cost overrun to ascertain if the findings can lead to more accurate budget estimates in the future by making more realistic allowances in estimates of the identified cost overrun risks, the SEM-PLS model would help in successfully managing the cost program

by providing a strong risk management plan and better risk contingency weightings in budget estimates. Furthermore it would prioritize cost overrun associated groups of risks based on their overall effects on the project cost overrun. This makes the risk assessment process valid. If every factor that poses a potential risk to the project was considered a high-priority, then nothing is a high priority because all risks will have an equal effect on the project cost overrun. Despite the best laid plans, unforeseen conditions, design errors and omissions, and scope revisions will occur in any construction project. Therefore, early planning for project risks can help reducing the undesirable effects on project cost. The model also aims to open a door for further research in Bahrain to develop relationships between the measures of the contingency percentage and construction project variables.

7.2 Recommendations

In order to complete construction projects within the budget, the following recommendations are specified to construction practitioners:

- Consider the contractor financial status and level of experience before awarding the job. Contractors with strong financial status have the ability to do their duties on time, on budget, and with required quality.
- Hire consultants with good technical knowledge and professional engineers, this help owners and contractors to run their projects with no or low cost overrun.

- Keep unqualified and inexperienced staff under supervision, unskilled employees imposes slow progress of work, thus increases the cost overrun chances of occurrence.
- Use managers with high level of work experience and knowledge on managing the critical conditions. Good managers have the ability to run the project successfully with cost limitations.
- Hire suitable number of labour with standard performance and high productivity, this prevents cost overrun and avoids slow progress of works.
- Use of Risk Assessment and cost control when estimating construction projects will help decision makers define unforeseen situations more reliably ahead of time, so that corrective measures can be better taken into account in project design and estimating.
- Prioritize of cost overrun factors in projects leads to better risk contingency weightings in budget estimates.
- Evaluate total cost carefully before undertaking a construction project contract. A contract price should not be over the financial ability of the company. Any financial problem in the project expenditures and payments will cause delay and cost overrun accordingly.
- Purchase of equipment and checking for material availability before beginning any construction activities is preferred. Shortage of equipment or materials always leads to delay and therefore cost overrun.
- The tight control of any construction project can limit variation in works during construction, and this is absolutely necessary for successful financial outcomes of projects.

7.2.1 Top cost overrun factors recommendations

Estimating procedures for construction projects should take into account the potential cost impacts resulting from each of the risk factors identified in this research. Failure to systematically evaluate the project risk potential of all of the risk factors could have severe impact on the favourable outcomes of projects and on the overall construction program. The research found that “frequent design changes” was the highest contributing risk factor to project cost overrun in construction projects. Therefore, to further manage cost overruns from design changes the following measures need to be followed:

1. Consider a detailed feedback from similar previously constructed projects.
2. Assure site surveys and measurements.
3. Involve construction and maintenance engineers as a part of the design team from the start to eliminate most of the changes.

The cost overrun developed model showed that shortage of materials, fluctuation of materials prices and equipment availability and failure, are good indicators of “non human resources” group. Construction managers should not overlook these basic elements of day-to-day operation because NHR has shown a significant influence on cost overrun.

The change in the scope of the project is good indicator of “project management and contract administration” group. PMCA has shown a significant influence on cost overrun. As construction progresses, the project manager may become aware of items overlooked in the initial design that seem desirable to be included now. Required field changes to add new

details or reconcile unforeseen conflicts or incompatibilities with the initial design often prove more expensive than had they been included in the original scope. Such changes will delay and cost the construction progress. It is better to find these deficiencies or conflicts early, so it's crucial for the project manager to make a thorough review of the contractor's drawings, specifications and procurement documents. Any problems found early provide significant savings in cost and time as well.

7.3 Research Limitations

During the program of this research, some possible limitation were identified and discussed as follow:

7.3.1 Model limitations

Studying the relationship of multiple variables is especially difficult because the desired knowledge is complex in nature and many different combinations of factors need to be examined. There are a number of ways that statistical techniques can be mis-applied to problems in the real world. These types of errors can lead to invalid or inaccurate results. For structure-activity correlation, Partial Least Squares (PLS) has many advantages over regression, including the ability to robustly handle more descriptor variables than compounds, and non-orthogonal descriptors, while providing more predictive accuracy and a much lower risk of chance correlation. The major limitations are a higher risk of overlooking 'real' correlations and sensitivity to the relative scaling of the descriptor variables.

7.3.2 Generalization limitations

This research was confined to the study of building construction projects in the kingdom of Bahrain. Since construction costs can be specific to geographical/economic areas and time periods, caution is required when comparing the data and results to another situation.

7.3.3 Data limitations

Data consisted of actual projects data and structured survey. Survey response bias is a general term for a wide range of cognitive biases that influence the responses of participants away from an accurate or truthful response. It can be induced or caused by a number of factors; all relating to the idea of human subjects, and can have a large impact on the validity of surveys. The actual projects data was complemented with the use of documentation and interviews, still the quality of the research could have been improved with studying a larger number of actual projects, and more people, who had worked on those projects, could have been interviewed. The use of more sources could have improved the quality of the research. However, this was not done due to time constraints and the difficulty of finding companies willing to participate in this topic.

7.4 Further research

The literature research has shown that contingency percentages of around 5-10% are commonly allocated to project estimates to allow for unforeseen and risks. Based on the research prioritizing and percentages, a project manager can indicate if a higher order percentage in contingency may be required for his project. The random dramatic increase in the percentage of project contingency across construction programs will have certain negative effects. Therefore, further research can be undertaken to identify if any project variables (Table 7.1) have a relationship to the accuracy of project cost contingency. Any project variables that are found to have a relationship can be used to predict more accurate project cost contingencies. After the contingency percentage is established for a certain construction project, the allocation of contingencies in Table 6.10 could be applied. There are many project variables that can be used to determine the total contingency percentage to the estimated cost, such as project location, work volume, dilapidation level, client type, project duration, and project type.

References

Abd El-Razek, M.; Bassioni, H.; and Mobarak, A. (2008). Causes of Delay in Building Construction: Projects in Egypt. *Journal of Construction Engineering and Management-ASCE*, 134(11), 831–841.

Aibinu, A.A.; and Al-Lawati, A.M. (2010). Using PLS-SEM Technique to Model Construction Organizations' Willingness to Participate in E-bidding. *Automat. Constr.*, 19, 714-724.

Aibinu, A.A.; Ling, F.Y.Y.; and Ofori, G. (2011). Structural Equation Modelling of Organizational Justice and Cooperative Behaviour in the Construction Project Claims Process: Contractor's.

Akter, S.; D'Ambra, J.; and Ray, P. (2011). An Evaluation of PLS Based Complex Models: The Roles of Power Analysis, Predictive Relevance and GoF Index. In *Proceedings of the Seventeenth Americas Conference on Information Systems*, Detroit, Michigan August 2011.

Amoudi, O.; and Al Brashdi, L. (2016). World Academy of Science, Engineering and Technology International Journal of Civil and Environmental Engineering, 4 (2).

Arizona Board of Regents (2002), Construction Cost Control and Professional Fee Guidelines. (ABOR-CCG).

Allahaim, L.; and Liu, L. (2012). Cost Overrun Causes Framework in Infrastructure Projects. Faculty of Engineering and IT, the University of Sydney and College of Architecture and Planning, King Saud University.

Arcila, S. G. (2012). Avoiding Cost Overruns in Construction Projects in the United Kingdom, Warwick Manufacturing Group.

Arditi, D.; Akan, G.; and Gurdamar, S. (1985). Cost Overruns in Public Projects. *International Journal of Project Management* 3 (4), 218-224.

Asmi, A.; Putra, J.; Abdulrahman, I. (2012). Controlling Cost Overrun Factors in Construction Projects in Malaysia Research. *Journal of Applied Sciences, Engineering and Technology* 5(8), 2621-2629.

Assaf, S. and Al-Hejji, S., (2006). Causes of Cost overrun in Large Construction Projects. *International Journal of Project Management* , 24: 349-357.

Auditor General of Western Australia, (2012). Managing Capital Projects, Perth, Australia: Office of the Auditor General of Western Australia, <http://tinyurl.com/l9ymlqu>.

Avots, I. (1983). Cost-Relevance Analysis for Overrun Control, *International Journal of Project Management*, 1 (3), 142-148.

Bacon, L. D. (1999). Using LISREL and PLS to Measure Customer Satisfaction, Sawtooth Software Conference Proceedings, La Jolla, California, 2 (5), 305-306.

Bahrain Economic Yearbook, (2014). Bahrain Economic Development Board. <http://www.bahrainedb.com/>.

Bass, B.; Avolio, B.; Jung, D.; and Berson, Y. (2003). Predicting Unit Performance by Assessing Transformational and Transactional Leadership. *Journal of Applied Psychology* 88(2), 207–218.

Cantarelli, C. (2011). Cost Overruns in Large-Scale Transport Infrastructure Projects. A Theoretical and Empirical Exploration for the Netherlands and Worldwide. Thesis Series T2010/15, the Netherlands Trail Research School.

Chantelin, Y.M.; Vinzi, V.E.; and Tenenhaus, M. (2002). State-of-Art on PLS Path Modelling Through the Available Software, HEC Research Papers Series n.764/2002, HEC School of Business and Management. Jouy-en-josas (France), HEC Research Papers Series.

Chileshe, N.; and Fianko, Y. (2012). An Evaluation of Risk Factors Impacting Construction Projects in Ghana. *Engineering Design Technology* (10) 306.

Chin, W. W. (1998). The Partial Least Squares Approach to Structural Equation Modelling. In G. A. Marcoulides (Ed.), *Modern Methods for Business Research* (295–336). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Chin, W. W.; Marcolin, B. L.; and Newsted, P. R. (2003). A Partial Least Squares Latent Variable Modelling Approach For Measuring Interaction Effects: Results From A Monte Carlo Simulation Study And Electronic Mail Emotion/Adoption Study, *Information Systems Research*, 14(2), 189-217.

CII (1989). Measuring the Cost of Quality Deviations in Design and Construction. RS 2-2, Construction Industry Institute, the University of Texas at Austin, Austin, Tex.

Cohen, J. (1992). A Power Primer. *Psychological Bulletin*, 112(1), 155-159.

Creative Research Systems, (2001). The Survey System.
www.surveysystem.com.

Creedy, G. (2006). Risk Factors Leading to Cost Overrun in the Delivery of Highway Construction Projects. Queensland University of Technology. Australia.

Davison, A. C.; and Hinkley, D. V. (1997), Lausanne and Santa Barbara. Bootstrap Methods and Their Application.

Diab, M. F.; and Varma, A. (2012). Using Risk Assessment to Improve Highway Construction Project Performance, 48th ASC Annual International Conference Proceedings by the Associated Schools of Construction. Egypt.

Eden, C.; Williams, T.; and Ackermann, F. (2005). Analyzing Project Cost Overruns: Comparing the “Measured Mile” Analysis and System Dynamics Modelling. *International Journal of Project Management* 23(2), pp. 135-139.

Efron, B.; and Tibshirani, R. (1986). Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy: Rejoinder. *Statist. Sci.*, Volume 1, Number 1.

Elinwa, U.; and Buba, S. (1993). Construction Cost Factors in Nigeria. *Journal of Construction Engineering and Management*. 119, (4) 698-714.

Faridi, A. and El-Sayegh, S., (2006). Significant Causes Causing Cost Overrun in the UAE Construction Industry. *Construction Management and Economics*, 14: 1167-76.

Flyvbjerg, B. (1997). The dark side of planning: Rationality and “Realrationalität”. Eds. Mandelbaum S., Mazza L., and Burchell & R., Explorations in planning theory. Center for Urban Policy Research Press. The State University of New Jersey, New Brunswick.

Flyvbjerg, B.; Holm, M. S.; and Buhl, S. (2002). Underestimating Costs in Public Works Projects. Journal of the American Planning Association.

Flyvbjerg, B.; Berzelius, N.; and Rothengatter, W. (2003). Megaprojects and Risk: An Anatomy of Ambition. Cambridge, United Kingdom.

Fornell, C.; and Larcker, D.F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. Journal of Marketing Research, 18 (1).

Frimpong, Y.; Oluwoye, J.; and Crawford, L.; (2002). Causes of Delay and Cost Overruns in Construction of Groundwater Projects in a Developing Countries; Ghana as a Case Study.

Gelinas, N. (2007). Lessons of Boston’s Big Dig. City Journal. http://www.cityjournal.org/html/17_4_big_dig.html

Al-Ghonamy, A., (2015). Engineering Faculty, University of Hail, Hail, Saudi Arabia. International Journal of Engineering & Technology IJET-IJENS, 15 (04), 35-42.

Gkritza, K.; and Labi, S. (2008). Estimating Cost Discrepancies in Highway Contracts, a Multistep Econometric Approach.

Gregory, R.; and Mendelsohn, R. (1992) Risk Analysis, 13(3), Perceived Risk, Dread, and Benefits.

Guadagnoli, E. and Velicer, F. (1988). Relation of Sample Size to the Stability of Component Patterns. *Psychological Bulletin*, 103 (2), 265-75.

Haenlein, M.; and Kaplan, A. M. (2004). *Users of the World, Unite! The Challenges and Opportunities of Social Media*. ESCP Europe, 79 Avenue de la Re´publique, F-75011 Paris, France.

Hair, J. F.; Hult, G. T. M.; Ringle, C. M.; and Sarstedt, M. (2013). *A Primer on Partial Least Squares Structural Equation Modelling (PLS-SEM)*. Thousand Oaks: Sage.

Han, S. (2005). *Estimation of Cost Overrun Risk in International Projects by Using Fuzzy Set Theory*. Middle East Technical University, Turkey.

Harman, H. H. (1967). *Modern Factor Analysis*. 2nd edition Chicago: University of Chicago Press.

Hartman, F. T., (2000). *Don't Park Your Brain Outside: A Practical Guide to Improving Shareholder Value with S.M.A.R.T. Management* Hardcover, Project Management Institute, 2000 - Business & Economics.

Henseler, J.; Ringle, C.; and Sinkovics, R. (2009). The Use of Partial Least Squares Path Modeling in International Marketing. *Advances in International Marketing*, 20, 277-320.

Henseler, J.; Wang, H.; Vinzi, V.; Chin, W.W. (2013). Modelling Customer Satisfaction: A Comparative Performance Evaluation of Covariance Structure Analysis Versus Partial Least Squares, in *Handbook of Partial Least Squares: Concepts, Methods and Application*. Springer, Germany, 307-325.

Hlaing N.N., Singh D., Tiong R.L.K. & Ehrlich M (2002). Perceptions of Singapore Construction Contractors on Construction Risk Identification. *Journal of Financial Management of Property and Construction*, 13 (2), 85-95.

Hulland, J. (1999). Use of Partial Least Squares (PLS) in Strategic Management Research: a Review of Four Recent Studies. *Strategic Management Journal*, 20(2), 195–204.

Hutcheson, G. and Sofroniou, N. (1999). *The Multivariate Social Scientist: Introductory Statistics Using Generalised Linear Models*. Thousand Oaks, CA: Sage Publications.

Hwang, H.; Malhotra, N. K.; Kim, Y.; Tomiuk, M. A.; and Hong, S. (2010). A Comparative Study on Parameter Recovery of Three Approaches to Structural Equation Modelling. *Journal of Marketing Research*, 47, 699-712.

Ibadov, N.; and Kulejewski, j. (2014). The Assessment of Construction Project Risks with the Use of Fuzzy Sets Theory. *Technical Transactions, Civil Engineering*, 1-B/2014.

Jackson, S. (2002). *Project Cost Overruns and Risk Management*. School of Construction Management and Engineering. The University of Reading, White knights, United Kingdom.

Kaiser, H. F. (1974). The Varimax Criterion for Analysis Rotation in Factor Analysis. *Psychometrika*, 23, 187-200.

Kaliba, C.; Muya M.; and Mumba, K. (2008). Cost Escalation and Schedule Delays in Road Construction Projects in Zambia. *International Journal of Project Management*.

Kaming, P.F.; Olomolaiye, P.O.; Holt, G.D.; and Harris, F.C. (1997). Factors Influencing Construction Time and Cost Overruns on High-Rise Projects in Indonesia. *Construction Management and Economics*, 15, 83-94.

Khandekar, S. D.; Tejale, D. S.; and Patil, J. R. (2015). Analysis of Construction Project Cost Overrun by Statistical Method. *International Journal of Advance Research in Computer Science and Management Studies*. India.

Koushki, P. A.; Al-Rashid, K.; and Kartam, N. (2004). Delays and Cost Increases in the Construction of Private Residential Projects in Kuwait. *Kuwait University*.

Le-Hoai, L.; Lee, Y. D.; and Lee, J. Y. (2008). Delay and Cost Overruns in Vietnam Large Construction Projects: A Comparison with Other Selected Countries. *Journal of Construction Management*, 367-377.

Lohmöller, J.-B. (1989). *Latent Variable Path Modelling with Partial Least Squares*, Heidelberg. Germany: Physica a-Verlag, eBook Published by Springer.

Love, P. E. D.; Sing, C. P.; Wang, W.; Irani, Z.; and Thwala, D. W. (2012). Overruns in Transportation Infrastructure Projects. *Structure and Infrastructure Engineering*, 1-19.

Mahamid, I.; and Dmaidi, N. (2013). Risks Leading to Cost Overrun in Building Construction from Consultants' Perspective Organization. *Technology and Management in Construction*. *International Journal* 5(2).

Mansfield, N.; Ugwu, O.; and Doran, T. (1994). Causes of Delay and Cost Overruns in Nigerian Construction Projects. *International Journal of Project Management*, 12 (4), 254-260.

Marcoulides, G. A.; and Saunders, C. (2006). Editor's Comments – PLS: A Silver Bullet, *MIS Quarterly*, 30(2). Management Information Systems Quarterly, California State University, and Fullerton University of Central Florida.

Al-Momani, A., (1996). Construction Cost Prediction for Public School Buildings in Jordan. *Construction Management and Economics*, 14: 311-317.

Monecke, A.; and Leisch, F. (2012). SEMPLS: Structural Equation Modeling Using Partial Least Squares. *Journal of Statistical Software*, May 2012, Volume 48, Issue 3.

Morris, M. (1990). Preparing a Landscaping Ordinance, American Planning Association Planning Advisory Service, Chicago.

Al-Najjar, J. (2008). Factors Influencing Time and Cost Overruns on Construction Projects in the Gaza Strip, The Islamic University: Gaza.

National Audit Office, (2012). The London 2012 Olympic Games and Paralympic Games: Post-Games Review, HC 794, Session 2012–13, TSO, London.

Nijkamp, P.; and Ubbels, B. (1999). How Reliable are Estimates of Infrastructure Costs? A Comparative Analysis. *International Journal of Transport Economics*, 26, 23-53.

Nunnally, S. W. (1998). North Carolina State University, Construction Methods and Management, (Seventh Edition), Prentice Hall, USA.

Odeck, J. (2004). Cost Overruns in Road Construction – What are Their Sizes and Determinants *Transport Policy*, 11(1).

Okpala, D.; and Aniekwu, A. (1988), Cause of High Cost of Construction in Nigeria. *Journal of Construction Engineering and Management*, ASCE , 114 (2) 223-34.

Petter, S.; Straub, D.; and Rai, A. (2007). Specifying Formative Constructs in Information Systems Research, *MIS Quarterly*, 31 (4), 623-656.

Sharma, S.; and Goyal, P. (2014). Access Cost Overrun Assessment Model in Fuzzy Environment. *American Journal of Engineering Research (AJER)*. Government Engineering College, Ajmer, India.

Slovic, P.; Fischhoff, B.; and Lichtenstein, S. (1980). Facts and Fears: Understanding Perceived Risk, Eds. in Schwing, R. and Albers, W. A.; *Societal Risk Assessment: How Safe Is Safe Enough?* (Plenum, New York), 181-214.

Sosik, J. J.; Kahai, S. S.; and Piovoso, M. J. (2009) Silver Bullet or Voodoo Statistics? A Primer for Using the Partial Least Squares Data Analytic Technique in Group and Organization Research. *Emerald Management Reviews: Group & Organization Management*, 34(1), 5-36.

Tenenhaus, M.; Vinzi, E.; Chatelin, Y. M.; and Lauro, C. (2005). PLS Path Modelling. *Computational Statistics & Data Analysis*, 48, 159–205.

The U.S. Federal Highway Administration Office of International Programs, (2007). Guide to the Risk Assessment and Allocation Process in Highway Construction. Report, July, 2007. Retrieved from <http://international.fhwa.dot.gov/riskassess/images/riskmatrix.cfm>.

Tipili, L. G.; Ojeba, P. O.; and Ilyasu, M. (2014). *Global Journal of Environmental Science and Technology*, 2(5), 048-054. Nigeria.

- Vinzi, E.; Chin, W.W.; Henseler, J.; and Wang, H. (2010). Review of the Book Handbook of Partial Least Squares: Concepts, Methods and Applications, International Journal of Business Science & Applied Management. 6 (2), 52-54.
- Wixom, B. H.; and Watson, H. J. (2001). An Empirical Investigation of the Factors Affecting Data Warehousing Success. MIS Quarterly 25(1), 17-38.
- Wold, H. (1982). Soft Modelling: Intermediate between Traditional Model Building and Data Analysis. Mathematical Statistics, 6, 333–346.
- Wong, K. K. (2010). Handling Small Survey Sample Size and Skewed Dataset with Partial Least Square Path Modelling. Vue: The Magazine of the Marketing Research and Intelligence Association, November, 20-23.
- Wong, P. S. P.; and Cheung, S. O. (2005). Structural Equation Model of Trust and Partnering Success. Journal of Management in Engineering, 21(2), 70-80.
- Wood, G.; and Haber, J. (1998). Nursing research: Methods, critical appraisal, and utilization (Fourth Edition). St. Louis, MO: Mosby.
- Yang, J. B.; and Ou, S. F. (2008). Using Structural Equation Modelling to Analyze Relationships Among Key Causes of Delay in Construction. Canadian Journal of Civil Engineering, 35, 321-332. <http://dx.doi.org/10.1139/L07-101>.
- Al-Zarooni, S. and Abdou, A., (2000). Risk Management in Pre-Design Stage and its Potential Benefits for UAE Public Projects. In, 28th World Congress on Housing Challenges for the 21st Century, 15-19 April, Abu Dhabi, UAE, 109-18.

Zou, P.; Zhang, G.; and Wang, J. (2007). Understanding the Key Risks in Construction Projects in China. *International Journal of Project Management* 25, 601–614.

APPENDIX ONE

QUESTIONNAIRE FORMAT

Cost Overrun in Construction Projects in Bahrain:

Organization Name: _____

Job title: _____

Part I: Please select one option for each of the following question:

a) Type of Organization

- a) Client
- b) Consultant
- c) Contractor

b) Value of Project

- a) Up to 300, 000 (Grade C)
- b) Up to BHD1, 000,000 (Grade B)
- c) Up to BHD6, 000,000 (Grade A)
- d) Over BHD 6, 000,000 (Grade AA)

c) Level of qualification

- a) None
- b) Diploma
- c) Bachelor's degree
- d) Master's degree
- e) PhD

d) Work Experience

- a) 0-5 years
- b) 6-10 years
- c) 11-15 years
- d) 16-20 years
- e) More than 20 years

Frequency of cost overrun experienced

Less than 5% of projects

5% to 15%

15% to 35%

35% to 50%

50% to 75%

More than 75%

Project delivery method

Design-Bid-Build (DBB)

Design-Build (DB)

Construction Management (CMAR)

Integrated Project Delivery (IPD)

Type of construction projects

Private

Government

Private and Government

Type of construction contract:

Unit price contract

Lump sum contract

Cost plus contract

Incentive contracts

Percentage of construction contracts

Type of estimating methods:

Parametric

Factored

Basic price

Unit rate

Approach of contingency analysis :

Probability based

Historical values

Decision analysis

Experience based

Other

Projects location:

Bahrain only

Bahrain & GCC countries only

Worldwide

Part II:

On scale 1-5 rate the cost overrun factors severity on project cost and frequency of occurrence in projects?

1 = not significant; 2 = slightly significant; 3 = moderately significant; 4 = very significant; 5 = extremely significant

Factors		Frequency of occurrence	Degree of Severity on project cost
a) Project Management and Contract Administration Related Factors:			
1	Poor project management		
2	Change in the scope of the project		
3	Delays in decisions making		
4	Inaccurate quantity take-off		
b) Design and Documentation Related Factors:			
5	Frequent design changes		
6	Mistakes and errors in design		
7	Incomplete design at the time of tender		
8	Poor design and delays in design		
9	Delay preparation and approval of drawings		
c) Contractor's Site Management Related Factors:			
10	Poor site management and supervision		
11	Incompetent subcontractors		
12	Schedule delay		
13	Inadequate planning and scheduling		
14	Lack of experience		
15	Inaccurate time and cost estimates		
16	Mistakes during construction		

17	Inadequate monitoring and control		
d) Financial Management Related Factors:			
18	Cash flow		
19	Poor financial control on site		
20	Financial difficulties of owner		
21	Delay in progress payment by owner		
22	Delay payment to supplier /subcontractor		
23	Contractual claims		
e) Information and Communication Related Factors:			
24	Lack of coordination between parties		
25	Slow information flow between parties		
26	Lack of communication between parties		
f) Human Resources (Workforce) Related Factors:			
27	Labor productivity		
28	Shortage of site workers		
29	Shortage of technical personnel		
30	High cost of labor		
31	Labor absenteeism		
g) Non-Human Resource Related Factors:			
32	Fluctuation of prices of materials		
33	Shortages of materials		
34	Equipment availability and failure		
35	Late delivery of materials and equipment		
h) Environmental Related Factors:			
36	Level and number of competitors		

37	economic instability		
38	effects of weather		
39	government policies		
40	inadequate production of raw materials by the country		
41	monopoly by suppliers		
42	number of projects going at the same time		
43	political situation		
44	project location		
45	social and cultural impacts		

Please list any other factors or information that you feel we could add to this research:

.....

.....

.....

.....

.....

.....

.....

.....

...Thank You ...

APPENDIX TWO
ACTUAL PROJECT DATA
FOR THE PERIOD 2008-09 to 2012-13

Forty construction projects from contracting companies in Bahrain:

Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/JHR/08/01	Open Tender	1,250,000	1,667,000	33.36%
Cost overrun reasons	Inaccurate quantity take-off Shortage of site workers			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/JHR/08/02	Agreed Price	412,500	467,200	13.26%
Cost overrun reasons	Lack of communication between parties Frequent design changes			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/CHP/08/03	Agreed Price	272,000	313,500	15.26%
Cost overrun reasons	Delays in decisions making			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/CHP/08/04	Open Tender	2,785,000	3,072,000	10.31%
Cost overrun	Delay payment to supplier /subcontractor			

reasons	Delay in progress payment by owner			
	Cash flow			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/CPC/08/05	Open Tender	1,748,000	1,822,750	4.28%
Cost overrun reasons	Mistakes during construction Inaccurate quantity take-off			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/CPC/08/06	Agreed Price	678,250	794,670	17.16%
Cost overrun reasons	Lack of coordination between parties Late delivery of materials and equipment			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/CPC/08/07	Open Tender	3,500,000	3,750,000	7.14%
Cost overrun reasons	Inaccurate time and cost estimates			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/KHJ/08/08	Agreed	910,000	1,125,000	23.63%

	Price			
Cost overrun reasons	Mistakes and errors in design Frequent design changes			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/KHJ/08/09	Agreed Price	800,000	876,500	9.56%
Cost overrun reasons	Mistakes and errors in design			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/NAS/09/10	Open Tender	740,000	810,400	9.51%
Cost overrun reasons	Schedule delay Delay preparation and approval of drawings Frequent design changes			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/NAS/09/11	Open Tender	3,760,000	3,978,000	5.80%
Cost overrun reasons	Poor project management Delays in decisions making			

Project number	Description of works	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/NAS/09/12	Open Tender	1,542,000	1,695,000	9.92%
Cost overrun reasons	Frequent design changes Mistakes during construction			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/HFR/09/13	Open Tender	678,000	721,000	6.34%
Cost overrun reasons	Mistakes and errors in design Incomplete design at the time of tender			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/HFR/09/14	Open Tendre	2,242,000	2,760,400	23.12%
Cost overrun reasons	Schedule delay			

Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/HFR/09/15	Agreed Price	212,000	287,000	35.38%

Cost overrun reasons	Poor design and delays in design Delay preparation and approval of drawings Mistakes during construction			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/DWT/10/16	Open Tendre	1,345,000	1,770,000	31.60%
Cost overrun reasons	Incomplete design at the time of tender Inaccurate time and cost estimates			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/DWT/10/17	Agreed Price	684,500	705,000	2.99%
Cost overrun reasons	Frequent design changes Delay preparation and approval of drawings Schedule delay Inadequate planning and scheduling			

Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/DWT/10/18	Agreed Price	818,600	876,800	7.11%
Cost overrun	Lack of experience			

reasons	Inadequate monitoring and control Poor financial control on site Lack of coordination between parties Slow information flow between parties Shortage of site workers			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/CHP/10/19	Open Tendre	4,280,000	4,697,000	9.74%
Cost overrun reasons	Labor productivity Shortage of site workers Labor absenteeism			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/CHP/10/20	Open Tendre	954,000	997,800	4.59%
Cost overrun reasons	Schedule delay Slow information flow between parties Shortage of technical personnel			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/WRD/10/21	Agreed Price	540,000	615,300	13.94%
Cost overrun	Frequent design changes			

reasons	Delay preparation and approval of drawings			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/4DT/10/22	Agreed Price	95,000	112,600	18.53%
Cost overrun reasons	Change in the scope of the project Poor site management and supervision			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/WRD/10/23	Open Tendre	768,000	806,000	4.95%
Cost overrun reasons	Inadequate planning and scheduling Lack of experience Inadequate monitoring and control			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/4DT/10/24	Agreed Price	112,000	198,300	77.05%
Cost overrun reasons	Lack of experience Mistakes during construction			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/AHJ/10/25	Open	1,780,000	1,922,000	7.98%

	Tendre			
Cost overrun reasons	Inaccurate quantity take-off Schedule delay Inaccurate time and cost estimates			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/AHJ/10/26	Agreed Price	430,000	471,700	9.70%
Cost overrun reasons	Schedule delay Inaccurate time and cost estimates			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/AHJ/10/27	Open Tendre	948,000	1,145,000	20.78%
Cost overrun reasons	Frequent design changes			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/AHD/10/28	Open Tendre	3,377,400	3,491,800	3.39%
Cost overrun reasons	Incomplete design at the time of tender Cash flow Shortages of materials Equipment availability and failure			

	Late delivery of materials and equipment monopoly by suppliers			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/AHD/11/29	Agreed Price	256,000	270,000	5.47%
Cost overrun reasons	effects of weather			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/AMH/11/30	Open Tendre	5,452,800	5,568,200	2.12%
Cost overrun reasons	Frequent design changes			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/AMH/11/31	Agreed Price	914,500	968,000	5.85%
Cost overrun reasons	Shortages of materials Late delivery of materials and equipment			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/AMH/11/32	Open	485,000	513,200	5.81%

	Tendre			
Cost overrun reasons	Delays in decisions making Inaccurate quantity take-off Schedule delay			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/AMH/11/33	Open Tendre	668,500	697,000	4.26%
Cost overrun reasons	Shortage of site workers Shortages of materials number of projects going at the same time			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/HFR/11/34	Agreed Price	178,000	192,000	7.87%
Cost overrun reasons	Cash flow Delay in progress payment by owner Labor productivity High cost of labor			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/HFR/11/35	Open Tendre	1,245,000	1,298,700	4.31%
Cost overrun	Contractual claims			

reasons	Level and number of competitors			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/DIV/11/36	Open Tendre	2,785,600	2,945,250	5.73%
Cost overrun reasons	number of projects going at the same time Incomplete design at the time of tender Shortage of site workers Shortages of materials			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/DCC/12/37	Open Tendre	111,700	126,200	12.98%
Cost overrun reasons	Poor project management Delays in decisions making Inaccurate quantity take-off Frequent design changes			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/DCC/12/38	Agreed Price	278,400	296,000	6.32%
Cost overrun reasons	Poor design and delays in design Inadequate planning and scheduling			

	Inaccurate time and cost estimates			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/MCC/12/39	Open Tendre	2,012,000	2,178,300	8.27%
Cost overrun reasons	Inaccurate time and cost estimates Cash flow Delay in progress payment by owner Delay payment to supplier /subcontractor High cost of labor			
Project number	Method of delivery	Estimated cost BHD	Actual cost BHD	% of cost overrun
CP/MCC/12/40	Agreed Price	378,000	391,000	3.44%
Cost overrun reasons	Inadequate planning and scheduling Lack of experience			

*** Reference number:**

Example: CP/CHP/08/04

CP: Construction project

CHP: Company name (ID)

08: Year of construction

04: Serial number

APPENDIX THREE
RELIABILITY TESTS OF SURVEY
AND ACTUAL DATA

Cronbach's Alpha reliability was computed using SPSS software to evaluate the reliability of the survey, actual and combined data. The combined sample size of 74 respondents and 40 actual construction projects was selected to conduct the study, since it had the most reliable results.

1. Combined Data Reliability Test

Cronbach's Alpha reliability method was applied to check the reliability of all factors in the combined, survey and actual data. The reliability coefficient values are given in Tables 1.1, 1.2 and 1.3 below:

Table 1.1: combined data reliability test

Cronbach's Alpha	No. of Items
0.826	45

Table 1.2: survey data reliability test

Cronbach's Alpha	N of Items
.654	45

Table 1.3: actual data reliability test

Cronbach's Alpha	N of Items
.757	45

2. Factor Analysis

Table 2.1 and 2.2 are showing the KMO and Bartlett's Test Results, for combined, survey and actual data.

Table 2.1: combined data KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	0.734
Approx. Chi-Square	4417.802
Bartlett's Test of Sphericity	df
	990
	Sig.
	0.000

Table 2.2: survey data KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.438
Approx. Chi-Square	1589.649
Bartlett's Test of Sphericity	df
	990
	Sig.
	.000

Actual data matrix was not positive definite

APPENDIX FOUR
PLS-SEM
THEORETICAL BACKGROUND

PLS path models consist of two components: the structural model and the measurement model. PLS path models only permit recursive relationships and can be expressed as simple connected digraphs. A digraph is called simple if it has no loops and at most one arc between any pair of nodes. A digraph is connected if an undirected path between any two nodes exists; consequently no node is isolated from the rest.

1. The Structural Model

In the structural model, also called inner model, the latent variables (LVs) are related with each other according to substantive theory. LVs are divided into two classes, exogenous and endogenous. Exogenous LVs do not have any predecessor in the structural model, all others are endogenous.

For the benefit simplicity the notation we use for the structural model dismisses the difference between exogenous and endogenous variables and we start with the compact form

$$Y = Y B + Z \quad (1)$$

Where, Y : denotes the matrix for the latent variables, both exogenous and endogenous.

Z : the error terms, Z are assumed to be centred, i.e., $E [Z] = 0$.

B : elements of the coefficients matrix, B are restricted to zero where the elements of the adjacency matrix D are zero.

2. The Measurement Model

The measurement model or outer model relates observed variables to their latent variables (LVs). Often observed variables are referred to as manifest variables or indicators, latent variables as factors. Within the PLS framework one manifest variable can only be related to one LV. All manifest variables related to one LV are called a block. So each LV has its own block of observed variables. A block must contain at least one MV. The way a block can be related to an LV can be either reflective (as in the cost overrun model) or formative.

Without loss of generality following assumptions can be made:

1. All MVs contained in the data matrix X are scaled to have zero mean and unit variance.
2. Each block of MVs X_g is already transformed to be positively correlated for all LVs y_g , $g = 1, \dots, G$.

Reflective measurement: In the reflective way each block of MVs reflects its LV and can be written as the multivariate regression:

$$X_g = y_g w_g^T + F_g, \quad E[F_g | y_g] = 0 \quad (2)$$

So w_g^T can be estimated by least squares as

$$\begin{aligned} W_g^T &= (y_g^T y_g)^{-1} y_g^T X_g \\ &= \text{VAR}(y_g)^{-1} \text{COV}(y_g, X_g) \\ &= \text{COV}(y_g, X_g) \end{aligned}$$

$$= \text{COR}(y_g, X_g). \quad (3)$$

PLS algorithm estimates all the LVs y_g , $g = 1, \dots, G$, as linear combination of their MVs under the constraint to have unit variance. All MVs was scaled to zero mean and unit variance. Consequently, the equality above is valid. The latent variable y_g is measured by the block X_g consisting of observed variables, $x_1, x_2, x_3, \dots$, in a reflective way (Figure 1).

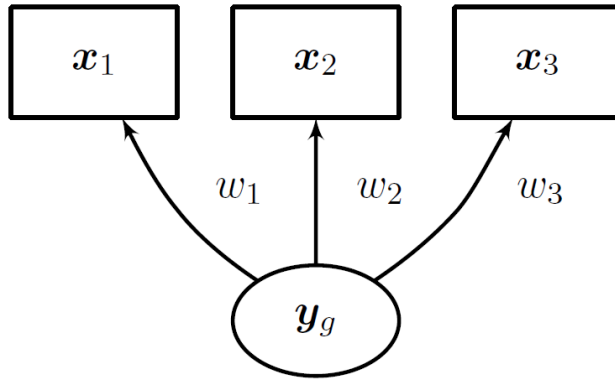


Figure 1: Reflective model path diagram.

3. The Partial Least Squares (PLS) Algorithm

The PLS algorithm aims at estimating the values for LVs (factor scores) by an iterative procedure. Figure 2 depicts the flowchart of the algorithm. The idea is to first construct each LV by the sum of its MVs. Then in the inner approximation, each LV is reconstructed by the means of its neighboring LVs. In the outer approximation the best linear combination must be find to express each LV by means of its MVs; the coefficients are referred to as outer weights. Finally, in step 4, each LV is constructed as weighted sum or linear combination of its MVs. After each step the LVs are scaled to have

zero mean and unit variance. The algorithm stops if the relative change for all the outer weights is smaller than a predefined tolerance.

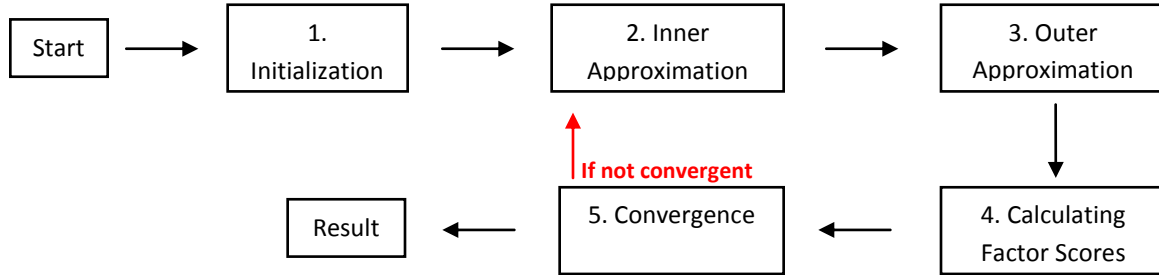


Figure 2: The diagram depicts the flowchart for the PLS algorithm.

3.1 Step 1 Initialization

Each LV is constructed as a weighted sum of their MVs. It was assumed that all the MVs, $X_1, \dots, X_K$, to be scaled ($\text{mean}(X_i) = 0$ and $\text{VAR}(X_i) = 1$). In the initialization all the weights equal one. As a sum of centered variables all the LVs are also centered ($\text{mean} = 0$). But they still have to be scaled to have unit variance ($\text{var} = 1$).

$$Y = X M \quad (4)$$

$$\hat{y}_g = \frac{y_g}{\sqrt{\text{VAR}(y_g)}} \quad , g = 1, \dots, G \quad (5)$$

The LVs are initialized as: $\hat{Y} = (\hat{y}_1, \dots, \hat{y}_G)$.

3.2 Step 2 Inner Approximations

In the inner approximation each LV is estimated as a weighted sum of its neighboring LVs. Again the recomputed LVS is scaled to have unit variance.

$$Y = \tilde{Y} \hat{E}$$

$$\hat{y}_g = \frac{\tilde{y}_g}{\sqrt{\text{VAR}(\tilde{y}_g)}} \quad , g = 1, \dots, G \quad (6)$$

The LVs are initialized as: $\tilde{Y} = (\tilde{y}_1, \dots, \tilde{y}_G)$.

3.3 Step 3 Outer Approximation)

For the initialization all weights were one, the weights are recalculated on the basis of the LV values from the inner approximation (Step 2). The weights can be estimated as a multivariate regression coefficient with the block of MVs as response and the LV as regressor:

$$\begin{aligned} \hat{w}_g^T &= (\tilde{y}_g^T \tilde{y}_g)^{-1} \tilde{y}_g^T X_g \\ &= \text{COR}(\tilde{y}_g, X_g), \end{aligned} \quad (7)$$

3.4 Step 4: Calculating Factor Scores

The outer weights vectors, $w_1, \dots, w_G$, in an outer weights matrix $\mathbf{W}$, will be used to estimate the factor scores by means of the MVs:

$$\hat{Y} = XW \quad (8)$$

$$\hat{Y}_g = \frac{\tilde{Y}_g}{\sqrt{\text{VAR}(\hat{Y}_g)}} \quad , g = 1, \dots, G \quad (9)$$

Resulting in the outer estimation: $\hat{Y} = (\hat{y}_1, \dots, \hat{y}_G)$.

3.5 Step 5 Convergence

If the relative change of all the outer weights from one iteration to the next are smaller than a predefined tolerance,

$$\left| \frac{\hat{W}_{oldkg} - \hat{W}_{newkg}}{\hat{W}_{newkg}} \right| < \text{tolerance} \quad \forall k = 1, \dots, K \wedge g = 1, \dots, G, \quad (10)$$

The estimation of factor scores done in Step 4 is taken to be final. Otherwise Step 2 must be recalculated.

4. Weighting schemes

The weighting scheme is used for the estimation of the inner weights in Step 2 of the PLS algorithm. Originally Wold (1982) proposed the centroid weighting scheme. Later Lohmöller (1989) introduced two other schemes, factorial and path weighting. The adjacency matrix D for the LVs in the model is representing the structural part of the model. Contrary to the matrix M for the measurement model, D accounts for the directionality. For every $d_{ij} = 1$, there is an arc from node i , the head of the arc, to node j , the tail of the arc. It could also be said that, the columns indicate the successor, whereas the rows indicate the predecessors. The adjacency matrix D facilitates the calculation of the inner weights. For all the weighting schemes, each LV is constructed as a weighted sum of the LVs it is related with. The weighting schemes differ in the way the relation is defined. Generally the inner estimate can be expressed $\tilde{Y}$ as matrix product of the outer estimate $\hat{Y}$ and the matrix of inner weights E :

$$\tilde{Y} = \hat{Y} E \quad (11)$$

Denote $R = \text{COR}(\hat{Y})$, the empirical correlation matrix for the LVs resulting from the outer estimation, and $C = D + D^T$ a symmetrical matrix indicating whether two LVs are neighbours.

Centroid weighting scheme: Following the centroid weighting scheme, the matrix of inner weights E takes the form:

$$e_{ij} = \begin{cases} \text{sign}(r_{ij}) & , \text{ for } c_{ij} = 1, \\ 0 & , \text{ else} \end{cases} , i, j = 1, \dots, G. \quad (12)$$

Factorial weighting scheme: The factorial weighting scheme,

$$e_{ij} = \begin{cases} T_{ij} & , \text{ for } c_{ij} = 1, \\ 0 & , \text{ else} \end{cases} , i, j = 1, \dots, G. \quad (13)$$

Path weighting scheme: For the path weighting scheme (or structural scheme) the predecessors and successor of a LV play a different role in the relation. The out-neighbourhood, or successor set of a node i is defined as the set of tails of arcs going from i . Likewise, an in-neighbourhood, or predecessor set of a node i is the set of heads of arcs going into i . A head is representing the start/initial node of an arch, a tail its end/terminal node. The relation for one specific LV y_i with its successor is determined by their correlation, for the predecessors it is determined by a multiple regression.

$$y_i = y_i^{\text{pred}} \gamma + z_i, \quad E[z_i] = 0, \quad i = 1, \dots, G$$

With y_i^{pred} the predecessor set of the LV y_i . Denoting y_i^{succ} the successor set of the LV y_i the elements of the inner weight matrix E are

$$e_{ij} = \begin{cases} \gamma_j & , \text{ for } j \in y_i^{\text{pred}}, \\ \text{COR}(y_i, y_j) & , \text{ for } j \in y_i^{\text{succ}}, \\ 0 & , \text{ else} \end{cases} \quad (14)$$

5. Calculation of Path Coefficients, Total Effects and Loadings

Once the factor scores are estimated by PLS algorithm, the path coefficients can be estimated by ordinary least squares (OLS), according to the structural model. For each LV $\hat{y}_g$, $g = 1, \dots, G$, the path coefficient is the regression coefficient on its predecessor set $\hat{y}_i^{\text{pred}}$

$$\hat{\beta}_g = (\hat{y}_g^{\text{predT}} \hat{y}_g^{\text{pred}})^{-1} \hat{y}_g^{\text{predT}} y_g$$

$$\text{COR}(\hat{y}_g^{\text{pred}}, \hat{y}_g^{\text{pred}})^{-1} \text{COR}(\hat{y}_g^{\text{pred}}, \hat{y}_g) \quad (15)$$

The elements $\hat{b}_{ij}$, $i, j = 1, \dots, G$, is obtained of the estimated matrix of path coefficients $\hat{B}$:

$$\hat{\beta}_{ij} = \begin{cases} \hat{\beta}_{gj} & , \text{ for } j = y_i^{\text{pred}}, \\ 0 & , \text{ else} \end{cases} \quad (16)$$

The matrix $\hat{B}$ can be interpreted as transition matrix for the structural model. The matrix of total effects $\hat{T}$ can be calculated as the sum of the 1 to G step transition matrices:

$$\hat{T} = \sum_{g=1}^G \hat{B}_g. \quad (17)$$

$\hat{B}_g$ expands to $\hat{B} \cdot \hat{B} \cdot \dots \cdot \hat{B}$, e.g., $\hat{B}^2$ contains all the indirect effects mediated by only one LV. The cross and outer loadings are estimated as:

$$\hat{\Lambda}_{\text{cross}} = \text{COR}(X, \hat{Y}) \quad (18)$$

$$\hat{\lambda}_{\text{outer } kg} = \begin{cases} \hat{\lambda}_{\text{cross } kg} & , \text{ if } m_{kg} = 1, \\ 0 & , \text{ else.} \end{cases} \quad (19)$$

Where, M is the adjacency matrix for the measurement model.

Notations:

Symbol	Description
N	Number of observations
K	Number of observed variables
G	Number of latent variables
X	$N \times K$ matrix of observed variables (MVs); each variable standardized
X_g, g = 1, ..., G	Block of observed variables (MVs) belonging to latent y _g
F_g	$N \times K$ matrix measurement error for reflective block X _g
δ_g	Measurement error vector of length N for a formative LV y _g
Y	$N \times G$ matrix of for the latent variables (LVs)
Z	$N \times G$ matrix of the structural model error terms
$\tilde{Y}$	$N \times G$ matrix: inner approximation/estimation of factor scores
$\hat{Y}$	$N \times G$ matrix: outer approximation/estimation of factor scores
M	Adjacency matrix ($K \times G$) for the measurement model
D	Adjacency matrix ($G \times G$) for the structural model
W	$K \times G$ matrix of outer weights
E	$G \times G$ matrix of inner weights
B	$G \times G$ matrix of path coefficients
T	$G \times G$ matrix of the total effects
Λ^{cross}	$K \times G$ matrix of cross loadings
Λ^{outer}	$K \times G$ matrix of outer loadings

ملخص الدراسة

إن الغاية الأساسية لقطاع البناء والتشييد هو إيجاد منشأة قادرة على تلبية إحتياجاته الوظيفية بالكفاءة المطلوبة، في غضون إطار زمني محدد وميزانية مقدرة. وعلى الرغم من أهمية هذه الميزانية ، إلا أن مشاريع البناء لا تزال تتجاوزها بشكل روتيني.

يعزى تجاوز التكاليف في مشاريع البناء إلى العديد من الأسباب كحدوث أخطاء في التصميم وعدم الكفاءة الإدارية وانعدام التنسيق بين الأطراف وغياب الرقابة المالية وحتى الفساد. عادة ما يكون لهذه الأسباب تأثير كبير على التكلفة النهائية للمشروع، ولكن من الصعب تحديد الأثر المالي لكل منها على الميزانية. وبالتالي يكون الاعتماد بشكل كبير على الخبرة والحدس عند إعداد تقديرات أولية، تؤدي في كثير من الأحيان لإهمال معظم هذه الأسباب والعوامل عند احتساب التكلفة النهائية. لذا فإن قرارات البناء لمعظم المشاريع الانشائية مبنية إلى حد كبير على تقديرات غير واقعية تؤدي بلا شك إلى تجاوز ميزانيتها المقدرة فيما بعد.

إن الهدف الرئيسي من هذا البحث هو دراسة العوامل المؤثرة على تكاليف مشاريع البناء في البحرين، ووضع نموذج يمكن أن يساعد في التوصل إلى تقديرات للتكلفة النهائية بشكل أكثر موثوقية وواقعية في المراحل الأولى للمشروع.

تتحدد فرضية الدراسة في أنه من المتوقع وجود فروق في أثار العوامل المختلفة على تكاليف المشاريع الانشائية في البحرين. كما اعتمدت هذه الدراسة على استراتيجيتين احدهما تتمثل في استفتاء يشمل العاملين في قطاع البناء من مقاولين و استشاريين و أصحاب مشاريع، و الاخرى عبارة دراسة تحليلية لمشاريع بناء سابقة مع مقابلات للعاملين عليها.

إستخدمت البيانات التي تم جمعها لدراسة العلاقة الأساسية بين المصادر المؤثرة على تكلفة المشاريع الانشائية باستخدام أسلوب تحليل متعدد المتغيرات لتقدير وتحليل العلاقات السببية في نماذج المسار. وقد تم تطوير نموذج لتقييم العوامل المسببة لتجاوز التكاليف. كما تم إيجاد أبرز العوامل المؤثرة على تكلفة المشروع النهائية و المتمثلة في التغييرات و التعديلات المستمرة في التصميم و الأخطاء في مرحلة البناء بالإضافة الى التأخير في الجدول الزمني للمشروع. ومن المؤمل أن هذه النتائج سوف توجه الجهود الرامية إلى تحسين أداء صناعة البناء والتشييد في المستقبل.

جامعة البحرين



قسم الهندسة الميكانيكية

كلية الهندسة

أسباب و آثار تجاوز التكاليف على المشاريع الإنشائية في البحرين

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

إعداد

هداية عبدالسميع محمد أبو صفية

20071812

إشراف

د. سعد سليمان

أستاذ

جامعة البحرين

مملكة البحرين

سبتمبر, 2016