A new methodology for tracking the performance of subcontractors in the construction industry

Author:
Almohssen, Abdulaziz

Publication Date:
2018

DOI:
https://doi.org/10.26190/unsworks/3444

A New Methodology for Tracking the Performance of Subcontractors in the Construction Industry

by

Abdulaziz Saud Almohssen

BEng, King Faisal University

MEng, UNSW Australia

A thesis presented in fulfilment of the requirements for the degree of Doctor of Philosophy

School of Civil and Environmental Engineering

Faculty of Engineering

The University of New South Wales

May 2018

PLEASE TYPE

THE UNIVERSITY OF NEW SOUTH WALES
Thesis/Dissertation Sheet

Surname or Family name: Almohssen

First name: Abdulaziz

Other name/s: Saud

Abbreviation for degree as given in the University calendar: PhD

School: Civil and Environmental Engineering

Faculty: Engineering

Title: A New Methodology for Tracking the Performance of Subcontractors in the Construction Industry

Abstract 350 words maximum: (PLEASE TYPE)

Forecasting the performance of subcontractors is fraught with difficulty. Any particular measurement will have both information about the subcontractors performance and also noise due to random effects. This thesis aims to present a methodology for separating the underlying performance from the noise in the measurements. A secondary aim is to examine how this affects the optimum time for updating historical records of subcontractors.

A case study has been adopted to test this methodology using data collected from Saudi Arabia. Data was collected in two phases. The first phase involved interviewing experts in assessing subcontractors to explore the importance of tracking the performance of subcontractors, the most important performance factors, and the frequency of updating historical records. The second phase involved collecting data about subcontractors' performance from historical records.

The results of the interviews show that different organisations focus on different factors, but they have strong agreement that work quality and safety are important. They also have different frequencies for updating their historical records, ranging from 1 to 5 years. The performance questions were classified also in two groups using factor analysis: management questions and technical questions. The expected change over time in subcontractors' performance was studied by using Markov chains. The noise content of these measurements was studied by comparing with hidden Markov models using the Baum Welch algorithm. A methodology was also provided that enables tracking of the loss of accuracy over time based on entropy. The results of the case study show that subcontractors improve over time in technical performance faster than in management performance. The updating time of the historical records based on the case study is recommended to be annually for technical questions and every two years for management questions.

This research demonstrates that hidden Markov models provide a new strategy for forecasting subcontractors' performance and reducing the effect of randomness to increase accuracy. A limitation of this work is that it is based on a single case study.

FOR OFFICE USE ONLY

Date of completion of requirements for Award:

ORIGINALITY STATEMENT

'I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.'

Signed ....

Date ....

COPYRIGHT STATEMENT

'I hereby grant the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all proprietary rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.
I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstract International (this is applicable to doctoral theses only).
I have either used no substantial portions of copyright material in my thesis or I have obtained permission to use copyright material; where permission has not been granted I have applied/will apply for a partial restriction of the digital copy of my thesis or dissertation.'

Signed ....

Date ....

AUTHENTICITY STATEMENT

'I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.'

Signed ....

Date ....

Dedication

This PhD thesis and my entire achievements are dedicated to my mother, Farida, and my father, Saud, who kept supporting me in my whole educational journey. The success that I have achieved would not be possible without their encouragement, care and prayers.

I would also dedicate this thesis to my wife, Ohoud, who joined me in this journey and shared all good and hard times I had gone through; with patience, sacrifices, support and efforts, and to my lovely sons, Saud and Nasser, whom I wish all blissfulness and wellbeing.

A big thank you to all of you is not enough to express my feelings towards everything you tried to do.

Acknowledgements

My first thanks would surely go to my supervisor, Steven Davis, whom I grateful to have been able to complete my doctoral studies under his guidance and immense support and also his continuous encouragement during the whole period of my PhD studies. He always inspired me while working on this thesis where completing it would not be possible without him. His openness, brain storming sessions and the creative ideas that he provided to overcome all difficulties and to fulfil the requirements of this thesis were really helpful and appreciated. I am considered lucky to have had him as my supervisors where he was generous and keen to provide this thesis which is expected to contribute to knowledge. Big thank you Davis for all efforts and time you have had done to me in the past 4 years. Thanks for everything I got to learn from you.

I would additionally like to thank David Carmichael, Ali Akbar Nezhad, Johnson Xuesong Shen, Travis Waller and Taha Rashidi for their useful insight during my reviews.

Above all, I am deeply thankful and grateful to the government of Saudi Arabia , represented by Saudi Cultural Mission in Australia, and the University of Imam Abdulrhman bin Faisal for their financial and moral support, without which I would not have acquired a privileged education.

Peace be upon you

Abdulziz Almohssem

Abstract

Forecasting the performance of subcontractors is fraught with difficulty. Any particular measurement will have both information about the subcontractors performance and also noise due to random effects. This thesis aims to present a methodology for separating the underlying performance from the noise in the measurements. A secondary aim is to examine how this affects the optimum time for updating historical records of subcontractors.

A case study has been adopted to test this methodology using data collected from Saudi Arabia. Data was collected in two phases. The first phase involved interviewing experts in assessing subcontractors to explore the importance of tracking the performance of subcontractors, the most important performance factors, and the frequency of updating historical records. The second phase involved collecting data about subcontractors' performance from historical records.

The results of the interviews show that different organisations focus on different factors, but they have strong agreement that work quality and safety are important. They also have different frequencies for updating their historical records, ranging from 1 to 5 years. The performance questions were classified also in two groups using factor analysis: management questions and technical questions. The expected change over time in subcontractors' performance was studied by using Markov chains. The noise content of these measurements was studied by comparing with hidden Markov models using the Baum Welch algorithm. A methodology was also provided that enables tracking of the loss of accuracy over time based on entropy. The results of the case study show that subcontractors improve over time in technical performance faster than in management performance. The updating time of the historical records based on the case study is recommended to be annually for technical questions and every two years for management questions.

This research demonstrates that hidden Markov models provide a new strategy for forecasting subcontractors' performance and reducing the effect of randomness to be annually for technical questions and every two years for management questions.

This research demonstrates that hidden Markov models provide a new strategy for forecasting subcontractors' performance and reducing the effect of randomness to increase accuracy. A limitation of this work is that it is based on a single case study

Table of Contents

Lists of Figures

Lists of Tables

Chapter 1 Introduction

1.1 Background

Subcontracting is a widespread practice in the construction field in most countries. For example, in Turkey there are twice as many subcontractors as contractors (Ulubeyli et al., 2010). The success of a construction project is significantly determined by the quality and reputation of the subcontractors selected (Dulung and Pheng, 2005). Maturana et al. (2007) reports that payments to subcontractors typically represent 60% to 70% of total project value. Also, some technical works and specific activities need to be accomplished by specialised subcontractors. According to Greenwood (2001), the improper selection of subcontractors leads to time and cost over-runs and contractor disputes, which are common issues in many construction projects. The importance of subcontractors increases as projects become larger and more complex (Wang and Yuan, 2011).

Contractors use different techniques and strategies to select subcontractors for construction projects. These techniques and strategies vary according to the factors that each contractor considers more significant. Price, quality and experience in similar projects are examples of aspects that contractors consider. For some of these aspects it is necessary to have access to past performance data.

According to Alfeld (1988), identifying past performance provides a baseline to benchmark and measure future performance against. Any variation in the performance level compared to this baseline can be measured. The variation could be positive or negative. Further management techniques should be used to determine the root causes of this variation to help with determining how to either enhance or eliminate the variation (Cox et al., 2003). The importance of performance in construction projects leads to the fact that subcontracting a large amount of work introduces risk for the main contractor, (Cooke and Williams, 2013). Therefore, the main contractor should efficiently manage the performance of subcontractors to assure that the expected performance is achieved.

Performance measurement has been increasingly used in many countries using different techniques and strategies, such as UK, Bassioni et al. (2004), Singapore, Hartmann and Caerteling (2010), Korea, Eom et al. (2008). Performance measures are usually assigned based on project objectives so that these performance measures are tied to success. Cox et al. (2003) points out that performance is a more comprehensive measure that includes productivity as one of the factors to be considered. Oluwoye et al. (1996) differentiates between effectiveness and efficiency. The difference is that effectiveness is related to the fulfilment of the explicit and implicit goals, whereas efficiency is related to minimizing the resources used to achieve those goals.

One of the main areas that have been studied in regards to performance in the construction industry is identifying the key performance indicators (KPIs) used in different countries. Similar research has been carried out into critical success factors (CSFs). This research has shown a great deal of consistency between the CSFs and KPIs tracked across many different countries.

1.2 Problem Statement

However, the change of performance of subcontractors over time has been overlooked. The change of performance of a subcontractor has a crucial importance since subcontractors can increase their capability both in the range of tasks that it can perform, and in how well it performs its tasks. Similarly, if subcontractors do not practice particular skills then their performance in these areas can drop. This leads to the questions of how does subcontractor performance change over time and how can this be measured. Furthermore it leads to the question of how long performance records can be considered current.

1.3 Aim of the research

The aim of this research is to fill the gap of studying the change in performance of subcontractors by developing a methodology to analyse the performance of subcontractors over time. This includes identifying long-term factors that could be used in tracking the performance of subcontractors in construction projects. The thesis also shows how to test for the consistency of performance ratings given by different contractors and how to examine the internal structure of the performance questions used to help uncover detailed insights later. A methodology will also be developed to identify the optimum frequency for updating the historical performance records of subcontractors based on the speed at which subcontractor performance varies and old information becomes obsolete.

This overall aim has been divided into a set of objectives that are shown in Table 1.1 with their corresponding methods. The research methodology of achieving these objectives is shown in Figure 1.1

Table 1.1. The objectives of the research with their corresponding hypothesis

Chapter	Objective	Method
4	Studying the relationship between the importance of the performance factors and the performance ratings given by contractors.	Identifying the most important performance factors used by different contractors and then collecting performance data to find out if contractors are more critical when assessing these important factors by using correlation and regression analysis.
5	Testing the consistency of different performance ratings provided by different contractors and identifying any internal structure of the performance questions.	Consistency of the different performance ratings will be tested using Cronbach's alpha whereas the internal structure of the performance factors will be identified using factor analysis.
6	Validating the Markovity of the performance data and testing whether different data sets can be combined.	Markovity will be tested using Kullback's logarithms. Rating sequences given by individual contractors will be tested separately to rating sequences deriving from multiple contractors.
7	Analysing the change in the performance of subcontractors over time and also the optimum updating frequency of historical records.	Markov chains will be used to model the performance of subcontractors. The frequency of the optimum updating time will be analysed using the drop in probability and the increase in variance and Shannon information entropy of the outcome.
8	Improving the signal to noise ratio in performance measurements and associated updating frequency.	Hidden Markov models, calculated using the Baum Welch algorithm, will be used to get a more accurate indication of the underlying performance separate from the random effects involved in individual measurements. Updating frequency will be similar to above, but using the new model

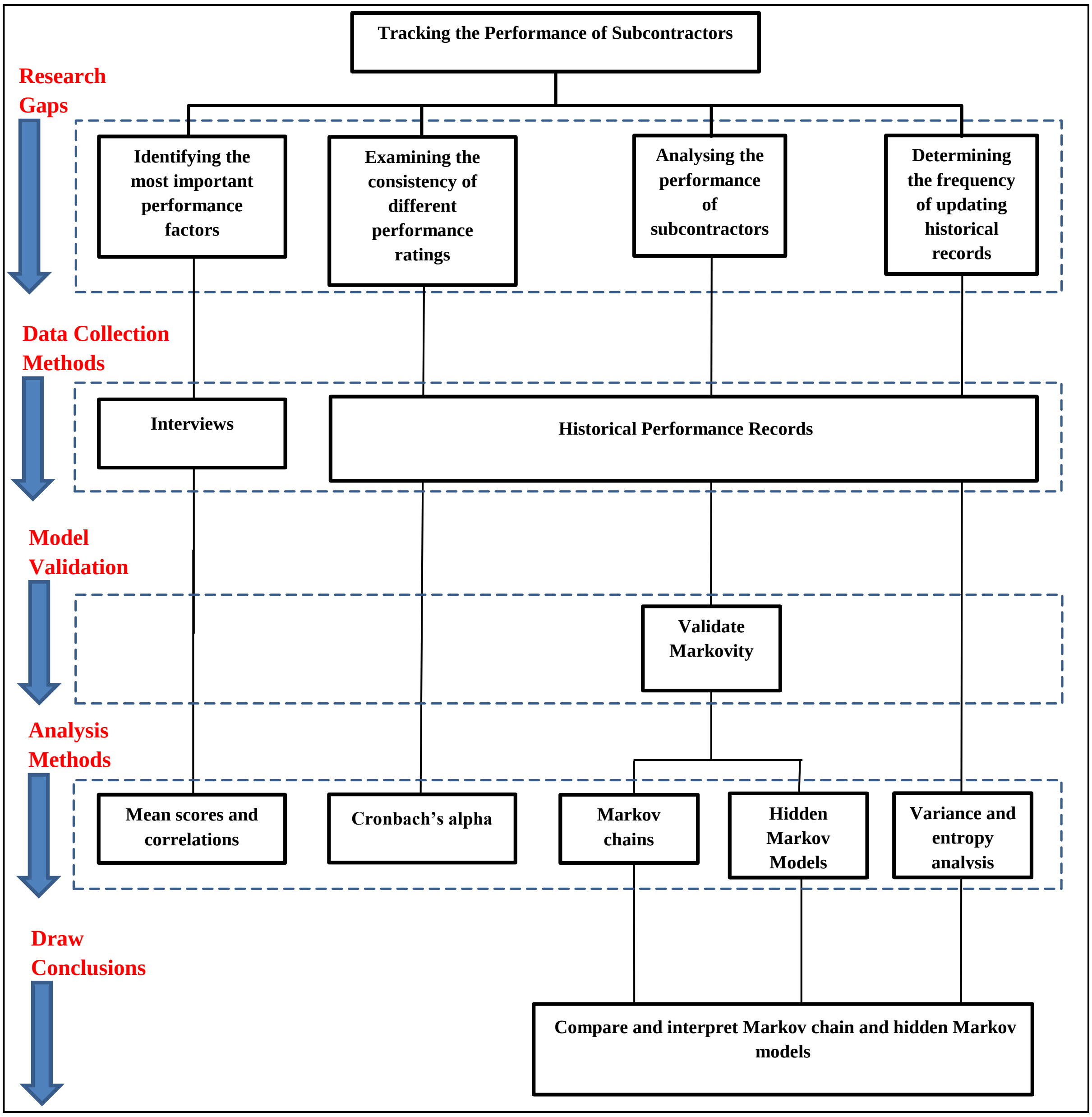

Figure 1.1. Research methodology of analysing the performance of subcontractors

1.4 Research Contributions

The main contribution of this research is the development of a new methodology for tracking the performance of subcontractors. This methodology will be validated by applying it to data collected from Saudi Arabia.

This application demonstrates that hidden Markov models can successfully be used to remove some of the randomness involved in performance measurements. The resulting underlying performance level is much more stable than the raw performance measurements.

A further contribution is the use of variance analysis and Shannon entropy analysis to determine the appropriate frequency of updating performance records.

1.5 Research Overview

To achieve the stated objectives and contributions, this thesis contains nine more chapters, which are:

Chapter 2: Provides an extensive literature review to identify all the different techniques and strategies that have been used in studying and analysing the performance in the construction industry, generally, and performance of subcontractors, specifically.

Chapter 3: Explains and justifies all methods used in collecting the data for the thesis.

Chapter 4: Tests if contractors are more critical when they were assessing the performance of subcontractors based on more important performance factors.

Chapter 5: Tests the consistency of different ratings of performance given by contractors and identifies the internal structure of the performance questions.

Chapter 6: Validates the applicability of using Markov chains to model subcontractor performance for rating sequences by either individual or multiple contractors.

Chapter 7: Analyses the performance of subcontractors based on the identified performance measures from the collected data set to forecast their performance change and the limiting probabilities for different performance levels based on their current performance data using Markov chains. It also determines the frequency of updating historical records based on the variance and Shannon entropy.

Chapter 8: Analyses the different levels of subcontractor performance using hidden Markov models. It also determines the frequency of updating based on the variance and Shannon entropy applied to the hidden Markov models.

Chapter 9: discusses the results from each of the previous chapters and draws them all together to highlight the contributions of the thesis.

Chapter 10: Presents the concluded remarks derived from this research, lists the limitations and provides recommendations for future research.

1 Chapter 2 Literature Review

2.1 Introduction

This chapter will review all studies related to the performance of contractors and subcontractors in the construction industry. Firstly it will address the importance of performance in construction projects and also to shed the light on the key performance indicators that are used to measure performance. Secondly, the literature will address different techniques and strategies used for measuring performance in different countries. Thirdly, the literature will discuss the need and existence of public registries of past performance of subcontractors; and their uses for selecting subcontractors for future projects. Finally, the use of Markov chains and hidden Markov models for modelling performance of people or organisations is discussed.

All of the studies that address different strategies and techniques used to identify or analyse the performance in the construction industry are summarised in Table 2.2.

Table 2.1 Summary of the different studies in the literature that use different strategies and techniques to indentify and analyse the performance in the construction industry.

Study Area	Authors	Year	Aim/s of the Study
The Importance of Performance in the Construction Industry	Alfeld	1988	Identifying the importance of past performance
	Oluwoye et al.	1996	The importance of differentiating the effectiveness and efficiency of the performance.
	Cox et al.	2003	The use of performance in construction projects in UK.
	Bassioni et al.	2004	The importance of performance in UK construction Projects
	Cooke and Williams	2013	The risk of subcontracting a large amount of a construction project.
Key Performance Indicators (KPI's)	Cooke-Davies	2002	How the KPI's are used to decide the success or the failure of a construction project.
	Cox et al.	2003	Exploring the possible groups of KPI's in UK projects and how to measure them.
	Chan and Chan	2004	Exploring the possible groups of KPI's projects and how to measure them.
	Horta et al.	2009	Addressing some limitations of KPI's and proposing a new method to measure them.
Study Area	**Authors**	**Year**	**Aim/s of the Study**

Key Performance Indicators (KPI's)	Skibniewski and Ghosh	2009	Proposing a new method for dividing the KPI's and identifying their importance in taking decisions to prevent the failure of a construction project.
	Haponava and Al-Jibouri	2009	The use of KPI's in the early stages of a construction projects.
	Yuan et al.	2009	Using KPI's in establishing a framework for assessing the performance in Public-Private Partnership (PPP) in construction projects.
	Ogunlana and Toor	2010	Dividing the KPI's in more than two groups and discussing the results of Cox et al. 2003.
	Radujković et al.	2010	Identifying the most important KPI's in Eastern Europe.
	Ali et al.	2013	Identifying the most important KPI's in Saudi Arabia
Critical Success Factors (CSF's)	Sanvido et al.	1992	Identifying the importance of CSF's in construction projects.
	Chua et al.	1999	Introducing two new approaches to identify the importance of intangible CSF's.
	Chan et al	2004	Identifying the most important CSF's in Hong Kong in PPP to minimise conflicts and to improve the project performance
	Li et al.	2005	Examining the importance of unclear CSF's in UK.
	Lu et al.	2008	This study conducted to reduce the large number of CSF's that were used in China.
	Aksorn and Hadikusumo	2008	Identifying the CSFs that influence the implementation of the safety program in Thai.
	Ng and Tang	2010	Establish a set of CSFs in Hong Kong to improve the performance of subcontractors from the organisational and project perspective.
	Hwang and Lim	2012	Identifying the CSF's in Singapore.
	Williams	2016	Studying the multidimensionality of CSF's.
Performance Measurement Techniques for Selecting Subcontractors	Ng et al.	2003	Examining the importance of different factors for selecting subcontractors in Hong Kong.
	Derek Lavelle et al.	2007	Studying the effect of using the price as the dominant factor for selecting subcontractors compared to other important factors such as past performance.
	Maturana et al.	2007	Addressing the risk of selecting subcontractors with the absence of past performance information.
	Eom et al.	2008	Testing the use of Balance Scored method to select subcontractors in Korea.
	Arslan et al.	2008	Proposing Web-based sub-contractor evaluation system to select subcontractors in USA.
Performance Measurement Techniques for Selecting Subcontractors	Hartmann and Caerteling	2010	Setting 4 criteria to be used in selecting subcontractor for construction projects in Singapore.
	Alencar and De Almeida	2010	Establishing a multi-criteria group decision model for selecting subcontractors for projects in Brazil.
	Cheng et al.	2011	Presenting an evolutionary fuzzy hybrid neural network to evaluate the performance of

			subcontractors during the project period.
Study Area	**Authors**	**Year**	**Aim/s of the Study**
Performance Registries.	Ng	2007	Studying the importance of Voluntary Subcontractor Registration Scheme (VSRS) in Hong Kong
		2008	Studying the importance of the Singapore List of Trade Subcontractors (SLOTS) in improving the performance of subcontractors.
	Ng et al.	2009	Providing a model for selecting subcontractors that shows the importance of each factor.
	Oak	2012	Investigating how reputation is reflecting the performance of subcontractors.
	Spagnolo	2012	Studying the use of historical performance records in selecting subcontractors in public and private projects in Europe.
	Costa and Tavares	2013	This study showed that how subcontractors used their reputation when information about their past performance were not available.
Subcontractors and Partnership	Famakin et al.	2012	Studying the effects of the partnership on the performance in construction projects in Nigeria.
	Meng	2012	identifying the influence of the supply chain on the project performance and o assess the impact of the supply chain relationship on construction performance in UK projects
	Emuze and Smallwood	2014	Identifying the level of collaboration among project partners in South Africa and its effects on performance.
Subcontractors and Partnership	Eom	2015	To identify the effects of the long term partnership between contractors and subcontractors on their performance in construction projects in Korea
	Lee et al.	2017	Examining a proposed framework for to be used by Korean subcontractors in international projects where the performance considered significantly important.
New Entrants	Artto et al.	2008	Examining the risks of the relationship between contractors and subcontractors on the performance of construction projects.
	Lau and Rowlinson	2009	Identifying what affect the building trust on the relationships in the construction industry.
	Autry and Golicic	2010	Testing the hypothesis that that past performance of any two parties would provide more accurate measures of predicting future performance.
	Ochieng and Price	2010	Identifying the factors that affect the communication in international projects between two different multicultural mangers in Kenya and UK.
	Spagnolo	2012	Examining the possibilities of selecting new subcontractors that have no past performance records.
	Costa and Tavares	2013	Introducing an Expected Future Performance Reward instead of past performance to make equal chances for all applicants to be considered in selection phase.

2.2 Importance of Performance Measurement

The performance in construction projects is an important tool that has different aspects used in measuring the success of a construction project. There are different studies addresses the importance of the performance in the construction industry, specifically for subcontractor. These studies show that how the measuring the performance is important where it has been used as a comprehensive tool to track the progress of any task in a construction project to ensure that the required goals of the projects are met.

According to Bassioni et al., (2004), performance measurement has been increasingly used in the UK construction industry using different techniques and strategies,. These performance measures are selected based on project objectives. Project success is therefore tied to these performance measures. Cox et al. (2003) points out that performance is a more comprehensive measure that includes productivity as one of the factors to be considered. However, Oluwoye et al. (1996) notes that the importance of the performance should involves a better understanding of the differences between effectiveness and efficiency. The difference is that the effectiveness is related to the fulfilment of the explicit and implicit goals, whereas efficiency is related to minimizing the resources used to achieve those goals.

According to Alfeld (1988), identifying past performance is also important where it provides a baseline to benchmark and measure future performance against. Any variation in the performance level compared to this baseline can be measured. The variation could be positive or negative. Further management techniques should be used to determine the root causes of this variation to help with determining how to either enhance or eliminate the variation (Cox et al., 2003). As the importance of performance in construction projects has a crucial importance to ensure the success of a construction projects, this leads to the fact that more risk is associated with the main contractors since a large amount of work is subcontracted, (Cooke and Williams, 2013). Therefore, the main contractor should efficiently manage the performance of subcontractors to assure that the expected performance is achieved.

To summarise these studies, the importance of performance involves measuring them within the projects as there are some risks have to be monitored to ensure the success of any construction project. The importance of past performance as a factor used in selecting subcontractors will be investigated in the interviews that will be conducted with different organisations.

2.3 Key Performance Indicators (KPIs)

Different studies focus more on the Key Performance Indicators which have been widely used as measure of performance in construction projects. Cox et al. (2003) claims that accurate analysis of construction performance can only be achieved when the key performance indicators are determined and monitored. These key indicators are defined as the compilations of data measures used to assess the performance of a construction operation. The evaluations of the key performance indicators are used to compare the actual and estimated performance of a particular task in terms of effectiveness, efficiency and quality with regard to workmanship and product. They divide construction project performance indicators into quantitative and qualitative indicators. Ten quantitative performance indicators are addressed: units/man-hour, $/unit, cost, on-time completion, resource management, quality control/rework, percent complete, earned man-hours, lost time accounting and punch list. The qualitative performance indicators are safety, turnover, absenteeism and motivation. The quantitative indicators have crucial importance in comparison with the qualitative ones. That is because the quantitative indicators can be measured in terms of dollars, units or man-hours, such as cost, on time completion and percent complete. On the other hand, the qualitative indicators are not as reliable to measure performance because of their perceived difficulty or inability to be measured.

Chan and Chan (2004) divides the Key Performance Measures into two groups, objective and subjective measures, where the first group is measured using mathematical equations and the second group is measured by asking the opinions of the stakeholders. The result of this study validates the results of the study conducted by Cox et al. (2003). It divides the Key Performance Indicators into two groups where one group is about the quantitative indicators that are measure mathematically and the other group is about the qualitative indicators that are hard to measure where asking the stakeholders about their opinion is suggested to resolve this difficulty.

In studying the difference between the success factors, which help in achieving the success of a project, and the success criteria of a project, which are used to decide the success or failure of the project, Cooke-Davies (2002) defines the key performance indicators as factors used in project success criteria that are used in deciding the success or failure of a project.

Ogunlana and Toor (2010) discusses the recommendation of Cox et al. (2003) to identify a common set of key performance indicators to be used in measuring the project performance at the project level. They claim that this recommendation is difficult and impractical to be generalised because every project has some certain and unique features and limitations. On the other hand, they suggest that the key performance indicators on different types of projects should be comprehended in terms of sharing the benefits and overcoming drawbacks from these projects to be then helpful in expanding the list of KPIs for future projects. Based on that, they conducted a study to investigate if there are different perceptions of KPIs between different stakeholders, clients, consultant and contractors, on mega construction projects. The results of this study show that there are different perceptions on traditional KPIs like time, budget and meeting the specifications, where these three indicators are classified as the three elements of the iron triangle named by Atkinson (1999). However, there is a sort of agreement on some of the qualitative indicators like minimised disputes and stakeholders' expectations. On the other hand, the general ranking of the KPIs for these stakeholders does not reflect the actual ranking for them as groups. There are some indicators that have significant importance for some groups while other groups report different indicators being the most important. For example, the Efficiency is the most important for the client whereas Safety is considered to be the most important for the design consultant. The results of Ogunlana and Toor (2010) validates the result of the study conducted by Skibniewski and Ghosh (2009) where it suggested that more studies are required to study the possibility of generlising the KPI's.

Horta et al. (2009) discusses two important limitations of the KPIs in their study on construction companies. They state that regardless of the general acceptance of using key performance indicators, there are some theoretical and empirical limitations to their use. The first limitation is that each of the KPIs only examines one aspect of the activity. This means that the overall performance relies on the analysis of several indicators. Therefore, each KPI assesses one dimension. Where there are some activities that have multi-dimensions to measure their performance it becomes difficult to compare their performance. This problem was solved by normalising and averaging the scores of the individual KPIs to obtain a composite measure. The other limitation is that the KPIs do not directly lead in a straight forward manner to improvement targets because each indicator needs to be compared to some benchmark value regardless of the other aspects of the activity that are not accounted for in that indicator. In addition, if any indicator for an activity has a poor result then action is required to achieve the target level but the effect of this action cannot be estimated confidently. Therefore, they introduced a method to be used in assessing the overall performance of a company in construction by combining the KPIs with a frontier method data envelopment analysis (DEA). Their study assessed the organisational performance and operational performance separately for a group of Portuguese contractors. The results of this study show that it is important to use separate KPIs for measuring these two different performance types so that remedial activities can be targeted more precisely.

Skibniewski and Ghosh (2009) claims that all research about KPIs focuses on industries or measures functional performance for a specific situation while there is no focus on defining a general framework of different types of KPIs. Another point they stress is that the KPIs should influence a business decision in some time scale and that makes the decision process difficult if there is no time constraint for the decision makers. Thus, they proposed that the KPIs be divided into two groups: Hard time related KPIs and soft time related KPIs. The hard time KPIs indicate any critical phase that needs to dealt with immediately (within one day), whereas the soft time KPIs indicate areas where the project should be improved, but where in the meantime the project can be continued with the degraded performance status. As a result, this study introduces a new method of how to take a decision in dealing with different KPI's according to their importance.

According to Haponava and Al-Jibouri (2009), KPIs have been developed in recent years by including the measurements of other aspects of project performance. However, the majority of these developments are mainly used for benchmarking purposes whereas controlling the performance during projects has been rarely developed. Moreover, controlling the performance during projects is essential in all different stages of construction projects. As a result, their study aims to identify the KPIs that provide control of the process in the early phases of the project. The result of this study is that there are five significant KPIs relevant for process control. These are: initial problem definition, management of client requirements, alignment of stakeholders' requirements, design solution and stakeholder involvement.

Yuan et al. (2009) states that there is less attention paid in the public-private partnership (PPP) to the significant impact of the process factors that may affect the performance. In addition, the need to improve the performance of the process is identified by the performance objectives and KPIs. Therefore, the study they conducted using goal setting theory selected 15 performance objectives to find out the significance and differences of these objectives. The results show that of the 15 objectives all of the interviewees agreed on the importance of 11 objects. For the remaining four objectives (budget constraints, risks, revenue and guarantees) the support by the interviewees was mixed. Therefore, a framework for KPIs is established to be used in assessing the performance of PPP projects.

(Radujković et al., 2010) conducted a study of more than 30 south eastern European construction companies to determine which KPIs are important. There was a low level of awareness of KPIs by the participants in the study resulting in markedly different perceptions of KPIs and a general lack of recognition of their importance in managing the performance of the process. These problems resulted in identification of 37 KPIs. However, further investigation confirmed that they mainly focus on the iron triangle elements from Atkinson (1999), time, cost, and quality. The study recommends conducting more studies to find a common framework of KPIs.

In Saudi Arabia, a study was conducted to identify the key performance indicators to be implemented in terms of measuring the performance at the company level, (Ali et al., 2013). The data of this study was collected from a survey, which contained 47 key performance factors identified from the literature, was sent to a selected sample of large size construction companies. The results show that there were 10 important key performance indicators contribute to the success of construction projects. One important result of this study is that the importance of financial measurement has less important compared to other key performance indicators. In contrast, external customer satisfaction, safety, business efficiency, and effectiveness of planning have more importance as key performance indicators. This study is benchmarking a system that is expected to improve the performance of Saudi Arabic construction industry.

Therefore, the studies in this section of identifying the key performance indicators in different countries show that they have been testing different objectives in terms of achieving the success in construction projects for different players such as contractors, subcontractors and owners. They also tested different measures in terms of shortlisting the KPI's that may help in meeting the minimum accepted level of performance. This shows that the exact list of KPI's used to track performance in this research project is not critical since there is little agreement between different authors. This problem will be identified in the interviews conducted with some organisations that have historical records to find out the most important performance factors from a common performance historical record.

2.4 Critical Success Factors (CSF)

In this section, the studies related to the critical success factors in managing construction projects will be identified to be then compared with Key Performance Indicators to find out the differences between them.

Sanvido et al. (1992) defines the critical success factors as those used to predict success on projects and represent managerial or enterprise areas in a way that give special and continual attention to ensure high performance of an organisation or a manager. Thus, there were different studies have been conducted to identify those CSFs on different aspects for construction projects to ensure the success and maintain high performance.

According to Chua et al. (1999), there were some research conducted using quantitative measures of different factors to identify the critical success factors for projects success. However, these different factors are only limited to the project management efforts where they cannot be comprehended to intangible factors or not to be used on the absence of the data of hard performance. The other way of identifying the CSFs can be based on expert opinions where they use the experience they have to legitimate the listed CSFs and then testing them. Two approaches have been used to identify expert opinions on factors contributing to the project success. The first approach was used in Chua et al. (1999) to measure the intangible Critical Success Factors. Experts were asked to list the factors they considered were critical. This was the same approach used by Chan and Chan (2004) to measure the importance of qualitative Key Performance Indicators. However, consistency of KPI's identified by this approach is lacking due to the different names and scope of the selected success factors by experts. Therefore an alternative approach has been used to give the experts a list of success factors to assess their importance using a prepared scale (Chan and Kumaraswamy, 1997). The limitation of the second approach was that there is no assurance of the consistency of experts assessments if there list of factors has more than few factors. Thus, Chua et al. (1999) conducted a study to identify the CSFs for construction projects using the analytical hierarchy process (AHP) based on the expert knowledge and judgment in the industry. Their study involved 67 factors divided into four main project aspects that are: namely, project characteristics, contractual arrangements, project participants, and interactive processes. Since the previous studies of identifying the CFSs ranging from general to more specific strategies, they decided to use of the AHP. The results of their study showed that there were three important success project objectives which are: budget performance, schedule performance and quality performance. These three objectives were all equal in importance because their relative weights were comparable. The final results of their study show that there were different agreement levels of experts on some sets of CSFs where experts stressed on the importance of project characteristics and contractual agreement in assessing the CSFs.

In UK, a study conducted for construction projects of public-private partnership (PPP) through Private Finance Initiatives (PFI) to identify the CSFs which were not entirely clear, (Li et al., 2005). The survey has 18 success factors to examine their importance using mean score values. The results of this study show that there were three most important success factors, which are: strong private consortium, appropriate risk allocation and available financial market. On the other hand, there are two less important factors for project success, which are: shared authority and social support. These 18 factors were divided into five groups that are considered as the five elements of CSFs for PPP/PFI in UK. These five groups are: effective procurement, project implementability, government guarantee, favourable economic conditions and available financial market. Thus, the identification of the most important success factors would give an insight for improving the performance of projects in UK to achieve the expected success.

In Hong Kong, there was a study conducted by Chan et al. (2004) to find out the CSFs for partnering in construction projects to minimise conflicts and to improve the project performance. The study introduced 10 factors that were extracted to 41 variables using factor analysis. The results of the study show that there are five most important factors contributing to the success of the partnering in Hong Kong construction industry. These five success factors were: the establishment and communication of a conflict resolution strategy, a willingness to share resources among project participants, a clear definition of responsibilities, a commitment to a win-win attitude, and regular monitoring of partnering process. However, the results of this study could be used as a benchmark measure for future projects in Hong Kong to need to test its validity.

Lu et al. (2008) conducted a study on contractors in China in terms of identifying CSFs that may allow reducing the wide number of different factors used in achieving the project success to make them less but more manageable. They used 48 CSFs to identify the most important factors for contractors. The results showed that there are 35 critical success factors rated as critical in determining the competitiveness of a contractor. These 35 factors were divided into eight groups, which are: project management skills, organization structure, resources, competitive strategy, relationships, bidding, marketing, and technology. These eight groups were named as supercritical success factors (SCFS). They claimed that these 35 factors will help contractors in improving their competitiveness while they are working with limited resources. However, Lu et al. (2008) stated that the result of their study has to be periodically updated due to the rapid development in the Chinese construction industry. Another issue regarding the result of this study is that even though the 48 CSF's were reduced to 35 factors, this is still too many, requiring high effort to monitor them to ensure the success of the construction projects. As a result, it is recommended that the number of Critical Success be reduced.

Aksorn and Hadikusumo (2008) conducted a study in Thai construction industry to find out the CSFs that influence the implementation of the safety program. Their study had 16 factors to be ranked by respondents from 80 medium to large scale construction projects. The results of the study show that the management support is the most important success factor that influences the safety program in Thai construction projects. Also, the other 15 success factors have similar importance where no one could be ranked as less important or to be neglected. These 16 factors were then grouped in four categories to tested using factor analysis. These four categories were: worker involvement, safety prevention and control system, safety arrangement and management commitment. To ensure the reliability of the results of their studies, three case studies were conducted to test these 16 factors. The results proved that the ranking of these factors is meeting their importance which resulted in addressing the influence of the standard of safety performance in Thai construction projects has been recognised and implemented.

Ng and Tang (2010) conducted a study on labour intensive subcontractors in Hong Kong to establish a set of CSFs. The aim of this study was to improve the performance of subcontractors from the organisational and project perspective. The importance of this study to subcontractors is to help them in achieving a satisfactory outcome to be more competent in the construction industry. This study used 29 CSFs to assess their importance; the results show that there are nine CSFs that had been more agreement on their importance, which were: time completion, profit, programme/planning, cash flow, management level leadership, relationship with main contractors/client/consultant, staff team spirit, staff qualification/skill of labour and growth in revenue. After that, a factor analysis was conducted resulting in having three critical CSFs, which were: timely completions, profit, programme/planning, cash flow and management level leadership. Finally, the CSFs results in this study were grouped in three categories: managerial performance, financial performance and labour-intensive specific factors. This method that Ng and Tang (2010) have used to reduce the number of CFS's would be suitable to deal with the afore mentioned problem in Lu et al. (2008) in having too many CFS's.

In Singapore, Hwang and Lim (2012) conducted a study to identify CSFs based on the objective of different project players. So, 32 CSFs were identified and then classified into four groups: project characteristics, contractual arrangements, project participants, and interactive processes. This study was analysed using analytic hierarchy process (AHP) to then find the critical success factors out of the 32 identified ones. The results of this study showed that there were different success factors considered by consultants, owners and contractors. Consultants considered the quality and owner satisfaction as the most important whereas contractors gave more importance to the time completion and owners for quality and schedule conformance.

The definition and importance of the success factors for construction projects have more focus from different researchers, however, the multidimensionality of the success factors where they are combined in complex interactions have not been studies, (Williams, 2016). This study shows how success factors contribute to the performance of construction projects and how the paths of the projects could be mapped and analysed from root causes to success criteria. The results show that the project success has different interacting criteria that need different interacting factors. The overall results have no different success factors but it has more detailed process in rooting the causes of projects so the risks will be identified and eliminated. One result of this study regarding the performance shows that it is affected by the culture, communication and stakeholders' engagement. However, this result is hard to be generalised to short term view.

Therefore, the studies in this section of identifying the critical success factors in different countries show that they have been testing different objectives in term of achieving the success in construction projects for different players such as contractors, subcontractors and owners. They also tested different measures in terms of shortlisting the CSFs which may help in either developing or maintaining high performance to be targeted.

Based on different studies in the previous section and this section, the common goal of KPIs and CSFs is to achieve project success. However, the critical success factors have more focus on the critical aspects of the project that need to be identified in term of ensuring the success of the project so high performance is expected to be maintained. On the other hand, the key performance indicators are tools used to control the success of the project by maintaining and controlling the minimum accepted level of performance. Moreover, achieving the success of projects starts by identifying the CSFs and then to be controlled by the KPIs.

2.5 Performance Measurement Techniques

Hartmann and Caerteling (2010) specify four different criteria for selecting subcontractors: price, technical know-how, quality and cooperation. A survey was conducted in Singapore to test the hypothesis that subcontractors vary in their performance across these criteria in different projects and that the contractor has to balance the criteria to select the most appropriate subcontractors. The result of the survey is that price is the dominant criteria that contractors mostly rely on when selecting subcontractors.

Although the criteria used in Hartmann and Caerteling (2010) are important, it seems that for research purposes a larger number of criteria should have been investigated.

Eom et al. (2008) shows that a model delivered from the balanced scorecard (BSC) could be used by contractors to select subcontractors based on a strategy that evaluates the subcontractors. This strategy is proposed to be used for contractors that plan to have long-term partnering. It is structured based on the construction industry in Korea. The four perspectives of this study in evaluating subcontractors are finance, services, process and improvement. All of these perspectives have sub-criteria to analyse the expected performance of a subcontractor in the four different categories. The outcomes of this model are that main contractors place primary importance on subcontractor service and financial stability, while subcontractors care more about their own technical capability, competitiveness and growth.

The Web-based sub-contractor evaluation system (WEBSES) was proposed to improve the process of selecting subcontractors and save time when using multiple criteria. This is proposed as an alternative the common head contractor strategy of focusing exclusively on lowest bid price, (Arslan et al., 2008). The paper describes the application of WEBSES in a mid-sized construction company in USA. The main criteria, each containing their own sub-criteria, examined in this study were: cost, time, quality and adequacy. The results of this study for Arslan et al. (2008) are that the selection of subcontractors is improved when the importance of bid price is reduced.

Alencar and De Almeida (2010) establish a multi-criteria group decision model for selecting subcontractors for projects in Brazil. This model is managed by the contractor, client and consultant to analyse the selected subcontractors for the bid and then to evaluate them according to the criteria of cost, quality, time, culture, design and experience. This model uses lowest maximum regret for selecting subcontractors.

Derek Lavelle et al. (2007) claims that price is not always the dominant factor in evaluating and selecting subcontractors based on several studies. The result of this study was that the health and safety records and past performance of a subcontractor have more importance in comparison with price, which was ranked as the third factor. This result was validated by Arslan et al. (2008) which showed that the importance of price should be reduced in selecting subcontractors for a construction project.

Assessing the performance of subcontractors is an issue, particularly when there is no information about past performance. The risk here is that poor performance may extend the completion time of the project (Maturana et al., 2007).

Cheng et al. (2011) presents an evolutionary fuzzy hybrid neural network to evaluate the performance of subcontractors during the project period depending on twelve factors. This information is then available to be used as a reference in selecting subcontractors in the future. However, this technique has difficulties in evaluating new subcontractors due to the lack of a previous evaluation.

Ng et al. (2003) examines 26 factors identified from a web-based search for different countries. The results show that there are ten key criteria to be used in identifying suitable subcontractors for construction projects. These criteria were based on a survey of construction industry practitioners in Hong Kong. These criteria were ranked by clients/consultants, contractors and subcontractors. There was a slight difference in assessing the importance of selection criteria by each group in Ng et al. (2003). The consultants were more interested in quality while contractors more interested in reducing the contractual risks. The subcontractor group was interested in increasing their own competitiveness. These different attitudes towards the importance of selection factors might increase the risk of disputes between these three parties. It might also affect the success of construction projects where other important factors have been identified. This result would also be investigated in Chapter 4 to find out whether different important performance factors would affect the overall performance assessment of subcontractors.

2.6 Performance Registries

Past Performance is an indicator of work quality and professionalism and contributes to a subcontractor's reputation. For example, (Ng et al., 2009) provides a model for selecting subcontractors where the reputation of the subcontractor is the first factor in the Critical Success Factors (CSFs). Similarly, according to the model used by Oak (2012), reputation plays a key role in selecting a subcontractor by a principal contractor. This reputation also increases the opportunity for a subcontractor to win jobs since it reflects a positive indication of high performance and the history of a company. However, Costa and Tavares (2013) states that subcontractors with long experience may not have records indicating their performance, because they are relying on their reputation. Of course, the problem that then arises is if the subcontractor uses image management to cause their reputation to be better than their performance deserves.

To keep track of past performance of subcontractors, some countries have established registries containing historical data about the performance of subcontractors in construction projects. Below are some examples of these registries.

Hong Kong has established a Voluntary Subcontractor Registration Scheme (VSRS). This was developed based on the recommendations of the Construction Industry Review Committee in 2001 (Development Bureau, Hong Kong). The aim of this scheme is to build up subcontractors who are capable, responsible and have specialized skills and high standards of professional ethics. Clients are legally required to select contractors who collaborate with registered subcontractors. Some clients require contractors to state the name of subcontractors during the tendering stage, Ng (2007) .The importance of the VSRS registry comes when subcontractors are suspended from tendering or removed from the registry when they fail to perform satisfactorily.

In Singapore, the Singapore List of Trade Subcontractors (SLOTS) is a centralized registry of subcontractors recording five attributes: company status, personnel resources, financial capability, track record and performance (Kim and Huynh, 2008). It is a requirement of certain clients to employ contractors that only deal with registered subcontractors in Singapore. Moreover, subcontractors who fail to perform satisfactory work may be suspended or removed from the registry. This registry is effective because it centralizes its data, hence making it available to more participants, and also minimizes the time and cost for contractors and subcontractors respectively.

In Saudi Arabia, the Deputy Minister for Contractors classification maintains historical records assessing and classifying both contractors and subcontractors. It measures performance according to six factors: Project Management (planning, organization and follow up), Work Quality and Compliance with the specifications, Compliance with the time schedule, Project Staff Level (Competence, Experience, Qualifications), Availability of the necessary equipment and systems and extent of their efficiency, and

Application of the security and safety procedures. Each factor is assessed using a five point scale from weak to excellent. For work on government projects the contractors are required to submit assessment of their subcontractors. For private projects subcontractors can request that their contractor do the same. Records are retained for four years. Based on these assessments each subcontractor is given a public rating.

Some countries only have registries for contractors and not subcontractors. For example the United States has set up a common platform to be used for future selection through The Federal Acquisitions Streamlining Act. The US government relies on past performance of contractors to select them for public projects and share their information through a common platform to be used for future selection. In EU, the use of past performance in selecting a contractor in public projects is limited, (Spagnolo, 2012). However, the past performance is considered in private procurement by different mechanisms. One main significant drawback for selection based on reputation is the lack of availability of past records. This highlights how important it is to keep past performance records.

2.7 Subcontractors and Partnership

In this section, all studies related to the effects of the partnership on the performance of the construction industry will be shown.

Famakin et al. (2012) studies the factors that affect the performance of partners in construction projects in Nigeria. The reason of conducting this study was that the construction industry has an increase pressure to have better options to overcome the challenges and critical issues that it has battling. A survey was conducted to collect data from partners and consultants in terms of assessing the importance of 20 factors that contribute to performance of partners in the construction industry. The results show that all of the factors have significant importance to the performance of partners with more importance given to communication, compatibility of objectives and mutual understanding among partners. This study provides a method of identifying the factors that contributing more significantly to the performance of partners.

The poor performance in construction projects in terms of time delays, cost overruns and quality defects was studies in UK construction industry to identify the influence of the supply chain on the project performance, (Meng, 2012). This study was conducted by sending a survey to assess the impact of the supply chain relationship on construction performance. The survey consists of 10 key indicators areas. The results of this study show that the relationship between projects parties was deteriorating. This possibility of resulting of poor performance because of this deterioration between the project parties is high. To reduce the possibility of poor performance, the strategic relationship between project parties is proposed to be adopted instead of traditional relationship. However, the adoption of the strategic relationship will not assure better performance unless more effort is to be undertaken. The strategic partnership is recommended to be practised in the long term to influence the performance of construction projects more positively.

A study was conducted in South Africa to identify the level of collaboration among project partners, (Emuze and Julian Smallwood, 2014). The studied partners were consultants, contractors, subcontractors, suppliers and manufacturers. The reason of studying these partners was that the performance of the construction industry was generating negative captions. These captions were about defects, reworks, injuries, delays and accidents. Therefore, the study was conducted using a quantitative survey. The results of the study show that there are different reasons justify the negative performance of the construction industry in South Africa. These reasons are poor problem solving of the exiting between project partners, poor use of modularisation, high number of different clients and lack of understanding the contract terms. However, this study only provides an insight into the poor performance because of the limited numbers of respondents. This study recommends that contractors have to do more effort in terms of overcoming the difficulties affecting the performance in any construction project with more emphasis on improving their relationships with subcontractors and suppliers.

Eom et al. (2015) conducted a study to identify the effects of the long term partnership between contractors and subcontractors in Korea. The main reason of this study was because of the important role that subcontractors have in the construction industry. So, identifying a new system to improve their relationships with contractors has a crucial importance. To achieve this aim, there were 7 partnership factors have been identified to be studied: subcontracting strategy, performance improvement, process innovation, information sharing, cooperation in collaboration, standardization of selection, and feedback of evaluation. The study was conducted by surveys and interviews. The results show that there is an agreement between contractors and subcontractors about the importance of these 7 factors. However, there is still a lack in executing them by that time. Thus, an e-procurement system was proposed to overcome this problem which aims to improve the collaboration between contractors and subcontractors. This improve is expected to have a positive effects on the performance of the construction projects because both parties need to continuously share information on regular basis and also need to avoid unnecessary information. On the other hand, this system is designed to for long term relationships, so, its applicability is limited.

Lee et al. (2017) found that the profits made by Korean subcontractors in international projects were low because of the lack of the financial and technical capabilities, shortage of experience, and lack of information. Therefore, their study aims to find out the possibility of implanting a win-win strategy for small and medium size Korean subcontractors in international construction projects by creating a framework. The framework considered two perspectives: the level of performance for subcontractors in terms of cost, schedule and quality, and the interface risks affecting the partnership between contractors and subcontractors. Determining the critical risks for both perspectives was studied by using 77 risk factors. The results of this study establish a two dimensional strategy matrix that considers the degree of the partnership between contractors and subcontractors; and the performance of subcontractors. This strategy matrix is expected to be used by subcontractors to assess any associated risks before start working a project with any contractors. This strategy will increase the possibility of the project success and improve the performance of subcontractors by making a good partnership. The result of Lee et al. (2017) is limited to Korean subcontractors where more studies are recommended to be conducted on subcontractors in other countries.

2.8 New Entrants

Using reputation in the selection of subcontractors phase for a complex construction project based on past tracked records may affect new entrants. The new entrants are not only new subcontractors, they could be subcontractors with short experience, subcontractors that have only worked in small size projects, subcontractors that do not have previous partnership with other contractors, subcontractors that have worked on different project types, and subcontractors competing for an international construction project. The opportunities of these new entrants to be selected in comparison with subcontractors who have long experience and partnership are reduced, (Spagnolo, 2012). Costa and Tavares (2013) suggests using expected future performance reward instead of past performance to make equal chances for all applicants to be considered in the selection phase. However, the potential variation in future performance is not considered and this is a big issue when past data is not available.

Artto et al. (2008) states that there has been less focus in addressing the risks of relationship the in the construction industry. As a result, a study was conducted to identify the risks of the relationship between contractors and subcontractors in terms of project business for contractors that resulted from inter-organisational of subcontractors. This study is limited to identify the risks in complex and dynamic projects. The data was collected by conducting several semi-structured interviews with two main global contractors. The results of this study show that there were four groups identified as risks sources from subcontractors based on the relationships of subcontractors with other subcontractors, the contractor's competitor the contractor's client and non-business actors. The four groups were contractor-subcontractor-subcontractor; contractor-subcontractor-competitor; contractor-subcontractor-client; and contractor-subcontractor-non business actor. The results of these four groups show that the relationship between contractors and subcontractors has only risks on two different layers. The first layer is about the temporary project where the risks are related to specific sales and delivery projects. The second layer is about the permanent business where the risks are related to changes in the position of the business players. Considering these two findings of this study, accepting new subcontractors to work with experienced contractors is expected to be difficult if these two risks are important to be overcome.

Building trust for some relationships in the construction industry is important because it reflects the quality of the relationship which will result in better project performance, (Lau and Rowlinson, 2009). However, Lau and Rowlinson (2009) state that the concept of trust has to be tested to provide better understanding of the factors affect it. As a result, a study was conducted by collecting data from 10 different partnering and non-partnering projects. One result of this study was that clients and contractors trust individuals whereas subcontractors and subcontractors trust organisations. The other result was that non-partnering projects would not have less trust than partnering projects. Thus, building trust between multi parties in construction projects has to consider more wide goals than economic and technology goals where middle managers have to acquire better understanding of trust since they undergo more relationship issues than the others. The use of new subcontractors might be encouraged as the results of this study show that trust in not related only for long relationship and trust is not limited to technological and economical goals.

The performance is generally expected to be better when two parties have strong relationships while they work on different projects, (Autry and Golicic, 2010). However, other research findings claim that past performance of any two parties would provide more accurate measures of predicting future performance. Autry and Golicic (2010) examines these two different opinions as they considered them snapshots that need to be strengthen by acquiring real data. Therefore, a study was conducted on 323 contractors-subcontractors relationships in the construction industry using strength performance spiral model. The results of this study show that a strong relationship between contractors and subcontractors is expected to improve the performance of construction projects. Moreover, the interactions of contractors with subcontractors will result in higher levels of performance since completing projects on scheduled time frame and budget are expected to be met. This result is applicable on both contractors and subcontractors. The findings of this study have some guidelines specifically for new subcontractors on how they could initiate new relationships with contractors and to strength their relationships to stay for long time in the market.

Trust and price have relative importance for contractors in selecting subcontractors for construction projects and also to assess their performance, (Hartmann and Caerteling, 2010). This was the result of a study conducted on Dutch residential building industry. One the other hand, this result is true when there are repeated partnering between contractors and subcontractors where contractors become more accurate in assessing the performance of subcontractors in terms of quality, technical know-how and cooperation in the past which they have strong contribution on future selection. One limitation of this study is that more studies have to be conducted to understand the trade-offs between price and trust that is made by contractors. One recommendation of this study is that subcontractors have to offer more competitive price to increase their chances to be selected by contractors. This recommendation of this study is one of the aspects that new subcontractors are expected to consider to increase their chances to enter the construction industry.

As stated in the introduction of this section, working in an international project considered as new entry for subcontractors. Ochieng and Price (2010) study the effect of the communication in international projects between two different multicultural mangers in Kenya and UK in terms of addressing the factors affect the communications.

The result of this study shows that successful communication between different mangers cultures would be achieved when the awareness of variations in different culture is demonstrated. So, the result of this study means that new subcontractors and contractors planning to work in new international projects have to consider the effect of understanding different cultures in terms of performing better which will result in assuring their success in international projects.

2.9 Markov Chain Processes

Markov chain is a stochastic process that generates two mechanisms, (Levinson et al., 1983). The first mechanism represents an underlying Markov chain that has a finite number of states. The second mechanism represents a set of random functions where each one is associated with each other state. Furthermore, Markov chain process deals with a case that has different number of states when it starts from one of these states and then moves to the next state. The movement from a state to the next one depends on the probability that called transition probability. The transition matrix could be used to forecast where the current state is irrelevant to one of the next states. That means the information of the current state is stationary or decayed and useless. One important property of Markov chain is to be absorbing, which means that the chain has at least one absorbing state. On the other hand, it would be transient if it is not absorbing.

According to Sirl (2005), the Markov Process is called a random process if its future is independent on the past. In addition, the Markov chain could be either time-homogeneous or discrete.

Markov chain process requires three conditions:

1- All states are finite.

2- The probability of each state is constant over time.

3- Each state is determined based on the previous state.

Markov chains had been used as a method to study to resolve the issue of predicting the location of individuals next location based on the observation of their driving behaviours over some period of time, (Gambs et al. 2012). This study had also used Markov chains to study the recent locations that the drivers had visited. In this study, a Mobility Markov Chains model has been introduced by developing a novel algorithm that can predict the next location of the individuals with respect to the previous visited location. The efficiency of the introduced model using three different database showed that the accuracy of the prediction using Markov chains ranged from 70% to 90%. As assessing the highway operation conditions have been usually measured either directly or by estimation, another study used Discrete Time Markov Chains (DTMC) to develop a model that can estimated travel time on freeway, (Yeon et al. 2008). In this study, the Markov model used two states to estimate the travel time which were whether the freeway was congested or not. Using Markov model in this study helped in comparing the expected travel time with actual where the result of the analysis showed that the estimates of the travel time using Markov model did not differ from the actual measured travel tine at the 99% confidence level.

More details about the Markov chains definition, Markov process and Markov properties are discussed in Chapter 7 to give better understanding for the readers of how this method was implemented in the thesis and how it could be used to model the performance data to then analyse it.

2.10 Hidden Markov models (HMM)

Hidden Markov models (HMMs) are a formal foundation for making probabilistic model of linear sequence and allows building any complex model by drawing an intuitive picture (Rabiner, 1989), (Durbin et al., 1998) and (Eddy, 1996). In addition, they help in identifying the sequence and the underlying state path, the labels. They deal with a large scale stochastic process and have successful application in different number of scientific areas such as engineering. It is a technique used in estimation and recognition. Its state duration is either a unit interval or geometrically distributed to make the underlying process Markovian (Yu and Kobayashi, 2003). It is a tool that is used for modelling time series data and represents the distributions of probability over observation sequences (Ghahramani, 2001). Hidden Markov model was derived from two assumptions. The first assumption is that the hidden states from observer, for some processes, generated the observation at time *t.* The other assumption is that the states of these hidden processes satisfy the Markov property, which is that the state at some time covers everything required to know about the history of the process and then to predict the future of the process. This means the current state of the process is independent of all prior states.

The extension of the hidden Markov model (HMM) is hidden semi-Markov model (HSMM) which is designed to allow general distribution for state durations., such as non-geometrical or non-exponential (Yu and Kobayashi, 2003). They claim that HSMM was firstly investigated by Ferguson in 1980. Whereas the conventional assumption of both HMM and HSMM is that there is one observable associated with the hidden state, there are multiple observation associated with the hidden state in some application. However, the difference between HMM and HSMM is the current state dependency. Moreover, the state in HMM the current state depends on the previous state through the most recent ones and then becomes a Markov process, but the HSMM is not. Based on the assumption of (Yu and Kobayashi, 2003), which is the state is called hidden when it is not directly observed, the observation patterns are classified into six types to estimate the observable output sequence. The six observation types are full, deterministic, random, state dependent, output dependent and multiple observations. Each pattern type has specific requirements to be considered in solving what required is.

The importance of the safety of the workers in construction sites has been studied before to identify any hazards that may cause an accident prone physical space surrounding the workers, (Rashid et al. 2017). These studies improved site safety for workers by using different techniques such as location-aware proximity sensing technique. However, the reliability of forecasting the impending hazardous scenarios before they occur still considered a major gap. As a result, data about of workers and site hazards were collected and then modelled with hidden Markov models to study the attitude of workers toward risk in construction sites and to predict their future positions; and detect imminent contact collisions. The result of this study showed that hidden Markov models was effective in robustly predicting potential collision events.

Hidden Markov models has been used to predict the human movement through the use of historical data of human locations, (Wesley et al. 2012). The proposed approach of this study clustered historical data of human locations according to their characteristics. Hidden Markov models were used in this study to deal with the location characteristics as unobservable parameters and also to deal with the effects of pervious actions of each individual. The result of this study showed that the accuracy of prediction using hidden Markov models was 13.85% for a small region. However, more studies can be conducted to compare the accuracy of hidden Markov models in small, medium and large region sizes.

More details about the hidden Markov models techniques and implementations are discussed in Chapter 8 to give better understanding for the readers of how this method was implemented in the thesis and how it could be used to model the performance data to then analyse it.

In conclusion, all of the studies in this chapter show how past performance is important in the construction industry. There were different techniques and strategies used to identify performance factors and also to analyse how important it is. However, the change of performance of subcontractors over time has been overlooked. The change of performance of a subcontractor has a crucial importance since subcontractors can increase their capability both in the range of tasks that it can perform, and in how well it performs its tasks. Similarly, if subcontractors do not practice particular skills then their performance in these areas can drop. This leads to the questions of how does subcontractor performance change over time and how can this be measured. Furthermore it leads to the question of how long performance records can be considered current.

Chapter 3 Data Collection

This chapter explains the methods used to collect all data used in the thesis. The explanation shows how and why set of data was collected. Then, the purpose of collecting the data will be linked to the problem statement of this research in terms of answering the research questions and hypothesises. At the end, the methods of analysing the data will be defined and the expected outcome would be stated.

3.1 What data?

As explained in the problem statement in the introduction, this thesis introduces a methodology for analysing the performance of subcontractors. This requires that firstly the most important performance factors are identified for use in the methodology and secondly that case study data be applied to the methodology to validate that it works. The data collected covered two important stages.

The first stage was to explore the importance of the performance of subcontractors compared to other factors that are used in selecting subcontractors, such as price. The different techniques used in assessing the performance of subcontractors were investigated by conducting an extensive literature review to identify the performance factors used in the performance registries in different countries. Collecting data to explore these stated goals also involved interviewing organisations and contractors in Saudi Arabia that have historical performance records to identify the performance factors used and to explore the techniques that they have been using to updated their historical records.

The other stage involved collecting performance data from historical records that was gathered from governmental performance records in Saudi Arabia because the performance of their subcontractors was selected as the case study. The data collected from that governmental historical record involved two data sets. More details about the reasons for collecting these two data sets are explained in the following sections of this chapter.

3.2 Exploratory Data:

3.2.1 Literature Review

As discussed in the literature review chapter, many researchers have considered the question of what performance attributes are used in selecting future subcontractors, how important they are, and how they are evaluated. The importance of the performance of subcontractors as one of the selection criteria has been identified in different countries, such as Korea, Eom et al. (2008), Singapore, Hartmann and Caerteling (2010), Saudi Arabia, Ali et al. (2013), and UK, Cox et al. (2003). These studies showed that the performance is one of the most important criteria in selecting subcontractors for future projects. However, these do not show details about how performance was evaluated and what criteria and scales were used to measure it. Instead they focused on comparing how important the past performance is compared to other issues, such as price. However, both Singapore and Saudi Arabia, for example, have government departments that collect and disseminate data on the performance of their subcontractors. The web pages of these government departments provide details about the procedures used to evaluate the performance of subcontractors and the criteria used, (Singapore Contractors Association Limited, 2017) and (DEPUTY MINISTRY FOR CONTRACTOR CLASSIFICATION IN SAUDI ARABIA, 2015). The scales used and the factors of performance are known but the methods used to aggregate the data are not revealed by the relevant agencies.

3.2.2 Interviews

The purpose of conducting interviews with organisations and contractors that have historical performance records was to explore the differences between the factors used in selecting subcontractors and which of these factors require past performance data. This purpose involved identifying the factors they use in assessing the performance of subcontractors and also the techniques used in rating their performance. The other purpose of these interviews was to identify the updating time that each organisation uses to update its historical performance records

Given that the literature review did not find many examples of organisations that collect data on subcontractor performance and that even for the ones that do the techniques used in assessing and updating the performance are unknown, it was decided to identify them by interviewing contractors who have historical records of subcontractor past performance. This would have the dual benefit of finding out which performance factors are considered important by most companies and how consistent the opinions regarding these levels of importance are.

The reason why interviews were used instead of a survey was because this was meant as an exploratory data collection. Thus it was considered preferable to use a semi structured interview format so that discussion could be had with the interviewees to better understand the context of their answers. It also enabled more detailed exploration of methods that they used that were not present in the literature. Thus since it was not important to get statistics of the overall population there was no need to run a survey and for the interviews a small sample size was sufficient.

3.2.1.1 *The relation of the interview questions to the research aims*

All of the 6 questions that had been used in the interview were constructed based on the research questions. These six different questions were:

1. *What would you consider to be the main factors that should be used when selecting subcontractors?*
2. *Which of these factors need information related to past performance?*
3. *Does your company keep and use formal records about performance for subcontractor selection?*
4. *How long do you expect the data on each factor to be valid?*
5. *How would these factors be measured?*
6. *How would you evaluate subcontractors for which you do not have any of this historical data?*

The first question (about the factors used in selecting subcontractors) was asked to identify the different selection factors used for selecting subcontractors between different organisations. This will establish whether performance is important to particular organisations. It is important because price is usually a primary selection factor and this question answers whether other factors are considered at all.

The second question was to find out whether any of these selection factors need past performance data. The thesis is all about how to use past performance data to predict future performance. If the organisations are not interested in past performance then this becomes irrelevant.

The third question confirms the answer to the second question. If the organisation claims that it bases selection on factors that require past data, but does not have or use past data for making the decisions then its answers are inconsistent. More importantly since it is an interview this question allows the interviewer to determine if the company keeps its own data or if it relies on external sources of data, such as government records. Questions 4, 5 and 6 would only follow if the question 3 was answered in the positive.

The fourth question was to identify how frequently the organisations update their records. The purpose of this was to collect data that could later be compared with the recommended frequency of updating resulting from the analysis of the case study data.

The fifth question was asked to identify different measures and techniques used by these interviewed organisations and contractors in terms of understanding how different they are to the proposed model in this research, which is a hidden Markov model.

The purpose of the last question was to identify whether these organisations and contractors have different processes or techniques of assessing the performance of new subcontractors. This is to address the problem of making decisions that normally require past data when there is no past data. This question was asked to also identify any differences between the assessments of registered subcontractors and new subcontractor. The hidden Markov model requires past data for prediction purposes, so this question investigates current methods of dealing with this problem.

The interview questions were piloted by three engineers in Saudi Arabia who have good experience in assessing the performance of contractors and subcontractors in Saudi Arabia. They did not add any more questions to those proposed as they considered them to be sufficient based on the stated aims of the research. However, one question was removed about the importance of price compared to performance as they stated that price is only considered in the selection phase.

The interview process and questions were submitted to the UNSW Human Research Ethics Advisory Panel "H" and approved with reference number 08/2013/91.

The interviews were conducted in Saudi Arabia in 2014. 12 large organisations were approached for the interviews. It quickly became apparent that 7 of the organisations had strategic relationships with subcontractors. Therefore their selection criterion for subcontractors was simply whether or not they had a strategic relationship with them and so the full interview was not carried out.

This left five different contractors and organisations that keep historical subcontractor performance records. These organisations and contractors that were interviewed are all large size companies with at least 500 registered subcontractors in their systems.

For each contractor an initial interview was held with a manager to explain the basic details of the information that was sought and get permission to obtain this data. Then a more detailed interview with someone in the company who is more familiar with the process was carried out to obtain more detail.

3.2.2.1 *Interviews Answers*

The answers of each interview question from the interviewed organisations and contractors will be discussed in this section.

Q1: ***What would you consider to be the main factors that should be used when selecting subcontractors?***

As shown in Table 3.1, the answers of question 1 show that past performance is one of the factors used in selecting subcontractors for a construction project. So this validates the findings in the literature review that past performance is one of the important factors in selecting subcontractors as founded by Alfeld (1988), Maturana et al. (2007), Derek Lavelle et al. (2007) and Costa and Tavares (2013)

Table 3.1. The selection factors that were identified from interviewed organisations in Saudi Arabia.

Organisations	Selection Factors
O1	- Past performance in similar previous projects. - Safety records. - Technical capability. - Capacity. - Financial condition. - Labor resources/compatibility. - Quality assurance program. - Environmental program. - Price.
O2	- Capability. - Cost. - Work quality. - Past performance in similar previous projects.
O3	- Technical Requirements. - Financial Consideration. - Past performance in similar previous projects.
O4	- Compliance to the Scope of work. - Work procedures and methods. - Compliance to the project schedule. - Safety records. - Past performance in similar previous projects.

	- Availability of Equipment. - References from previous contractors. - Organization Chart.
O5	- Past performance in similar previous projects. - References from previous contractors. - Resources (equipment and workforce) - Number of current projects under construction.

Q1a: ***Discussion for question 1***

During the interviews it became apparent that it would be useful to have the interviewees rank the importance of the performance questions used in the governmental records. Therefore each of the 5 interviewed organisations were asked this in the discussion associated with question 1.

These organisations were asked to rank the performance questions of the governmental record from most (5) to least important (1) based on their own interests. Results are shown in Table 3.2. These results will be used in testing whether contractors are more critical when assessing the performance factors that they consider to be most important in Chapter 4

Table 3.2. The ranking given to the performance factors used in the governmental record by other in Saudi Arabia

Performance Factors	**O1**	**O2**	**O3**	**O4**	**O5**
Project Management (planning, organisation and follow up)	1	1	2	3	1
Work Quality and Compliance with the specifications	2	5	3	2	4
Compliance with time schedule	5	3	5	5	3
Project Staff Level	1	1	1	1	1
Availability of equipment	3	4	1	1	2
Application of safety procedures	4	2	4	4	5

Q2: ***Which of these factors need information related to past performance?***

As shown in Table 3.3, the answers of this question show that organisations use a variety of factors requiring past information. Each organisation that used a particular factor in Table 3.3 for making selection decisions also recorded the necessary data. As can be seen there is a wide variety of factors considered. While four of the factors are used by four organisations each, there are no factors that are universally used.

Table 3.3. Selection factors that require past information used by organisations in Saudi Arabia

No.	Performance factors	Organisations				
		O1	O2	O3	O4	O5
1	Compliance to the project schedule.		●	●	●	●
2	Safety adherence	●	●	●	●	
3	Availability of equipment and resources	●	●		●	●
4	Compliance to the scope of work				●	
5	Work quality	●	●		●	●
6	Work procedures and methods				●	
7	Environmental program	●				
8	Compliance with the cost of the project		●	●		

Q3: ***Does your company keep and use formal records about performance for subcontractor selection?***

The answer for this question was "Yes" from all organisations interviewed. Therefore, the interview continued onto questions 4 and 5 for all of the organisations.

In practice this question was actually a confirmation of the interviewee selection criteria that organisations would be interviewed if they recorded and retained performance data.

As mentioned previously, 7 of the organisations that were approached do not retain historical performance records for subcontractors. This was justified by them because of the fact that they have strategic partnerships with subcontractors that usually work for them on all projects. Moreover, they consider those (fixed) subcontractors as partners that are familiar with all procedures and strategies used by the contractors. Therefore, using the same subcontractors will save them time, cost and rework.

Q4: ***How long do you expect the data on each factor to be valid?***

One of the central questions of this research is how long data is valid for. As a practical matter it turned out in the interview discussions that it was simpler for the organisations to say how often they update the data as an indirect method of indicating how long they consider the data to be valid. Table 3.4 shows that the average of updating time for historical records is 2 years. Each organisation has its own strategy to update its record to track the performance of its registered subcontractors.

Table 3.4. The updating frequency for historical records

Organisations	Keeping historical records	Updating frequency for Historical Records (Years)
O1	Yes	1
O2	Yes	2
O3	Yes	2-3 for medium size projects 5 years for Large Projects
O4	Yes	2-3 years
O5	Yes	2 years

Organisation 1 stated that annual tracking of performance of subcontractors is required to assure their suitability of working in their projects since data older than 1 year may not provide an accurate measure of current performance.

Organisation 2 stated that 2 years is a reasonable updating time of performance records as subcontractors are expected to work in more projects within two years which will provide more accurate ratings of their performance than would be obtained from a single project. However, this organisation does not consider the performance of subcontractors that worked in projects with any other organisation. The reason is that it considers that it has its unique standard which is different to other organisations.

Organisation 3 only accepts subcontractors that have worked in medium and large projects where small projects are not expected to reflect the performance that would be expected on one of its projects. It has two different updating times of its performance records. As shown in Table 3.4, more time is given for larger projects since the size of the projects that would limit the number of the projects that a subcontractor can work in and so more time is required to get a representative sample of projects. However, this organisation does not consider the performance of subcontractors that worked in projects with any other organisation as it also considers that it has its unique standard.

Organisation 4 considers 2 to 3 years as enough time that subcontractors can work in different number of projects so their performance would be reasonably assessed.

Organisation 5 updated its performance records within 2 years as the maximum time and it has similar reason of considering this time similar to organisation 4.

As a result, these answers state that organisations that have performance records are frequently updating their records within 1 to 5 years. This result would be compared to the results of the data collected by applying Markov and hidden Markov models in chapter 7 and 8 to determine the optimum updating time bases on the performance data that was collected.

Q5: ***How would these factors be measured?***

As shown in Table 3.5, the answers show that different techniques are used by each organisation to assess the performance data of subcontractors. Two of the companies have formal procedures where data is recorded at the end of each project that is then used for assessing subcontractors for the next project (O2 and O4). Three of the organisations ask for data about past projects from the subcontractor instead. The method of using this data varies from putting into a formal algorithm to ensure that successful subcontractors possess all of the minimum requirements to using it as input for a discussion between the project team. Moreover, one of the organisations uses a different weighting system based on the requirements of each project.

Table 3.5. The different measures used in assessing performance data

Organisations	Methods of measuring performance data
O1	- Actual performance. - Prequalification data. - Bid evaluation criteria.
O2	- According the previous projects and send them department of quality review to measure them.
O3	- Measuring these factors differ based on the requirements of a project, there is not fixed weighting system for all projects. The criteria is set and signed by the bid review team prior to the evaluation.
O4	- In Percentage scale by acquiring a minimum score for each performance factor.
O5	- Discussing the performance data with project team members.

The level of transparency between the organisations also differed. Some published their assessment criteria so that the subcontractors could know and understand it, like the governmental performance records in Saudi Arabia, while others kept the criteria and decision process confidential and hidden from the subcontractors.

Q6: ***How would you evaluate subcontractors for which you do not have any of this historical data?***

The answers provided for this question, as shown in Table 3.6, show that there are different strategies in assessing subcontractors in the absence of past performance data. Two organisations prefer to avoid working with such subcontractors until they build a record that could be used to track their performance (O3 and O5). The other three organisations that would accept such subcontractors have techniques to assess their expected performance. For example, organisation 4 uses 7 criteria to make a decision in assessing the performance of a new subcontractor. They will accept any subcontractor that can demonstrate and meet these criteria. This demonstration does not have to be via past projects.

Table 3.6. The techniques used in assessing new subcontractors

Organisations	Assessing subcontractors with no past information
O1	- Through a prequalification process, pre-RFP - Through weighted evaluation criteria, post-RFP * RFP = Request for price
O2	- Profile and site visit.
O3	- Does not prefer to work with subcontractors that have no records, especially in mega projects.
O4	- Execute plans of the project. - Available equipment and resources - Similar projects *if applicable*. - Quality assurance. - Safety Plan - Experience of project manager and Org Chart.
O5	- Not hiring subcontractors with no past information

3.3 Testing Data

One aim of this research is to validate the proposed methodology for analysing the performance of subcontractors and how it changes from project to project over time. While interviews are satisfactory for discovering the criteria by which contractors assess subcontractors, in order to see the variation in subcontractor performance, it is better to look at the actual historical records so that faulty human recollections are avoided.

Generally speaking, past performance of subcontractors is kept in the form of rating scales. For example a scale may range from “Weak” to “Excellent”. Obtaining this data will enable the examination of how stable subcontractors are in terms of these scales by determining the various probabilities for changes in each rating for subcontractors between projects.

The collection of data to answer the research questions was acquired from the historical records of the Deputy Ministry for Contractors Classification in the Ministry of Municipal and Rural Affairs. The records cover contractors and subcontractors involved in 29 types of work. The classification is based on financial and technical aspects as indicators of the ability of a contractor or a subcontractor to undertake or work on a project. The financial aspect measures capability and ability of a contractor or subcontractor to undertake a project based on three criteria: budget, profits and financial percentage. The technical aspect is to measure the capability in 5 criteria: organisation management level, equipment, previous projects, site visit and performance evaluation. The data collected was only in regard to the performance evaluation criteria.

The performance factors used in assessing subcontractors collected consist of six factors. Factor one is about project management (planning, organization and follow up). This factor is to check the proposed planning and the follow up in previous projects. The second factor is about work quality and compliance with the specifications to measure the quality of workers and whether they can achieve the required specifications in projects. The third factor is about compliance with the time scheduled as an indicator of the ability to execute the project based on the time frame scheduled and the proposed planning. The fourth factor is about project staff level (competence, experience, qualifications), which measures the experience and qualifications of mangers, engineers and workers of a subcontractor. The fifth factor is about the availability of the necessary equipment and systems and extent of their efficiency. This factor is considered as an indicator of a subcontractor ability to work in a project based on the required equipment. The sixth factor is about application of the security and safety procedures as they were proposed in tendering process. All of these factors were assessed using a scale ranging from "weak" to "excellent".

The data was collected in two phases to answer the research questions and to meet the aims about the future forecasting of performance of subcontractors.

3.2.3 *Data Set 1*

Data set 1 involved information about the performance of subcontractors in construction projects in Saudi Arabia from 2009 to 2012. It included data from 60 subcontractors that worked for 82 contractors in 197 different projects. This data set was not restricted in terms of types of work, date or contractor ranking.

The reasons for collecting this data were to examine the conditional probabilities for each performance factor to change from one performance rating to another or the probability of staying in the same rating. This data will be used to test the proposed model by applying both Markov chains and hidden Markov models. Some of the subcontractors exclusively worked for a particular contractor, other subcontractors worked with different contractors for each project, and some subcontractors worked with a mix of the first two cases. Sorting this data shows that there are more subcontractors that worked for the same contractor than those that worked for different or a mix of the same and different contractors, see Table 3.7.

Table 3.7. The three contracting patterns found for subcontractors in Saudi Arabia from 2009-2012.

Patterns	Subcontractors	Contractors	Projects
Subcontractors that exclusively worked for an individual contractor	38	32	123 (60.41%)
Subcontractors that always worked for different contractors for each project	11	34	41 (9.64%)
Subcontractors that worked for multiple contractors, sometimes for multiple projects	11	32	33 (23.86%)
Total	60	82	197

As stated, the use of this data to analyse the change in the performance of subcontractors will be tested using Markov chains and will be validated using Kullback's algorithm in Chapter 6. This will enable determination of the transition probabilities from one state to another. The aim of testing the validity of this data was to show that the data possesses the Markovity property. This means that the only data needed to predict the future performance of subcontractors is their current level of performance. If the data has this property then it fits a first order Markov chain.

This data will be also used to answer a research question about whether the importance of a factor would affect how critically the contractors would assess that item. This hypothesis was tested by identifying the correlations and p-values of the regression analysis between the importance rankings of each organisation for the performance factors involved in the data collected and the mean performance scores given to subcontractors for particular performance questions. The results of testing this data are presented in Chapter 4.

However, data set 1 has a problem that it had a small sample involving subcontractors that worked for multiple contractors. This would prevent using this sample unless more data was to be collected, i.e. data set 2

3.2.4 Data Set 2

One aim of collecting data set 2 was to enlarge the size of the sample that involved subcontractors that worked for multiple contractors. The new data set comprises 30 subcontractors that worked for the same 4 different contractors in 120 projects. The Markoivty of data collected is also validated using Kullback' algorithm in Chapter 7 for the same reasons stated for validating data set 1.

The other aim of collecting data set 2 was to test the consistency of the performance ratings given by different contractors. The level of agreement and the quality of the performance evaluations in data set 2 is examined by using the Cronbach's alpha in Chapter 5.

Data set 2 was also used to answer the research question that is whether the performance questions have any internal structure. This question is answered in Chapter 5 using factor analysis.

3.4 Conclusion

The aim of this chapter is explain the data collection method for this thesis. The collection of this data involved two goals. The first goal was to explore the importance of performance data and to identify the different methods and strategies used in assessing subcontractor performance by conducting interviews with organisations that have historical records of performance. The second goal was to answer the research question and validate the methodology of using hidden Markov models to analyse the performance of subcontractors and to determine the optimum updating time of their historical records. Two data sets were collected to meet the second goal. The extracted sample of data from data set 1 and the use of data set 2 make two types of subcontractors. Type 1 involved subcontractors worked for the same contractors and type 2 involved subcontractors worked multiple contractors. This chapter showed the relevance of the data collected, either for exploring or testing, to the aims and questions of the research.

Chapter 4 Identifying the Importance of Performance Measures and Effect on Ratings

4.1 Introduction:

The aim of this chapter is to answer the following research question:

What effect does contractor perception of the importance of the particular measures of performance affect the rating of subcontractors for that performance measure? In other words, do contractors rate subcontractors more harshly on the performance measures that they believe are most important?

The aim is derived from literature that indicates that clients, contractors and subcontractors have different perspectives in considering the most important performance factors to be used in assessing the performance (reference). However, this literature has focussed on comparing the factors considered important by different parties. It has not studied the relation between the importance of each performance measure and the actual scores.

Moreover, the validity of ratings received from different contractors or clients' needs to be addressed. This is because of the fact that different parties in the construction industry may consider different factors to be important. Thus the performance of subcontractors in the common historical record may not reflect their actual performance. Thus this chapter will investigate the hypothesis that organisations that consider a particular measure to be more important will give harsher (lower) scores for that measure compared to other measures.

The following section will describe the methodology of collecting the data to investigate this issue. The data was collected from Saudi Arabia as a case study where both interviews were conducted and historical performance data was collected.

4.2 Methodology:

The methodology used in collecting data contains two phases.

The first phase is involved conducting interviews with organisations that have historical records for the performance of subcontractors. This included 5 large organisations. These organisations were asked a total of 7 questions, 2 of which are relevant to this chapter, see chapter 3 to see all of the questions used in the interviews. The first relevant question asked the organisation to list the performance measures that it recorded for its own records. The second relevant question asked them to rank their importance of the government performance measures.

The second phase involved identifying the performance ratings in the data described in chapter 3 where one of the 5 organisations that had been interviewed had performed the performance measurement.

Two methods will be used to test the hypothesis that organisations that consider a particular measure to be more important will give harsher (lower) scores for that measure compared to other measures. Firstly the correlation between the ranking that each organisation gives to a measure and the mean score given to its subcontractors will be determined for each organisation. Secondly a regression analysis will be carried out to find the p-value based on the t-statistic of the slope coefficient in the regression analysis to see if the slope really is significantly different to zero.

The mean score equation is given as:

$$MS = \frac{\sum(f \times s)}{N} \qquad \text{Equation 4.1}$$

Where

s is the score for each factor that is given by each organisation ranging from 1 to 5

f is the frequency of responses to each for each factor.

N: is the total number of responses regarding each factor.

Calculations were performed in Microsoft Excel. Since the p-value given in the regression analysis in Microsoft Excel is based on a two tailed test (i.e. the slope is not zero) and the hypothesis in this chapter calls for a one tailed test (i.e. the slope is negative) the p-values from Microsoft Excel will be divided by 2.

4.3 Results

The answers of the question about the most important factors used by the 5 organisations to assess the performance of subcontractors working in their projects are shown in Table 4.1. As can be seen from the results of this question, four performance measures were recorded by four of the organisations: compliance with time scheduled, safety adherence, work quality; and the availability of equipment and resources. There were no performance factors that were recorded by all five companies. Another four performance measures were recorded by only one or two organisations each.

Table 4.1. The most important factors used for assessing the performance of subcontractors in Saudi Arabia

	Performance factors	Organisations					Frequency	Ranking
No.		**O1**	**O2**	**O3**	**O4**	**O5**		
1	Compliance to the project schedule.		●	●	●	●	4	2
2	Safety adherence	●	●	●	●		4	1
3	Availability of equipment and resources	●	●		●	●	4	1
4	Compliance to the scope of work				●		1	6
5	Work quality	●	●		●	●	4	1
6	Work procedures and methods				●		1	6
7	Environmental program	●					1	6
8	Compliance with the cost of the project		●	●			2	4

The rankings of the governmentally collected performance factors according to the 5 organisations are shown in Table 4.2 where scoring 5 is the highest and scoring 1 is the lowest. The compliance with time scheduled, safety, and work quality were most consistently ranked highly as the most important performance factors whereas the project staff level was considered as the least important performance factor.

Table 4.2. The ranking given to the performance factors used in the governmental records by the interviewed organisations in Saudi Arabia

Performance Factors	O1	O2	O3	O4	O5
Project Management (planning, organisation and follow up)	1	1	2	3	1
Work Quality and Compliance with the specifications	2	5	3	2	4
Compliance with time schedule	5	3	5	5	3
Project Staff Level	1	1	1	1	1
Availability of equipment	3	4	1	1	2
Application of safety procedures	4	2	4	4	5

The results shown in Table 4.1 and Table 4.2 finalises the first phase of the methodology of this study. The results of the second phase of conducting this study are all shown below.

Table 4.3 shows the number of projects and the number of subcontractors for each of the contractor organisations.

Table 4.3. The number of subcontractors worked for Saudi organisations in different number of projects

Organisation	Subcontractors	Projects
1	5	19
2	8	28
3	8	24
4	8	28
5	9	27

Based on the sample taken for each organisation, the results of their assessments for subcontractors that worked for them are shown in Tables 4.4 to 4.8.

Table 4.4. The mean scores and ranking for the performance assessments that were given for subcontractors worked for organisation 1

Organisations	**Performance Factors**	**Performance Ratings**					**Mean**
		1	**2**	**3**	**4**	**5**	
Organisation 1	Project Management (planning, organisation and follow up)	0	2	5	9	3	3.684
	Work Quality and Compliance with the specifications	0	1	4	8	6	4.000
	Compliance with time schedule	1	0	4	12	2	3.737
	Project Staff Level	0	1	5	11	2	3.737
	Availability of equipment	0	2	4	11	2	3.684
	Application of safety procedures	1	1	8	8	1	3.368

Table 4.5: The mean scores and ranking for the performance assessments that were given for subcontractors worked for organisation 2

Organisations	**Performance Factors**	**Performance Ratings**					**Mean**
		1	**2**	**3**	**4**	**5**	
Organisation 2	Project Management (planning, organisation and follow up)	0	4	13	9	2	3.321
	Work Quality and Compliance with the specifications	0	2	11	13	2	1.818
	Compliance with time schedule	0	7	9	10	2	2.286
	Project Staff Level	0	1	13	8	4	2.655
	Availability of equipment	0	2	10	11	5	2.212
	Application of safety procedures	0	9	8	9	2	2.329

Table 4.6: The mean scores and ranking for the performance assessments that were given for subcontractors worked for organisation 3

Organisations	**Performance Factors**	**Performance Ratings**					**Mean**
		1	**2**	**3**	**4**	**5**	
Organisation 3	Project Management (planning, organisation and follow up)	0	0	5	6	13	4.333
	Work Quality and Compliance with the specifications	0	1	1	6	16	4.542
	Compliance with time schedule	0	0	8	2	14	4.250
	Project Staff Level	0	0	5	8	11	4.250
	Availability of equipment	0	0	2	12	10	4.333
	Application of safety procedures	0	2	3	9	10	4.125

Table 4.7: The mean scores and ranking for the performance assessments that were given for subcontractors worked for organisation 4

Organisations	**Performance Factors**	**Performance Ratings**					**Mean**
		1	**2**	**3**	**4**	**5**	
Organisation 4	Project Management (planning, organisation and follow up)	0	2	10	12	4	3.643
	Work Quality and Compliance with the specifications	0	0	6	9	13	4.250
	Compliance with time schedule	0	1	17	4	6	3.536
	Project Staff Level	0	2	4	18	4	3.857
	Availability of equipment	0	1	1	22	4	4.036
	Application of safety procedures	0	5	11	10	2	3.321

Table 4.8: The mean scores and ranking for the performance assessments that were given for subcontractors worked for organisation 5

Organisations	Performance Factors	Performance Ratings					Mean
		1	2	3	4	5	
Organisation 5	Project Management (planning, organisation and follow up)	0	0	4	8	15	4.407
	Work Quality and Compliance with the specifications	0	3	5	8	11	4.000
	Compliance with time schedule	0	3	4	7	13	4.111
	Project Staff Level	0	0	4	7	16	4.444
	Availability of equipment and resources	0	0	4	11	12	4.296
	Application of safety procedures	0	3	10	7	7	3.667

The results of the actual performance assessments of the 5 organisation that were given to subcontractors worked for them are matched with the ranking of the mean score for each performance question in Table 4.9.

Table 4.9. Results of the means score of subcontractro performance compared to the ranking given by organisations

Performance Factors	O1		O2		O3		O4		O5	
	M	R	M	R	M	R	M	R	M	R
Project Management (planning, organisation and follow up)	3.684	1	3.321	1	4.333	2	3.643	3	4.407	1
Work Quality and Compliance with the specifications	4.000	2	1.818	5	4.542	3	4.25	2	4.000	4
Compliance with time schedule	3.737	5	2.286	3	4.250	5	3.536	5	4.111	3
Project Staff Level	3.737	1	2.655	1	4.250	1	3.857	1	4.444	1
Availability of equipment	3.684	3	2.212	4	4.333	0.5	4.036	1	4.296	2
Application of safety procedures	3.368	4	2.329	2	4.125	4	3.321	4	3.667	5

These ranking of the actual assessments of performance will be then compared with the scores that were given by these organisations to the performance factors in the governmental record. The comparisons between these rankings will firstly be studied using the correlations between them for each organisation. They will then also be tested using regression analysis to determine the p-value for the slope of the regression line for each organisation.

The results of the correlations and P-values between the importance assigned by these 5 organisations to the performance factors in the governmental record and the performance ratings they gave to the subcontractors that worked for them are shown in Table 4.10. The results show that the correlation results for all organisations are negative. The probability that all of these would be negative if the null hypothesis is true is 0.03125. Therefore the null hypothesis is rejected and it is concluded that contractors rate subcontractors more harshly for performance factors that they consider more important. Also, the individual P-values are significant for 3 of the 5 organisations.

Table 4.10. Correlation between performance factor importance and bestowed ratings

Organisations	Correlation	One-tailed P-values
O1	-0.3623	0.2402
O2	-0.8558	0.0149
O3	-0.2322	0.3289
O4	-0.7527	0.0420
O5	-0.9827	0.0002

4.4 Discussion:

The results of the study conducted in this chapter show that contractors rate subcontractors more harshly for performance factors that they consider more important. That harshness is derived from the fact that the most important performance factors considered by these organisations will be used as the indicators of the performance of subcontractors that work for these organisations and so more attention will be paid to the factors considered important by each organisation. This result leads to the question of whether reconsideration is required regarding the performance assessments that are given to the governmental performance record from organisations that assign different levels of importance to different performance factors. This requires an analysis of the consistency of the performance ratings of different contractors for the same subcontractors to make a fair comparison. This analysis will be performed in the next chapter.

Since correlation does not imply causation an alternative view of these results is that instead of contractors rating more harshly on the factors that are considered important it could be that the indicated importance levels are actually assessing how important it is that particular performance factors be improved. The idea here is that contractors have a greater focus on improving the performance factors where they receive low performance than on the factors where they already receive high performance. Thus the low ratings cause the assignment of high importance rather than the other way around.

A deeper examination of the data in Table 4.2 shows that the three factors that were generally assigned the highest importance (time, safety and quality) are output factors. These things are received by the contractors. On the other hand the three factors that were generally assigned the lowest importance (staff level, project management, and equipment availability) are input factors. These are the things that enable the subcontractors to do their work. Thus more importance is attached to factors regarding the actual work of the subcontractors than to the factors that enable them to do work.

In particular it can be seen in the results is that all of the organisations assigned minimal importance to project staff level. Thus it may be that there is not much point in keeping track of this performance measure.

This leaves only 5 performance measures that are actually worth tracking. Comparing this to the results of the study of key performance indicators (KPIs) in Saudi Arabia by Ali et al. (2013) where 10 important performance indicators were identified show that less performance indicators are actually used when making selection decisions. This means that although the importance of more KPIs was acknowledged, some parties still use less performance factors.

4.6.Conclusion

The aim of this chapter is to test the hypothesis that the assessment of subcontractor performance was affected by the importance placed on each performance factor. The testing of this hypothesis was conducted by comparing the performance assessments given by organisations with the importance assigned to each question by those organisations. Data used came from a case study of Saudi Arabian contractors where 5 different organisations were interviewed to determine how important they considered each of the performance factors used in governmental records, and actual performance data were collected from the governmental record for evaluations of 38 subcontractors over a total of 126 projects.

The results of this chapter show that there are some different organisations consider different performance factors to be the most important. In fact, there were differences between the lists of performance factors collected by each organisation, although there were many similarities.

Generally the output factors of time, safety and quality were considered more important than the input factors of staff level, equipment availability and project management.

The results of comparing the ranking given by these organisations to the most important performance factors used in the common governmental record and their actual assessments for subcontractors worked for them show that there is a strong trend. This trend is represented as those subcontractors being more harshly assessed on the performance factors considered to be most important by the assessing contractor, whereas they were given higher ratings to the less important factors.

While it is plausible that the importance of the factors caused the low ratings the data could also be the result of low ratings causing particular factors to be considered important. The next chapter will investigate the consistency of ratings given by different contractors to a uniform set of subcontractors to see whether this effect of contractors rating important factors more harshly affects the overall validity of subcontractor performance rating.

Chapter 5 Testing the Consistency of the Performance Ratings of Subcontractors from Different Contractors and Identification of Internal Structure

5.1 Introduction

Data was collected from Saudi Arabia regarding the performance of subcontractors in two sets. The first set of data collected contained three categories of subcontractors in terms of the contractors that they had worked with. The first category included subcontractors that only worked for a single contractor. The second category included subcontractors that worked for multiple contractors, but not for the same contractor twice. The third category included those that worked for multiple contractors, including some ow which they had worked for multiple times.

A problem that was considered was whether ratings given by different contractors would be consistent with each other. Initially it was planned to test the consistency of the subgroup of data that involved multiple contractors. However, the sample size of subcontractors that worked for multiple contractors was too small to be tested. Therefore, another set of data was collected. This second set of data was specifically selected so that it included a group of subcontractors that had all worked for the same contractors at some stage. Thus it included exactly one project for each combination of subcontractor and contractor. The aim of this chapter is to test the internal consistency of the different performance ratings that were given by different contractors. This consistency is required to ensure that the results of the remainder of the thesis are valid.

The other aim of this chapter is to determine if there are any relationships between the different questions.

5.2 Methodology

The data collected for the performance of subcontractors that worked for multiple contractors included 30 subcontractors that worked for the same 4 different contractors giving a total number of 120 projects. The aim of collecting data for subcontractors that worked for the same multiple contractors will help in testing the consistency of the different performance ratings. The methodology that is used to meet the aims of this chapter will be as follows:

1- The consistency of the different ratings of performance will be tested using Cronbach's alpha. According to Tavakol and Dennick (2011), Cronbach's alpha was proposed by Lee Cronbach in 1951 for measuring the internal consistency of a test or a scale. The results of Cronbach's alpha range between 0 and 1. A result of at least 0.7 shows an acceptable reliability of the data tested. The meaning of the internal consistency results using Cronbach's alpha that was interpreted by George and Mallery (2003) are shown in Table 5.1 The internal consistency is considered poor when the result of alpha is less than 0.6 and not acceptable if it is less than 0.5.

Table 5.1: The equivalent internal consistency for Cronbach's alpha results (George and Mallery, 2003)

Cronbach's Alpha	Internal Consistency
$\alpha \geq 0.9$	Excellent
$0.9 > \alpha \geq 0.8$	Good
$0.8 > \alpha \geq 0.7$	Acceptable
$0.7 > \alpha \geq 0.6$	Questionable
$0.6 > \alpha \geq 0.5$	Poor
$0.5 > \alpha$	Unacceptable

The equation used for Cronbach's alpha is:

$$\alpha = \frac{K}{K-1}\left(1 - \frac{\sum_{i=1}^{K} \sigma_{Y_i}^2}{\sigma_X^2}\right)$$ Equation 5.1

where:

K is the number of components or items.

σ_X^2 is the variance of the observed total test scores.

$\sigma_{Y_i}^2$ is the variance of the component *i* of the current tested sample

Testing the consistency using Cronbach's alpha will be conducted using SPSS.

2- Determination of relationships between the performance measures will be carried out using Factor Analysis. According to Chan et al. (2004) and Aksorn and Hadikusumo (2008), Norusis (1993) defined factor analysis as a statistical technique that is used to identify a small number of factors can be used to represent the relationships among sets of many interrelated variables. The basic steps that factor analysis involves are summarised as:

1- Identifying the measured variables. In this case this refers to the six performance factors.

2- Calculating the correlation of the six performance factors used in the selected case of Saudi Arabia.

3- Extracting and rotating each component.

4- Identifying the number of components that represent regularities in the performance factors.

This analysis was conducted using SPSS.

5.3 Results

Consistency testing of the performance ratings for Cronbach alpha was carried out without excluding any of the data. The results of the consistency test of the performance ratings that were given by 4 different contractors for the same 30 subcontractors are shown in Table 5.2.

Table 5.2. Results of the consistency for the different performance ratings given for subcontractors by different contractors

Contractors	Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
1	0.894	0.909	6
2	0.914	0.921	6
3	0.938	0.940	6
4	0.895	0.903	6

The results of Cronbach’s alpha for all contractors fall into the good to excellent range. The results are classified as good for the performance ratings from contractors 1 and 4 whereas they are classified excellent from contractors 2 and 3.

For the factor analysis two components were found to have an eigenvalue of greater than 1.0. The scree plot for this is shown in Figure 5.1.

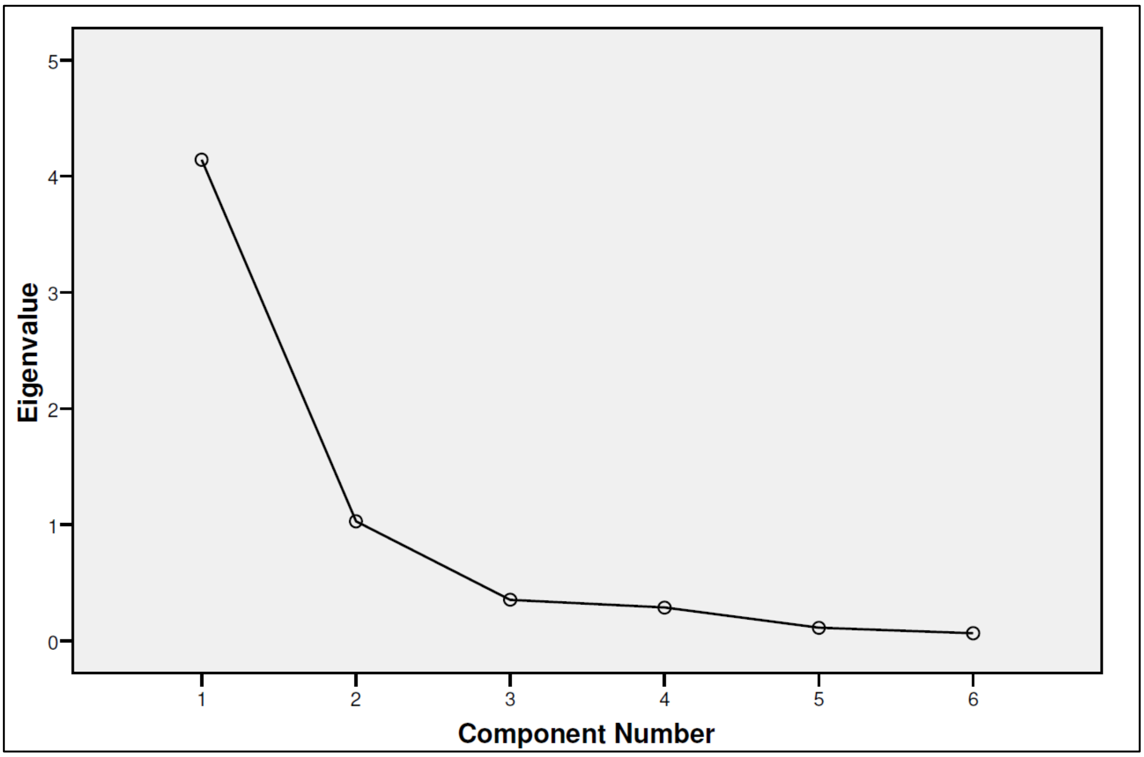

Figure 5.1. The scree plot of the Eigenvalue of 6 performance questions

Figure 5.1 shows a distinct break of slope at the second component. According to Chan et al. (2004) the gradual trailing off in plot is called the scree, which means that this scree resembles the debris that are formed on the foot of the mountain. As a result, the results of the total variances shown in the figure prove that the performance questions used in assessing the performance of subcontractors are represented by two components.

The results of the loading factors when rotated using the method of Varimax with Kaiser Normalization are shown in Table 5.3. The results show that questions 1, 3 and 5 represent one group; and questions 2, 4 and 6 represent another group. There is a large difference between the factor loadings for each question with its own group and the opposite group.

Table 5.3. The results of identifying the number of groups that the performance questions represent in 6 performance questions used from the data collected

Performance questions	**Component**	
	1	2
Q3	0.890	0.337
Q5	0.877	0.340
Q1	0.874	0.219
Q4	0.257	0.927
Q6	0.273	0.917
Q2	0.359	0.777

Two performance groups were extracted that accounted for 86% of the variance of the different performance ratings where they represent 69% and 17%, see Table 5.4. All factor loadings were greater than 0.7 while 5 of the 6 were greater than 0.85. Moreover, these results mean that all of the performance questions are consistent.

The formation of two groups of performance measures from the 6 performance questions can be clearly seen in Figure 5.2.

Table 5.4. Factor Structure of Principal Factors Extraction for performance questions

Performance Questions	Factor Loading	Percentage of variance explained	Cumulative percentage of variance explained
Group 1		69.044	69.044
Compliance with time schedule	0.890		
Availability of equipment	0.877		
Project Management (planning, organisation and follow up)	0.874		
Group 2		17.181	86.225
Project Staff Level	0.927		
Application of safety procedures	0.917		
Work Quality and Compliance with the specifications	0.777		

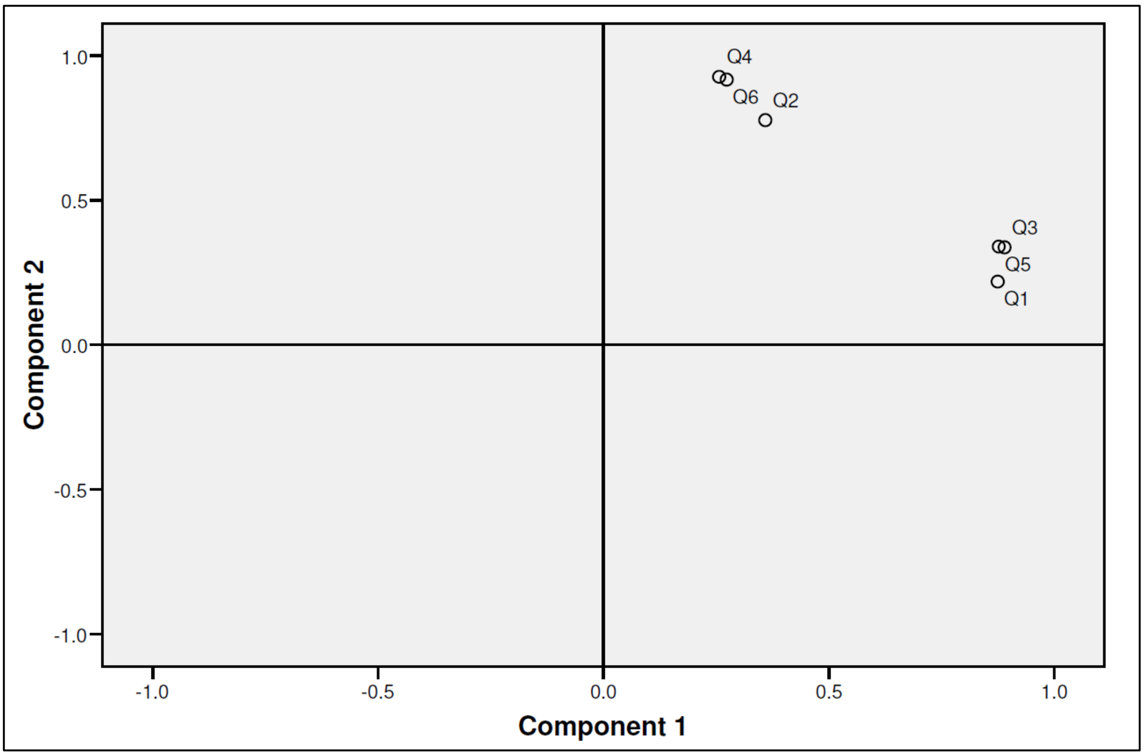

Figure 5.2. Component Plot of the performance questions showing the formation of two performance measures.

5.4 Discussion

The results of the consistency of the different performance ratings for subcontractors show that they are quite consistent and thus using rating assigned by different contractors presents no problem. This seems to be in opposition to the results from the previous chapter where individual contractors were harsher when they assessed performance factors that they considered important compared with other factors. This is consistent with the alternative hypothesis presented in the previous chapter that contractors simply consider the factors where they receive the lowest performance as being the most important because they have the most room for improvement. As a result, the existence of the relation between harshness and importance that was found in the previous chapter does not create a problem with using any of the data.

The result of factor analysis is that there are two groups of performance factors among the questions. An important part of factor analysis is identifying what these factors are.

Examination of the questions belonging to the two groups in Table 5.4 shows that the first three questions tend to be related to the management skill of the subcontractor. Thus the group comprising these three questions has been labelled management performance questions. In contrast the three questions in the second group tend to be about the technical skill of the workers (safety and quality) or the staffing level. Thus the group comprising these three questions has been labelled technical performance questions.

The results of the factor analysis will be used in the two following chapters about Markov chains and hidden Markov Models. The results of these two chapters will be analysed in terms of these two measures and their underlying performance questions. The analysis of the data based on these two results will help in identifying how these performance measures change over time.

5.5 Conclusion

In conclusion, this chapter aims to examine the consistency of the different performance ratings given for subcontractors by different contractors. The results of testing the consistency showed that different performance ratings are all highly consistent. Thus it appears that these performance ratings really do reflect the actual performance of subcontractors.

The other aim of this chapter is to determine if the performance questions used in assessing the performance of subcontractors based on the case study conducted in Saudi Arabia have internal structure. The results show that the 6 performance questions can be categorized into two groups. The first group was labelled management performance questions and includes project management, compliance with time scheduled and availability of equipment. The other group was labelled technical performance questions and includes work quality, project staff level and safety procedures. These two measures will be used in analysing the change in the performance of subcontractors using Markov chain and hidden Markov models in the following chapters. This will assist with exploring for different trends related to the different groups.

Chapter 6 Validating the Markovity of the Performance Data

6.1 Introduction

The overall aim of this thesis is to show that hidden Markov models are useful for identifying the performance level of individual subcontractors. To do this it is necessary to show that the data possesses the Markovity property. This means that the only data needed to predict the future performance of subcontractors is their current level of performance. If the data has this property then it fits a first order Markov chain.

Chapter 3 showed that very few subcontractors, (11 out of 60), in the first data set that was collected had worked for multiple contractors. Most of subcontractors, (38 out of 60), had worked for the same contractors on all of their projects, although it was generally a different contractor for each subcontractor. The small sample size of subcontractors that had worked for multiple contractors prevented comparisons of the performance between these two groups. However, it is important to make comparisons because working consistently for a single contractor may be the result of partnerships.

Thus, it was decided to collect a second set of data focussing on subcontractors that had worked for multiple contractors. This would increase the sample size for this group to enable comparisons with the group that worked for single contractors. It also has the added benefit of enabling the testing of the consistency of ratings given by different contractors (discussed in Chapter 5).

The aim of this chapter is to use the extra data collected in the second data set to compare the results in each performance question used in assessing the performance of subcontractors between subcontractors that worked with multiple contractors or single contractors. This comparison is expected to provide insight into the nature of the relationship between the sub-contractors and the contractors in the two situations of multiple vs single contractors. In studying these two groups, two scenarios were considered that might lead to different outcomes for the two different groups.

The first scenario is that subcontractors that worked for the same contractor for more than two projects are expected to have less variability in their performance than those with multiple contractors. Thus, the performance of these subcontractors would be expected to be more stable than subcontractors who always worked for different contractors. Interviews with contractors in Saudi Arabia revealed several reasons that might cause this. Firstly this might be the result of a partnership between a subcontractor and a contractor. Secondly, contractors are often not willing to change subcontractors because this incurs the cost of executing a process where the new subcontractors prove their capability to satisfy the requirements of the contractor's projects. Thirdly, even after vetting new subcontractors it is not guaranteed that they will not be worse than the existing subcontractors and hence might damage the reputation of the contractor. These answers were given in the context of asking contractors that did not keep performance records of their subcontracts why they did not do this.

The other scenario to be taken into account is that subcontractors that worked for multiple contractors might be more qualified and competitive. (The small sample size of the first data set appeared to indicate this).

Therefore, to study the differences between the performance ratings for both groups, Markov chains will be implemented on both data sets separately. But this implementation, that is going to be tested in chapter 7, is restricted to one hypothesis that needs to be tested. That hypothesis is about testing whether each data group is a first order Markov chains or not, and that is the main goal of this chapter.

To justify the reasons of testing whether each group is a first order or not is because of the fact that first order Markov chains is expected to give more accurate and reliable results than second or any other higher orders. Those results may be used to select a subcontractor or forecast its future performance.

So, testing whether each group is a first order Markov chains or not is now justified. In methodology section of this chapter, more details about how to test this hypothesis, what expected results are and what other considered scenarios are; will be all explained.

This chapter will examine the validity of testing the collected data using Markov chains. This means that the two data sets should satisfy the Markovity that requires that each subcontractor has at least three data points. The Markovity means that the availability of the past does not help with predicting the future once the present is known. Therefore, testing the Markovity requires measurements of performance from at least three different times so that there can be information of performance involved past, present and future to compare. The small set of data that contained subcontractors that worked for more than one subcontractor, but for multiple times for any of them was examined. This examination was for situations where they had either worked for the same contractor at least three times. Those identified cases in data set 1 were added to the same contractor group whereas the other identified cases where a subcontractor had worked for three different contractors were added to the different contractors group.

If the general nature of the best fit models for these two groups is similar then it will be possible to combine both of them with the mixed groups to make a larger data set for further analysis.

6.2 Methodology:

The hypothesis that needs to be tested is:

H1: *The probability distribution for the performance of each group is a first order Markov chain.*

The process of testing this hypothesis is shown in Figure 6.1.

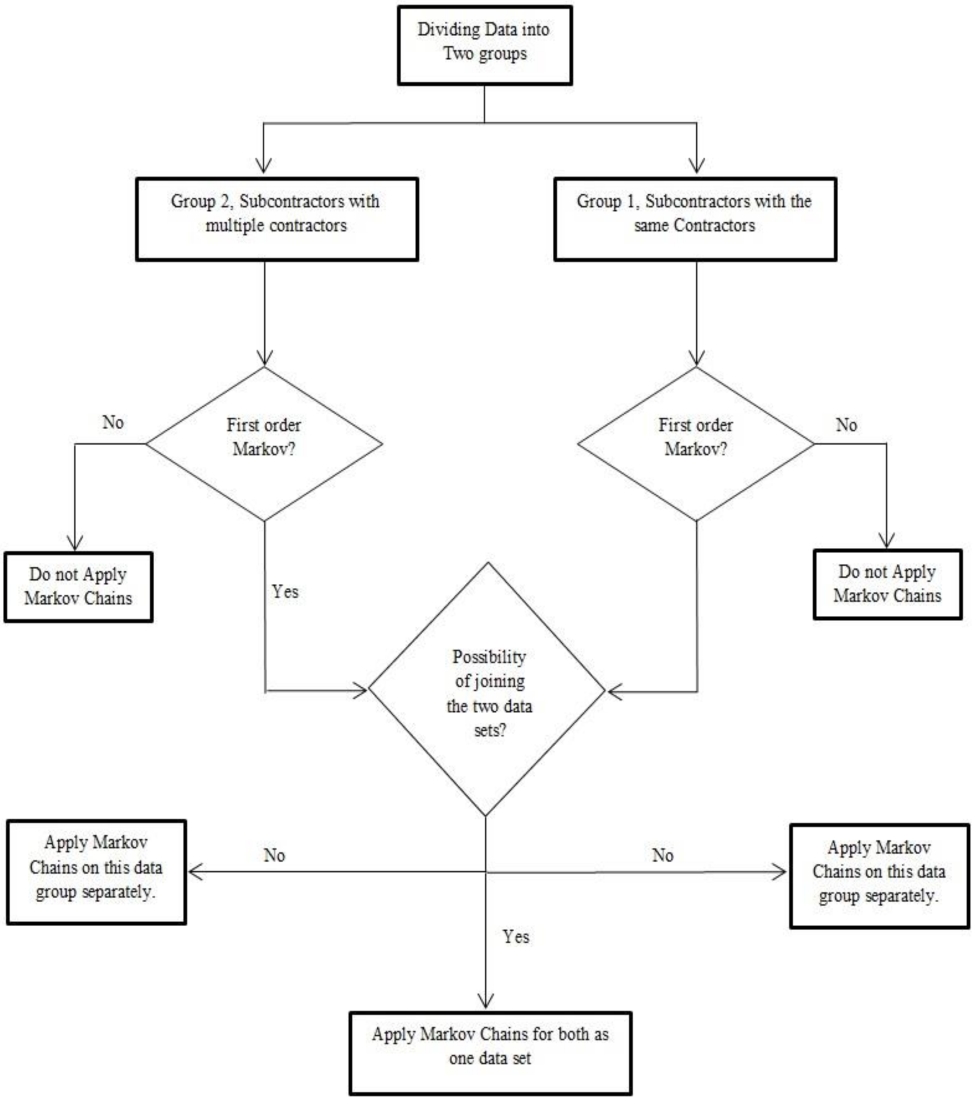

Figure 6.1. The process of validating the Markovity of the two performance data sets

The general methodology for testing this hypothesis will follow (Kullback et al., 1962). This methodology will be applied separately to the six performance questions in the data collected about the performance of subcontractors from the case study of Saudi Arabia. For more details about the performance questions refer to chapter 3.

The steps involved are the following:

1- It is assumed that the probability distribution of each performance question in each group is not a third or higher order Markov chain.

2- The probability distribution for each performance question in each group has to be determined to obtain the transition matrices for the Markov chains. Both first and second order transition matrices need to be obtained.

3- The values of the probabilities are then substituted into the equation for the Markovity component given in Table 6.1.

Table 6.1. The three components of testing if a probability distribution is a first order Markov Chain (Table 8.6 from (Kullback et al., 1962).

Component due to	Information	D. F.
Two-way independence	$2 \sum_{i=1}^{r} \sum_{j=1}^{r} f_{ij.} \ln \frac{n f_{ij.}}{f_{i..} f_{.j.}}$	$(r-1)^2$
Markovity (conditional independence)	$2 \sum_{i=1}^{r} \sum_{j=1}^{r} \sum_{k=1}^{r} f_{ijk} \ln \frac{f_{ijk}}{\frac{f_{ij.} f_{.jk}}{f_{.j.}}}$	$r(r-1)^2$
Two-way by one-way independence	$2 \sum_{i=1}^{r} \sum_{j=1}^{r} \sum_{k=1}^{r} f_{ijk} \ln \frac{n f_{ijk}}{f_{i..} f_{.jk}}$	$(r-1)(r^2-1)$

All entries in all of the three equations in Table 6.1 are additive. The equations in the table refer to:

- The two-way independency equation tests whether $P(E_j|E_i) = P(E_j)$, which would imply that the data does not follow a Markov chain at all.
- The Markovity equation tests whether the Markov chain is a first order Markov chain. A significant result would lead to concluding that the Markov chain is second order.
- The two-way by one-way independence equation is testing whether the second step transition matrices are homogenous with respect to the initial values.

$f_{i\,j\,k}$ is the probability that a subcontractor that starts with a rating of *i*, receives a rating of *j* in the next evaluation and a rating of *k* in the third evaluation.

Dots in the subscripts in the equation indicate that the value should be summed for that subscript. Thus $f_{i\,.\,.}$ is simply the probability that a subcontractor receives a rating of *i* in the initial evaluation.

In the following section there will be more explanation of how to convert the three equations to simple additions to be then used to find out the results for testing the proposed hypothesis.

6.2.1 Explaining the Logarithms of Kullback's Equations:

In this section, the logarithms of the three equations for the two- way independence, Markovity and the two-way by one-way independence will be explained in terms of how to convert them to be in simple additions.

1- The equation used to find the two-way independence is:

$$2\sum_{i=1}^{r}\sum_{j=1}^{r} f_{ij.} \ln \frac{n\, f_{ij.}}{f_{i..} f_{.j.}} \qquad \text{Equation 6.1}$$

Converting this equation to simple additions will be:

$$\sum_{i=1}^{r}\sum_{j=1}^{r} 2f_{ij.} \ln f_{ij.} + \sum_{i=1}^{r} 2f_{i..} \ln f_{i..} - \sum_{j=1}^{r}\sum_{k=1}^{r} 2f_{.jk} \ln f_{.jk} - \sum_{j=1}^{r} 2f_{.j.} \ln f_{.i.} \qquad \text{Equation 6.2}$$

2- The equation used to find the Markovity is:

$$2\sum_{i=1}^{r}\sum_{j=1}^{r}\sum_{k=1}^{r} f_{ijk} \ln \frac{\frac{f_{ijk}}{f_{ij.}f_{.j}}}{f_{.j.}} \qquad \text{Equation 6.3}$$

Converting this equation to simple additions will be:

$$\sum_{i=1}^{r}\sum_{j=1}^{r}\sum_{k=1}^{r} 2f_{ijk} \ln f_{ijk} + \sum_{j=1}^{r} 2f_{.j.} \ln f_{.i.} - \sum_{i=1}^{r}\sum_{j=1}^{r} 2f_{ij.} \ln f_{ij.} - \sum_{j=1}^{r}\sum_{k=1}^{r} 2f_{.jk} \ln f_{.jk} \qquad \text{Equation 6.4}$$

3- The equation used to find the two-way by one-way independence is:

$$2\sum_{i=1}^{r}\sum_{j=1}^{r}\sum_{k=1}^{r} f_{ijk} \ln \frac{nf_{ijk}}{f_{i..}f_{.jk}} \qquad \text{Equation 6.5}$$

Converting this equation to simple additions will be:

$$\sum_{i=1}^{r}\sum_{j=1}^{r}\sum_{k=1}^{r} 2f_{ijk} \ln f_{ijk} + \sum_{i=1}^{r} 2f_{i..} \ln f_{i..} - \sum_{j=1}^{r}\sum_{k=1}^{r} 2f_{.jk} \ln f_{.jk} - \sum_{j=1}^{r}\sum_{k=1}^{r} 2f_{.jk} \ln f_{.jk} \qquad \text{Equation 6.6}$$

After converting all of the three equations to simple additions, there are six main components need to be explained on how to substitute them with the values from the main given table of first order Markov chain. Firstly, the main table of the first order Markov chain consists of three states as shown in Table 6.2.

Table 6.2. An example of how the testing the Markovity could be obtained from a set of data

Question		Second State	Third State			
			A	B	C	TOTAL =
First State	A	A	A01	B01	C01	A01+B01+C01
		B	A02	B02	C02	A02+B02+C02
		C	A03	B03	C03	A03+B03+C03
	B	A	A11	B11	C11	A11+B11+C11
		B	A12	B12	C13	A12+B12+C12
		C	A13	B13	C13	A13+B13+C13
	C	A	A21	B21	C21	A21+B21+C21
		B	A22	B22	C23	A22+B22+C22
		C	A23	B23	C23	A23+B23+C23
		TOTAL				

From this table, the $f_{ij.}$ and $f_{.jk}$ tables will be established as follows:

Table 6.3. An illustrative example of how the values of *fij*.and *f.jk* can be obtained from the each data set of performance

Question		Second State			
		A	B	C	TOTAL
First State	A	A01+B01+C01	A02+B02+C02	A03+B03+C03	
	B	A11+B11+C11	A12+B12+C12	A13+B13+C13	
	C	A21+B21+C21	A22+B22+C22	A23+B23+C23	
	TOTAL				

Question		Third State			
		A	B	C	TOTAL
Second State	A	A01+A11+A21	B01+B11+B21	C01+C11+C21	
	B	A02+A12+A22	B02+B12+B22	C02+C12+C22	
	C	A03+A13+A23	B03+B13+B23	C03+C13+C23	
	TOTAL				

6.3 Results:

Table 6.4 shows the second order transition tables for the six performance measurements for the subcontractors that only worked for one contractor, while

Table 6.5 gives the same data for those that worked for multiple contractors. Note that the rating of performance is simplified here where 1 indicates the lowest performance ratings received and 5 indicting the highest performance rating. A major issue for analysing these tables as Markov chains is that very few subcontractors were given ratings of 1 or 2 in any of the questions. The small sample size makes the calculated probabilities very unreliable. In addition none of the subcontractors received a rating of 1 for their first evaluation in any set of three consecutive ratings and none received a rating of 1 or 2 in their third evaluation in any set of three consecutive ratings. This leads to a lot of zero counts in the tables. This leads to division by zero errors in calculating certain probabilities. The other problem is that the empty cells should not count toward the number of degrees of freedom used in the statistical tests and the empty cells make up a large proportion of the total number of cells.

Table 6.4. Frequencies of occurrence for sets of three consecutive ratings for subcontractors that consistently worked for the same contractor

Q1		Second State	Third State					
			1	2	3	4	5	Total
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	1	3	0	0	0	0	0	0
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
								0
		1	0	0	0	0	0	0
		2	0	0	0	1	0	1
	2	3	0	0	1	0	0	1
		4	0	1	0	0	0	1
		5	0	0	1	1	0	2
								0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
First State	3	3	0	0	6	1	0	7
		4	0	0	0	1	0	1
		5	0	0	0	0	0	0
								0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	4	3	0	0	0	1	0	1
		4	0	0	1	9	2	12
		5	0	0	0	2	2	4
								0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	5	3	0	0	1	0	0	1
		4	0	0	0	0	1	1
		5	0	0	4	0	10	14
		Total	0	1	14	16	15	46

Q2		Second State	Third State					
			1	2	3	4	5	Total
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	1	3	0	0	0	0	0	0
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
								0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	2	3	0	0	1	0	0	1
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
								0
		1	0	0	0	0	0	0
		2	0	0	1	0	0	1
First State	3	3	0	0	1	3	0	4
		4	0	1	1	0	0	2
		5	0	3	0	1	0	4
								0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	4	3	0	0	1	0	0	1
		4	0	0	2	8	1	11
		5	0	0	0	0	3	3
								0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	5	3	0	0	0	0	0	0
		4	0	0	0	0	2	2
		5	0	0	0	1	16	17
		Total	0	4	7	13	22	46

Q3		Second State	Third State					
			1	2	3	4	5	Total
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	1	3	0	0	0	0	0	0
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
								0
		1	0	0	0	0	0	0
		2	0	0	0	1	0	1
	2	3	0	0	0	0	0	0
		4	0	1	1	0	1	3
		5	0	0	0	0	0	0
								0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
First State	3	3	0	0	8	0	0	8
		4	0	0	0	2	0	2
		5	0	3	0	0	0	3
								0
		1	0	0	0	0	0	0
		2	0	0	0	1	0	1
	4	3	0	0	1	0	0	1
		4	0	0	0	8	1	9
		5	0	0	0	0	4	4
								0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	5	3	0	0	0	0	0	0
		4	0	0	0	1	1	2
		5	0	0	0	1	11	12
		Total	0	4	10	14	18	46

Q4		Second State	Third State					
			1	2	3	4	5	Total
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	1	3	0	0	0	0	0	0
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
								0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	2	3	0	0	0	0	0	0
		4	0	0	0	0	0	0
		5	0	0	0	0	1	1
								0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
First State	3	3	0	0	3	2	0	5
		4	0	2	0	2	0	4
		5	0	0	0	0	1	1
								0
		1	0	0	0	0	0	0
		2	0	0	1	0	0	1
	4	3	0	0	0	1	0	1
		4	0	0	1	10	2	13
		5	0	0	0	2	2	4
								0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	5	3	0	0	0	0	0	0
		4	0	0	0	2	3	5
		5	0	0	3	1	7	11
		Total	0	2	8	20	16	46

Q5		Second State	Third State					
			1	2	3	4	5	Total
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	1	3	0	0	0	0	0	0
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
								0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	2	3	0	0	0	1	0	1
		4	0	0	0	0	0	0
		5	0	0	0	1	0	1
								0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
First State	3	3	0	0	3	1	0	4
		4	0	1	1	0	0	2
		5	0	0	0	0	0	0
								0
		1	0	0	0	0	0	0
		2	0	0	1	0	0	1
	4	3	0	0	1	1	0	2
		4	0	0	1	11	3	15
		5	0	0	1	1	2	4
								0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	5	3	0	0	0	0	0	0
		4	0	0	0	3	1	4
		5	0	0	3	1	8	12
		Total	0	1	11	20	14	46

Q6		Second State	Third State					
			1	2	3	4	5	Total
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	1	3	0	0	0	0	0	0
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
								0
		1	0	0	0	0	0	0
		2	0	0	1	0	0	1
	2	3	0	1	1	0	0	2
		4	0	1	0	0	0	1
		5	0	1	0	0	1	2
								0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
First State	3	3	0	1	8	0	0	9
		4	0	0	2	1	0	3
		5	0	3	0	1	0	4
								0
		1	0	0	0	0	0	0
		2	0	0	1	1	0	2
	4	3	0	0	0	1	1	2
		4	0	0	0	5	1	6
		5	0	0	0	0	0	0
								0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	5	3	0	0	0	0	0	0
		4	0	0	1	1	0	2
		5	0	0	0	2	10	12
		Total	0	7	14	12	13	46

Table 6.5. Frequencies of occurrence for sets of three consecutive ratings for subcontractors that worked

for multiple contractors

Q1		Second State	Third State					
			1	2	3	4	5	Total
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	1	3	0	0	0	0	0	0
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
		1	0	0	0	0	0	0
		2	0	0	1	0	0	1
	2	3	0	0	4	0	0	4
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
		1	0	0	0	0	0	0
		2	0	0	1	1	0	2
First State	3	3	0	0	7	4	0	11
		4	0	0	3	4	1	8
		5	0	0	0	0	0	0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	4	3	0	1	2	2	0	5
		4	0	0	0	12	2	14
		5	0	0	0	4	6	10
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	5	3	0	0	0	0	0	0
		4	0	0	0	2	2	4
		5	0	0	0	5	5	10
		Total	0	1	18	34	16	69

Q2		Second State	Third State					
			1	2	3	4	5	Total
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	1	3	0	0	0	0	0	0
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
		1	0	0	0	0	0	0
		2	0	0	1	0	0	1
	2	3	0	0	1	1	0	2
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
First State	3	3	0	0	0	6	0	6
		4	0	1	6	6	2	15
		5	0	0	1	0	1	2
		1	0	0	0	0	0	0
		2	0	0	0	1	0	1
	4	3	0	0	4	3	2	9
		4	0	0	3	5	7	15
		5	0	0	2	3	2	7
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	5	3	0	0	0	0	0	0
		4	0	0	0	3	1	4
		5	0	0	0	2	5	7
		Overall	0	1	18	30	20	69

Q3		Second State	Third State					
			1	2	3	4	5	Total
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	1	3	0	0	0	0	0	0
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	2	3	0	0	2	0	0	2
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
		1	0	0	0	0	0	0
		2	0	0	1	0	0	1
First State	3	3	0	0	9	6	0	15
		4	0	0	6	5	0	11
		5	0	0	0	0	0	0
		1	0	0	0	0	0	0
		2	0	0	0	1	0	1
	4	3	0	0	2	2	0	4
		4	0	1	1	12	2	16
		5	0	0	0	4	4	8
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	5	3	0	0	0	0	0	0
		4	0	0	0	3	1	4
		5	0	0	0	3	4	7
		Overall	0	1	21	36	11	69

Q4		Second State	Third State					
			1	2	3	4	5	Total
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	1	3	0	0	0	0	0	0
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	2	3	0	0	0	0	0	0
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
		1	0	0	0	0	0	0
		2	0	0	0	1	0	1
First State	3	3	0	0	4	6	0	10
		4	0	0	6	7	2	15
		5	0	0	1	0	0	1
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	4	3	0	1	1	4	3	9
		4	0	0	3	10	6	19
		5	0	0	1	4	2	7
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	5	3	0	0	0	0	0	0
		4	0	0	0	1	2	3
		5	0	0	0	2	2	4
		Overall	0	1	16	35	17	69

Q5		Second State	Third State					
			1	2	3	4	5	Total
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	1	3	0	0	0	0	0	0
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	2	3	0	0	0	0	0	0
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
First State	3	3	0	0	6	6	0	12
		4	0	0	8	5	0	13
		5	0	0	0	0	0	0
		1	0	0	0	0	0	0
		2	0	0	0	1	0	1
	4	3	0	0	2	4	0	6
		4	0	1	3	10	2	16
		5	0	0	0	5	4	9
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	5	3	0	0	0	0	0	0
		4	0	0	0	3	1	4
		5	0	0	0	3	5	8
		Overall	0	1	19	37	12	69

Q6		Second State	Third State					
			1	2	3	4	5	Total
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	1	3	0	0	0	0	0	0
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	2	3	0	0	0	0	0	0
		4	0	0	0	0	0	0
		5	0	0	0	0	0	0
		1	0	0	0	0	0	0
		2	0	0	1	0	0	1
First State	3	3	0	0	4	7	1	12
		4	0	0	6	8	3	17
		5	0	0	1	0	0	1
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	4	3	0	0	1	4	2	7
		4	0	0	3	9	4	16
		5	0	0	1	2	3	6
		1	0	0	0	0	0	0
		2	0	0	0	0	0	0
	5	3	0	1	0	0	0	1
		4	0	0	0	1	2	3
		5	0	0	1	1	3	5
		Overall	0	1	18	32	18	69

To overcome the two stated problem of division by zero not counting the empty cells toward the number of degrees of freedom used in the statistical tests, it was decided to combine the results from ratings 1 to 3 into a single state for the purposes of modelling as a Markov chain. The resulting frequency tables can be seen in Table 6.6 and Table 6.7

Table 6.6. Frequencies of occurrence for sets of three consecutive ratings for subcontractors that consistently worked for the same contractors after combining ratings 1-3 into a single state

Q1		Second State	Thirde State			
			3	4	5	Total
First State	(1-3)	(1-3)	7	2	0	9
		4	1	1	0	2
		5	1	1	0	2
	4	(1-3)	0	1	0	1
		4	1	9	2	12
		5	0	2	2	4
	5	(1-3)	1	0	0	1
		4	0	0	1	1
		5	4	0	10	14
	Total		15	16	15	46

Q2		Second State	Thirde State			
			(1-3)	4	5	Total
First State	(1-3)	(1-3)	3	3	0	6
		4	2	0	0	2
		5	3	1	0	4
	4	(1-3)	1	0	0	1
		4	2	8	1	11
		5	0	0	3	3
	5	(1-3)	0	0	0	0
		4	0	0	2	2
		5	0	1	16	17
	Total		11	13	22	46

Q3		Second State	Thirde State			
			(1-3)	4	5	Total
First State	(1-3)	(1-3)	8	1	0	9
		4	2	2	1	5
		5	3	0	0	3
	4	(1-3)	1	1	0	2
		4	0	8	1	9
		5	0	0	4	4
	5	(1-3)	0	0	0	0
		4	0	1	1	2
		5	0	1	11	12
	Total		14	13	18	46

Q4		Second State	Thirde State			
			(1-3)	4	5	Total
First State	(1-3)	(1-3)	3	2	0	5
		4	2	2	0	4
		5	0	0	2	2
	4	(1-3)	1	1	0	2
		4	1	10	2	13
		5	0	2	2	4
	5	(1-3)	0	0	0	0
		4	0	2	3	5
		5	3	1	7	11
	Total		9	20	16	46

Q5		Second State	Thirde State			
			(1-3)	4	5	Total
First State	(1-3)	(1-3)	3	2	0	5
		4	2	0	0	2
		5	0	1	0	1
	4	(1-3)	2	1	0	3
		4	1	11	3	15
		5	1	1	2	4
	5	(1-3)	0	0	0	0
		4	0	3	1	4
		5	3	1	8	12
	Total		11	20	14	46

Q6		Second State	Thirde State			
			(1-3)	4	5	Total
First State	(1-3)	(1-3)	12	0	0	12
		4	3	1	0	4
		5	4	1	1	6
	4	(1-3)	1	2	1	2
		4	0	5	1	6
		5	0	0	0	0
	5	(1-3)	0	0	0	0
		4	1	1	0	2
		5	0	2	10	12
	Total		20	11	13	46

Table 6.7. Frequencies of occurrence for sets of three consecutive ratings for subcontractors that consistently worked for the multiple contractors after combining ratings 1-3 into a single state

Q1		Second State	Thirde State			Total
			3	4	5	
First State	(1-3)	(1-3)	13	5	0	18
		4	3	4	1	8
		5	0	0	0	0
	4	(1-3)	3	2	0	5
		4	0	12	2	14
		5	0	4	6	10
	5	(1-3)	0	0	0	0
		4	0	2	2	4
		5	0	5	5	10
	Total		26	29	14	69

Q2		Second State	Thirde State			Total
			(1-3)	4	5	
First State	(1-3)	(1-3)	2	7	0	9
		4	7	6	2	15
		5	1	0	1	2
	4	(1-3)	4	4	2	10
		4	3	5	7	15
		5	2	3	2	7
	5	(1-3)	0	0	0	0
		4	0	3	1	4
		5	0	2	5	7
	Total		26	32	11	69

Q3		Second State	Thirde State			Total
			(1-3)	4	5	
First State	(1-3)	(1-3)	12	6	0	18
		4	6	5	0	11
		5	0	0	0	0
	4	(1-3)	2	3	0	5
		4	2	12	2	16
		5	0	4	4	8
	5	(1-3)	0	0	0	0
		4	0	3	1	4
		5	0	3	4	7
	Total		29	29	11	69

Q4		Second State	Thirde State			Total
			(1-3)	4	5	
First State	(1-3)	(1-3)	4	7	0	11
		4	6	7	2	15
		5	1	0	0	1
	4	(1-3)	2	4	3	9
		4	3	10	6	19
		5	1	4	2	7
	5	(1-3)	0	0	0	0
		4	0	1	2	3
		5	0	2	2	4
	Total		27	35	7	69

Q5		Second State	Thirde State			Total
			(1-3)	4	5	
First State	(1-3)	(1-3)	6	6	0	12
		4	8	5	0	13
		5	0	0	0	0
	4	(1-3)	2	5	0	7
		4	4	10	2	16
		5	0	5	4	9
	5	(1-3)	0	0	0	0
		4	0	3	1	4
		5	0	3	5	8
	Total		25	32	12	69

Q6		Second State	Thirde State			Total
			(1-3)	4	5	
First State	(1-3)	(1-3)	5	7	1	13
		4	6	8	3	17
		5	1	0	0	1
	4	(1-3)	1	4	2	7
		4	3	9	4	16
		5	1	2	3	6
	5	(1-3)	1	0	0	1
		4	0	1	2	3
		5	1	1	3	5
	Total		31	29	9	69

Next are the values of $f_{ij.}$ that were calculated and are given in Table 6.8 and Table 6.9.

The values of $f_{.jk}$ that were calculated and are given in Table 6.10 and Table 6.11

Table 6.8. The results of calculating the *fij*. for sets of three consecutive ratings for subcontractors that consistently worked for the same contractors after combining ratings 1-3 into a single state

Q1 (fij)		Second State			Total
		(1-3)	4	5	
First State	(1-3)	9	2	2	13
	4	1	12	4	17
	5	1	1	14	16
	Total	11	15	20	46

Q2 (fij)		Second State			Total
		(1-3)	4	5	
First State	(1-3)	6	2	4	12
	4	1	11	3	15
	5	0	2	17	19
	Total	7	15	24	46

Q3 (fij)		Second State			Total
		(1-3)	4	5	
First State	(1-3)	9	5	3	17
	4	2	9	4	15
	5	0	2	12	14
	Total	11	16	19	46

Q4 (fij)		Second State			Total
		(1-3)	4	5	
First State	(1-3)	5	4	2	11
	4	2	13	4	19
	5	0	5	11	16
	Total	7	22	17	46

Q5 (fij)		Second State			Total
		(1-3)	4	5	
First State	(1-3)	5	2	1	8
	4	3	15	4	22
	5	0	4	12	16
	Total	8	21	17	46

Q6 (fij)		Second State			Total
		(1-3)	4	5	
First State	(1-3)	12	4	6	22
	4	4	6	0	10
	5	0	2	12	14
	Total	16	12	18	46

Table 6.9 The results of calculating the *fij*. for sets of three consecutive ratings for subcontractors that consistently worked for the multiple contractors after combining ratings 1-3 into a single state

Q1		fij. Second State			
		(1-3)	4	5	Total
First State	(1-3)	18	8	0	26
	4	5	14	10	29
	5	0	4	10	14
	Total	23	26	20	69

Q2		fij. Second State			
		(1-3)	4	5	Total
First State	(1-3)	9	15	2	26
	4	10	15	7	32
	5	0	4	7	11
	Total	19	34	16	69

Q3		fij. Second State			
		(1-3)	4	5	Total
First State	(1-3)	18	11	0	29
	4	5	16	8	29
	5	0	4	7	11
	Total	23	31	15	69

Q4		fij. Second State			
		(1-3)	4	5	Total
First State	(1-3)	11	15	1	27
	4	9	19	7	35
	5	0	3	4	7
	Total	20	37	12	69

Q5		fij. Second State			
		(1-3)	4	5	Total
First State	(1-3)	12	13	0	25
	4	7	16	9	32
	5	0	4	8	12
	Total	19	33	17	69

Q6		fij. Second State			
		(1-3)	4	5	Total
First State	(1-3)	13	17	1	31
	4	7	16	6	29
	5	1	3	5	9
	Total	21	36	12	69

Table 6.10. The results of calculating the *f.jk* for sets of three consecutive ratings for subcontractors that consistently worked for the same contractors after combining ratings 1-3 into a single state

Q1		f.jk Third state			
		(1-3)	4	5	Total
Seoncd State	(1-3)	8	3	0	11
	4	2	10	3	15
	5	5	3	12	20
	Total	15	16	15	46

Q2		f.jk Third state			
		(1-3)	4	5	Total
Seoncd State	(1-3)	4	3	0	7
	4	4	8	3	15
	5	3	2	19	24
	Total	11	13	22	46

Q3		f.jk Third state			
		(1-3)	4	5	Total
Seoncd State	(1-3)	9	2	0	11
	4	2	11	3	16
	5	3	1	15	19
	Total	14	14	18	46

Q4		f.jk Third state			
		(1-3)	4	5	Total
Seoncd State	(1-3)	4	3	0	7
	4	3	14	5	22
	5	3	3	11	17
	Total	10	20	16	46

Q5		f.jk Third state			
		(1-3)	4	5	Total
Seoncd State	(1-3)	5	3	0	8
	4	3	14	4	21
	5	4	3	10	17
	Total	12	20	14	46

Q6		f.jk Third state			
		(1-3)	4	5	Total
Seoncd State	(1-3)	13	2	1	16
	4	4	7	1	12
	5	4	3	11	18
	Total	21	12	13	46

Table 6.11. The results of calculating the *f.jk* for sets of three consecutive ratings for subcontractors that consistently worked for the multiple contractors after combining ratings 1-3 into a single state

Q1		f.jk			
		Third State			
		(1-3)	4	5	Total
Seoncd State	(1-3)	16	7	0	23
	4	3	18	5	26
	5	0	9	11	20
	Total	19	34	16	69

Q2		f.jk			
		Third State			
		(1-3)	4	5	Total
Seoncd State	(1-3)	6	11	2	19
	4	10	14	10	34
	5	3	5	8	16
	Total	19	30	20	69

Q3		f.jk			
		Third State			
		(1-3)	4	5	Total
Seoncd State	(1-3)	14	9	0	23
	4	8	20	3	31
	5	0	7	8	15
	Total	22	36	11	69

Q4		f.jk			
		Third State			
		(1-3)	4	5	Total
Seoncd State	(1-3)	6	11	3	20
	4	9	18	10	37
	5	2	6	4	12
	Total	17	35	17	69

Q5		f.jk			
		Third State			
		(1-3)	4	5	Total
Seoncd State	(1-3)	8	11	0	19
	4	12	18	3	33
	5	0	8	9	17
	Total	20	37	12	69

Q6		f.jk			
		Third State			
		(1-3)	4	5	Total
Seoncd State	(1-3)	7	11	3	21
	4	9	18	9	36
	5	3	3	6	12
	Total	19	32	18	69

Finally, the two-way independence, Markovity, and two-way by one way independence were calculated. Detailed calculations for all components for the first question for group one about subcontractors that worked for the same contractors are as follows:

$$\sum_{i=1}^{r}\sum_{j=1}^{r}\sum_{k=1}^{r} 2f_{ijk}\ln f_{ijk} = 2(7\ln 7) + 2(2\ln 2) + \cdots + 2(10\ln 10) = 1617.890$$

$$\sum_{j=1}^{r} 2f_{.j.}\ln f_{.i.} = 2(11\ln 11) + 2(15\ln 15) + 2(20\ln 20) = 253.824$$

$$\sum_{i=1}^{r}\sum_{j=1}^{r} 2f_{ij.}\ln f_{ij.} = 2(9\ln 9) + 2(2\ln 2) + \cdots + (2(1\ln 1) + 2(12\ln 12) = 189.717$$

$$\sum_{j=1}^{r}\sum_{k=1}^{r} 2f_{.jk}\ln f_{.jk} = 2(8\ln 8) + 2(3\ln 3) \ldots + 2(3\ln 3) + 2(12\ln 12) = 177.603$$

$$\sum_{i=1}^{r} 2f_{i..}\ln f_{i..} = 2(13\ln 13) + 2(17\ln 17) + 2(16\ln 16) = 251.741$$

$$2\ln n = 2f_{...}\ln f_{...} = 2(46\ln 46) = 352.235$$

So, by using the above calculated values, the three components are obtained as follows:

1- Two-way independence component value is:

189.717 + 352.235 - 251.741 - 253.824 = 36.387

The degree of freedom to be calculated as follows:

$(r-1)^2 = (3-1)^2 = 4$

If the null hypothesis is true then 36.387 comes from a chi squared distribution with 4 degrees of freedom.

Therefore the p-value is: 2.41E-07

2- The Markovity component value is:

1617.890 + 253.824 – 189.717 – 177.603 = 1504.395

The degree of freedom to be calculated as follows:

$r(r-1)^2 = 3(3-1)^2 = 12$

The p-value is: 0.043

3- The two-way by one way independence component value is:

1617.890 + 352.253 – 251.741 – 177.603 = 1540.782

The degree of freedom to be calculated as follows:

$(r-1)(r^2-1) = (3-1)(3^2-1) = 16$

The p-value is: 1.17E-06

These calculations are then applied on each question in the two groups of subcontractors to find out the Markovity to decide then if both groups or one of them is a first order Markov or not.

The summary of all results is given in Table 6.12 in p-values.

Table 6.12. The results of the Markovity for converting each group to a first order Markov Chains after extracting third states of 0 and 2; and removing second states of 1 and 2

	Components						
Same Contractors	**Question No.**	**Two way independent**	**D.F.**	**Markovity**	**D.F.**	**Two way by one-way independent**	**D.F.**
	1	2.41E-07	4	0.043	12	1.17E-06	16
	2	1.70E-06	4	0.001	12	4.47E-08	16
	3	4.06E-05	4	0.012	12	1.57E-05	16
	4	9.21E-04	4	0.227	12	5.57E-03	16
	5	4.41E-05	4	0.274	12	8.55E-04	16
	6	2.92E-06	4	0.001	12	1.58E-07	16

	Components						
Multiple Contractors	**Question No.**	**Two way independent**	**D.F.**	**Markovity**	**D.F.**	**Two way by one-way independent**	**D.F.**
	1	1.19E-08	4	0.569	12	6.88E-06	16
	2	2.58E-03	4	0.089	12	3.56E-03	16
	3	3.19E-07	4	0.509	12	6.73E-05	16
	4	7.63E-03	4	0.147	12	1.35E-02	16
	5	7.34E-06	4	0.506	12	6.80E-04	16
	6	6.70E-03	4	0.410	12	4.56E-02	16

6.4 Discussion:

The results for the Markovity in Table 6.12 show that group 1 is not accurately modelled by a first order Markov chain whereas group 2 is. The results for group 1, subcontractors that worked for the same contractors, show that the p-value for the Markovity test for each performance question is very small leading to rejecting the null hypothesis that it is a first order Markov chain. Whereas for group 2 the p-values seem to be more or less evenly spread from 0 to 1 and so the null hypothesis is accepted that the data is generated by a first order Markov chain process.

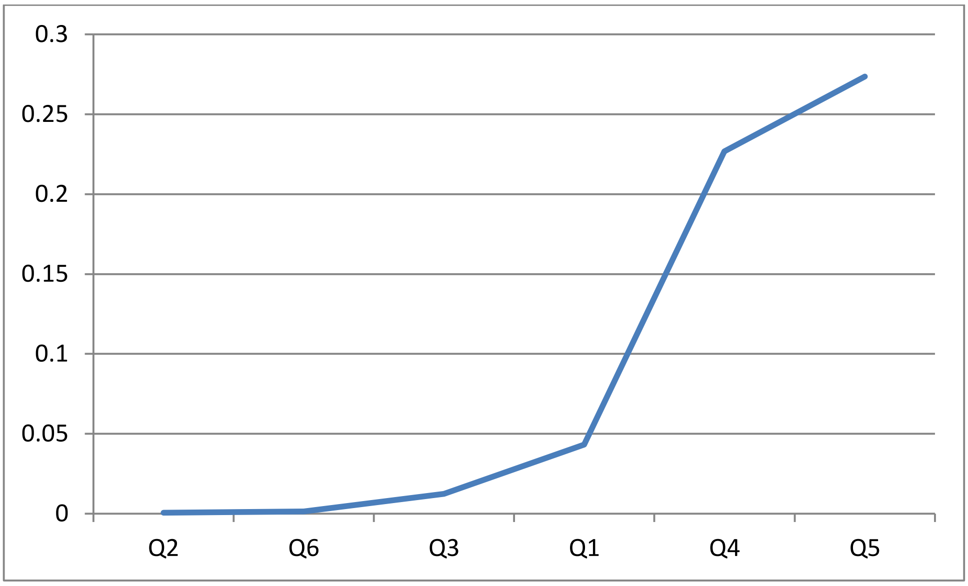

Figure 6.2. Markovity results for performance data set 1 for subcontractors that worked for the same contractors testing the possibility of accepting it as a first order Markov chain. Four of the six questions have p-values less than 0.05indicating that it is unlikely that the results are generated by a first order Markov chain.

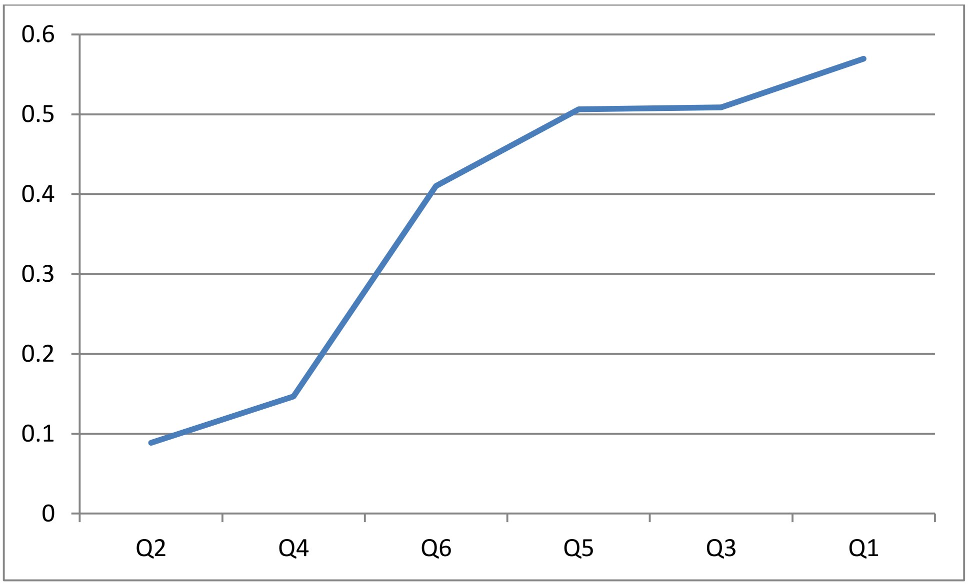

Figure 6.3. Markovity results for performance data set 2 for subcontractors that worked for the multiple contractors testing the possibility of accepting it as a first order Markov chain. The even spread of results is consistent with generation by a first order Markov chain.

This raises the question of what is the nature of the deviation of group 1 from a first order Markov chain. Looking at the data this appears to take the form of subcontractors having identical ratings three times in a row more often than would be expected if it was a first order Markov chain. To investigate this, the following calculations determine what the performance of subcontractors that worked with the same contractors would be expected to be if the data did come from a first order Markov chain. These expected values will later be compared with the actual values. To do this, the following steps are applied:

1- Converting the size of the transition matrix of each performance question from 5 by 5 to 3 by 3 as that was applied when the data was examined for Markovity. See the following example:

For question one about the performance of the project management of subcontractors, the first order transition matrix will be converted by grouping ratings 1 – 3 into a single rating / state:

Table 6.13. Transition Matrix of question 1 of the performance for subcontractors that worked with the same contractors assuming that the data fits a first order Markov chain

	1	**2**	**3**	**4**	**5**
1	0	0	0	0	0
2	0	1	1	2	2
3	0	0	15	3	0
4	0	1	2	22	7
5	0	0	6	4	26

→

	(1-3)	**4**	**5**	**Total**
(1-3)	17	5	2	24
4	3	22	7	32
5	6	4	26	36
Total	26	31	35	92

2- Next, the expected frequencies of occurrence for sets of three consecutive ratings will be generated by multiplying the frequencies for the first state in the three state sequences by the probability that it will transition to the next state using the probabilities in Table 6.13 and multiplied again by the probability that it will transition from this second state into the third state also using the probabilities from Table 6.13 see Equation 6.7.

$$Expected\ Performance = P(A|B) * P(A|B) * Total$$

$$Expected\ frequency(S_1 = x, S_2 = y, S_3 = z)$$

$$= Observed\ frequency(S_1 = x) * P(S_2 = y|S_1 = x) * P(S_2 = z|S_1 = y) \qquad \text{Equation 6.7}$$

where

$Observed\ frequency(S_1 = x)$ is the number of times in the set of two state transitions that the initial is state x

$P(S_2 = y|S_1 = x)$ is the probability of a transition from state x to state y using the values from Table 6.13

Table 6.14. Transition probabilities of question 1 of the performance for subcontractors that worked with the same contractors assuming that the data fits a first order Markov chain

	(1-3)	**4**	**5**
(1-3)	0.708	0.208	0.083
4	0.094	0.688	0.219
5	0.167	0.111	0.722

→

Q1		**Second State**	**Third State** (1-3)	**Third State** 4	**Third State** 5	**Total**
First State	**(1-3)**	**(1-3)**	7	2	0	**9**
		4	1	1	0	**2**
		5	1	1	0	**2**
			9	**4**	**0**	**13**
	4	**(1-3)**	0	1	0	**1**
		4	1	9	2	**12**
		5	0	2	2	**4**
			1	**12**	**4**	**17**
	5	**(1-3)**	1	0	0	**1**
		4	0	0	1	**1**
		5	4	0	10	**14**
		Total	**5**	**0**	**11**	**16**
		Total	**13**	**17**	**16**	**46**

These expected frequencies for question one for the subcontractors that worked for the same contractors are shown in Table 6.15

Table 6.15. The results of the expected frequencies for question one for the subcontractors that worked for the same contractors

Q1		Second State	Third State (1-3)	4	5	Total
First State	(1-3)	(1-3)	7.058	1.512	1.008	9.579
		4	0.257	1.368	0.428	2.053
		5	0.274	0.182	0.912	1.368
		Total	7.588	3.063	2.348	13
	4	3	1.566	0.336	0.224	2.125
		4	1.417	7.556	2.361	11.333
		5	0.708	0.472	2.361	3.542
		Total	3.691	8.363	4.946	17
	5	3	2.358	0.505	0.337	3.200
		4	0.267	1.422	0.444	2.133
		5	2.133	1.422	7.111	10.667
		Total	4.758	3.350	7.892	16

Table 6.16. The results of the observed frequencies for question one for the subcontractors that worked for the same contractors

Q1		Second State	Third State (1-3)	4	5	Total
First State	(1-3)	(1-3)	7.058	1.512	1.008	9.579
		4	0.257	1.368	0.428	2.053
		5	0.274	0.182	0.912	1.368
		Total	7.588	3.063	2.348	13
	4	3	1.566	0.336	0.224	2.125
		4	1.417	7.556	2.361	11.333
		5	0.708	0.472	2.361	3.542
		Total	3.691	8.363	4.946	17
	5	3	2.358	0.505	0.337	3.200
		4	0.267	1.422	0.444	2.133
		5	2.133	1.422	7.111	10.667
		Total	4.758	3.350	7.892	16
		Total	16.037	14.776	15.187	46

Q2		Second State	Third State (1-3)	4	5	Total
First State	(1-3)	(1-3)	3.000	1.286	1.714	6.000
		4	0.429	1.714	0.429	2.571
		5	0.294	0.294	2.841	3.429
		Total	3.722	3.294	4.984	12
	4	3	1.250	0.536	0.714	2.500
		4	1.667	6.667	1.667	10.000
		5	0.214	0.214	2.071	2.500
		Total	3.131	7.417	4.452	15
	5	3	0.814	0.349	0.465	1.629
		4	0.271	1.086	0.271	1.629
		5	1.349	1.349	13.044	15.743
		Total	2.435	2.784	13.781	19
		Total	9.289	13.495	23.217	46

Q3		Second State	Third State (1-3)	4	5	Total
First State	(1-3)	(1-3)	6.884	2.459	1.475	10.818
		4	0.644	2.254	0.966	3.864
		5	0.258	0.172	1.889	2.318
		Total	7.786	4.884	4.330	17
	4	3	1.591	0.568	0.341	2.500
		4	1.458	5.104	2.188	8.750
		5	0.417	0.278	3.056	3.750
		Total	3.466	5.950	5.584	15
	5	3	0.990	0.354	0.212	1.556
		4	0.173	0.605	0.259	1.037
		5	1.267	0.845	9.295	11.407
		Total	2.430	1.803	9.766	14
		Total	13.682	12.638	19.680	46

Q4		Second State	Third State (1-3)	4	5	Total
First State	(1-3)	(1-3)	2.396	2.053	0.684	5.133
		4	0.550	2.750	1.100	4.400
		5	0.169	0.338	0.959	1.467
		Total	3.115	5.142	2.743	11
	4	3	1.108	0.950	0.317	2.375
		4	1.484	7.422	2.969	11.875
		5	0.548	1.096	3.106	4.750
		Total	3.141	9.468	6.391	19
	5	3	0.862	0.738	0.246	1.846
		4	0.462	2.308	0.923	3.692
		5	1.207	2.414	6.840	10.462
		Total	2.530	5.460	8.009	16
		Total	8.786	20.070	17.144	46

Q5		Second State	Third State (1-3)	4	5	Total
First State	(1-3)	(1-3)	3.240	1.388	0.463	5.091
		4	0.321	1.540	0.321	2.182
		5	0.104	0.156	0.468	0.727
		Total	3.664	3.084	1.251	8
	4	3	2.059	0.882	0.294	3.235
		4	2.284	10.962	2.284	15.529
		5	0.462	0.693	2.080	3.235
		Total	4.805	12.538	4.658	22
	5	3	1.455	0.623	0.208	2.286
		4	0.504	2.420	0.504	3.429
		5	1.469	2.204	6.612	10.286
		Total	3.428	5.248	7.324	16
		Total	11.897	20.870	13.233	46

Q6		Second State	Third State (1-3)	4	5	Total
First State	(1-3)	(1-3)	10.096	1.923	2.884	14.903
		4	1.002	1.670	0.167	2.839
		5	0.681	0.681	2.895	4.258
		Total	11.779	4.274	5.947	22
	4	3	2.391	0.455	0.683	3.529
		4	2.076	3.460	0.346	5.882
		5	0.094	0.094	0.400	0.588
		Total	4.561	4.010	1.429	10
	5	3	1.517	0.289	0.434	2.240
		4	0.791	1.318	0.132	2.240
		5	1.523	1.523	6.474	9.520
		Total	3.831	3.130	7.039	14
		Total	20.171	11.414	14.415	46

Looking at Table 6.16 shows that even with a first order Markov chain it is expected that mostly subcontractors will get three ratings in a row that are the same. That is, they are more likely to stay in the same state that they were in than to move to higher or lower states. The important question is how this relates to the observed frequencies. Table 6.17 shows the difference between the observed and expected frequencies. (Positive implies observed is larger).

Table 6.17. The results of the difference between the observed and expected frequencies

Q1		Second State	Third State		
			(1-3)	4	5
First State	(1-3)	(1-3)	-0.058	0.488	-1.008
		4	0.743	-0.368	-0.428
		5	0.726	0.818	-0.912
	4	(1-3)	-1.566	0.664	-0.224
		4	-0.417	1.444	-0.361
		5	-0.708	1.528	-0.361
	5	(1-3)	-1.358	-0.505	-0.337
		4	-0.267	-1.422	0.556
		5	1.867	-1.422	2.889

Q2		Second State	Third State		
			(1-3)	4	5
First State	(1-3)	(1-3)	0.000	1.714	-1.714
		4	1.571	-1.714	-0.429
		5	2.706	0.706	-2.841
	4	(1-3)	-0.250	-0.536	-0.714
		4	0.333	1.333	-0.667
		5	-0.214	-0.214	0.929
	5	(1-3)	-0.814	-0.349	-0.465
		4	-0.271	-1.086	1.729
		5	-1.349	-0.349	2.956

Q3		Second State	Third State		
			(1-3)	4	5
First State	(1-3)	(1-3)	1.116	-1.459	-1.475
		4	1.356	-0.254	0.034
		5	2.742	-0.172	-1.889
	4	(1-3)	-0.591	0.432	-0.341
		4	-1.458	2.896	-1.188
		5	-0.417	-0.278	0.944
	5	(1-3)	-0.990	-0.354	-0.212
		4	-0.173	0.395	0.741
		5	-1.267	0.155	1.705

Q4		Second State	Third State		
			(1-3)	4	5
First State	(1-3)	(1-3)	0.604	-0.053	-0.684
		4	1.450	-0.750	-1.100
		5	-0.169	-0.338	1.041
	4	(1-3)	-0.108	0.050	-0.317
		4	-0.484	2.578	-0.969
		5	-0.548	0.904	-1.106
	5	(1-3)	-0.862	-0.738	-0.246
		4	-0.462	-0.308	2.077
		5	1.793	-1.414	0.160

Q5		Second State	Third State		
			(1-3)	4	5
First State	(1-3)	(1-3)	-0.240	0.612	-0.463
		4	1.679	-1.540	-0.321
		5	-0.104	0.844	-0.468
	4	(1-3)	-0.059	0.118	-0.294
		4	-1.284	0.038	0.716
		5	0.538	0.307	-0.080
	5	(1-3)	-1.455	-0.623	-0.208
		4	-0.504	0.580	0.496
		5	1.531	-1.204	1.388

Q6		Second State	Third State		
			(1-3)	4	5
First State	(1-3)	(1-3)	1.904	-1.923	-2.884
		4	1.998	-0.670	-0.167
		5	3.319	0.319	-1.895
	4	(1-3)	-1.391	1.545	0.317
		4	-2.076	1.540	0.654
		5	-0.094	-0.094	-0.400
	5	(1-3)	-1.517	-0.289	-0.434
		4	0.209	-0.318	-0.132
		5	-1.523	0.477	3.526

As can be seen in Table 6.17 the results show that even though the expected frequencies indicate that repeated identical observations are expected to be the most common outcome, the actual results are even more extreme than this, particularly for states 4 and 5. In other words, outcomes of 4,4,4 and 5,5,5 are considerably more common than expected for most questions.

The unexpectedly high frequency of the same state occurring three times is strongest for state 5. It is less strong for state 4 and weakest for state 3. From this it appears that when a contractor forms an opinion that a subcontractor is very good or excellent that opinion is harder to shake than an opinion that the subcontractor is merely good or below.

Analyses were performed to validate the two trends mentioned above.

The first analysis was performed to see how the expected and observed frequencies compared based on the variability of the subcontractor. For this the two step transitions were divided into three groups. In the first group of transitions both steps involved changing states. In the second group of transitions only one step involved changing state. That is, in the second group of transitions the subcontractor received the same rating for two consecutive projects, but the other rating was different. In the third group both transitions involved staying in the same state. That is, the subcontractor received the same rating three times in a row.

Table 6.18 shows the state transitions included in each group, i.e. the cells from Table 6.17. Table 6.19 uses question 1 as an example for how these are calculated. Table 6.20 shows the results for all questions

Table 6.18. How the values for each performance question to be extracted from Table 6.17 to identify the changes in performance

Change in performance		
Change twice	**Change once**	**Not change**
3-4-3	3-3-4	3-3-3
3-4-5	3-3-5	4-4-4
3-5-3	3-4-4	5-5-5
3-5-4	3-5-5	
4-3-4	4-3-3	
4-3-5	4-4-3	
4-5-3	4-4-5	
4-5-4	4-5-5	
5-3-4	5-3-3	
5-3-5	5-4-4	
5-4-3	5-5-3	
5-4-5	5-5-4	

Table 6.19. The results for the change in performance in question 1 for the different three groups by extracting the values from Table 6.17

Change in performance		
Change twice	**Change once**	**Not change**
0.743	0.488	-0.058
-0.428	-1.008	1.444
0.726	-0.368	2.889
0.818	-0.912	
0.664	-1.566	
-0.224	-0.417	
-0.708	-0.361	
1.528	-0.361	
-0.505	-1.358	
-0.337	-1.422	
-0.267	1.867	
0.556	-1.422	
Sum	**Sum**	**Sum**
2.567	**-6.842**	**4.275**

Table 6.20 Observed – expected for subcontractors that changed rating twice, once and not at all. Subcontractors were more likely than expected to change a lot, or not at all

Table 6.20. The results for the change for all performance for the different three groups by extracting the values from Table 6.17

	Change in performance		
Questions	**Change twice**	**Change once**	**Not change**
1	2.567	-6.842	4.275
2	3.519	-7.809	4.289
3	3.360	-9.076	5.717
4	0.562	-3.905	3.342
5	1.927	-3.113	3.753
6	6.497	-13.467	6.970

It can be seen in Table 6.20 that the subcontractors that never change and the subcontractors that change particularly frequently are over represented in the observed data compared to the expected frequencies resulting from a first order Markov chain. To summarise there are two over represented groups, firstly, a group with particularly low variability, and secondly a group with particularly high variability.

The other observation made above was that the group with low variability tended to be high performing subcontractors while the group with high variability tended to be lower performing subcontractors. Table 4 shows the differences between observed and expected for all questions for subcontractors in the low variability group.

Table 6.21. How frequently do subcontractors in group one that worked for the same contractors receive the same rating. Values are extracted from Table 6.17

	Observed frequency – expected frequency		
Questions	**3-3-3**	**4-4-4**	**5-5-5**
1	-0.058	1.444	2.889
2	0.000	1.333	2.956
3	1.116	2.896	1.705
4	0.604	2.578	0.160
5	-0.240	0.038	1.388
6	1.904	1.540	3.526

Table 6.21 shows that subcontractors that receive very good or excellent ratings three times in a row are more over represented in the data than subcontractors that receive good ratings or lower three times in a row.

The group with high variability is defined as all sequences of three successive assessments where the rating changed between both the first and second assessment and also changed between the second and third assessment. In order to investigate this group the various high variability state changes were categorised according to the average rating in each three rating sequence. For example 3-4-3 gives an average of 3.333.

Table 6.22 below shows how each three rating sequence is categorised.

Table 6.22. How the average ratings of performance questions would be extracted from Table 6.21 for the results of the difference between the expected and actual frequencies for subcontractors that worked for the same contractors.

Average Rating						
3.333	**3.667**	**=3.333 + 3.667**	**4.000**	**4.333**	**4.667**	**=4.333 + 4.667**
3-4-3	4-3-4	3-4-3 + 4-3-4	3-4-5	5-3-5	5-4-5	5-3-5 + 5-4-5
	3-5-3	3-5-3	3-5-4	4-5-4		5-4-5
			4-3-5			
			4-5-3			
			5-3-4			
			5-4-3			
=$\sum \mathbf{3.333}$	=$\sum \mathbf{3.667}$	=$\sum(\mathbf{3.333} + \mathbf{3.667})$	=$\sum \mathbf{4.000}$	=$\sum \mathbf{4.333}$	=$\sum \mathbf{4.667}$	=$\sum(\mathbf{4.333} + \mathbf{4.667})$

Table 6.23 shows the values of observed – expected for all cells listed in Table 5, once again using question 1 as an example. Table 6.24 then combines the results for all

questions.

Table 6.23. The results of the average ratings for question 1 of performance that was extracted from Table 6.21.

Average Rating						
3.333	**3.667**	**=3.333 + 3.667**	**4.000**	**4.333**	**4.667**	**=4.333 + 4.667**
0.743	0.664	1.407	-0.428	-0.337	0.556	0.219
	0.726	0.726	0.818	1.528		1.528
			-0.224			
			-0.708			
			-0.505			
			-0.267			
0.743	1.391	2.133	-1.314	1.191	0.556	1.747

Table 6.24. The results of the average ratings for question 1 of performance that was extracted from Table 6.21.

Observed frequency – expected frequency	**Average Rating**				
Questions	**3.333**	**3.667**	**4.000**	**4.333**	**4.667**
1	0.743	1.391	-1.314	1.191	0.556
2	1.571	2.170	-1.271	-0.680	1.729
3	1.356	3.174	-1.422	-0.490	0.741
4	1.450	-0.119	-3.503	0.658	2.077
5	1.679	0.014	-0.361	0.099	0.496
6	1.998	4.863	0.295	-0.528	-0.132

The different columns in Table 6.24 do not all represent the same number of cells. However, the results for 3.333 can be compared with 4.667, and also 3.667 can be compared with 4.333. Making these comparisons show that there is a greater overrepresentation of the low average rating results than of the high average rating results. That is, the overrepresentation for average = 3.333 is generally speaking larger than the over representation of average = 4.667 and also the overrepresentation for average = 3.667 is generally speaking larger than the over representation of average = 4.333. Thus it can be concluded that the high variability group tends to have a lower average performance.

After justifying the reasons which prevent the group 1 for subcontractors that worked for the same contractors from being a first order Markov chain, testing the possibility of combining both groups in one group regardless their pattern in working with the same or multiple contractors cannot be conducted.

6.5 Conclusion:

In conclusion, the aim of this chapter is to examine if each of the two sets of data of subcontractor performance that either worked for the same or multiple contractors is a first order Markov chain or not to be then tested using Markov chain and hidden Markov models. The secondary aim was to test the possibility of combining both groups, if they are first order Markov, as one group. The complete data set was divided into two groups. The first group comprised subcontractors that always worked for the same contractor and the second group comprised subcontractors that never worked for the same contractor twice. All of the subcontractors in the first group came from data set 1. The second group contained all of data set 2 as well as some of the subcontractors in data set 1. Subcontractors that worked for a contractor more than once and also worked for other contractors were not included in this analysis. Group 1 represented 46 subcontractors whereas group 2 represented 69 subcontractors. Examining this hypothesis was conducted by using the method described in (Kullback et al., 1962). The results show that only the data for group 2 about subcontractors that worked for multiple contractors are first order Markov chain whereas the data of group 1 is not.. This result prevents testing if both groups could be combined as one group regardless their work for the same or multiple contractors. The result of this is that in the next chapter only the second group of subcontractors will be analysed using Markov chain methods.

Group 1 was further studied to find out the reasons that prevent it from being a first order Markov chain. This involved comparing the expected and actual performance for subcontractors that worked for the same contractors. It was found that subcontractors are more likely to receive the same rating on consecutive jobs than would be expected from a first order Markov chain. This was particularly the case for subcontractors that tended to receive the higher ratings, i.e. states 4 and 5 but not for state (1-3). As a result, another analysis was conducted to examine what pattern might relate to state (1-3). This analysis was based on the observation that sets of three consecutive ratings where the rating changed in both steps was more frequent than would be expected from a first order Markov chain. The analysis showed that these extra variable sets of observations tended to be associated with subcontractors that on average had lower ratings. Thus, the results of this chapter lead to the fact that Markov chain will be implemented only on subcontractors that worked for multiple contractors in the following chapter. The results of the analysis and investigations in this chapter are expected to help in understanding the process of Markov chains on that data when forecasting the future performance of subcontractors and also when applying hidden Markov models to deal with subcontractors that have incomplete records so the proposed methodology of using hidden Markov models is validated.

Chapter 7 Analysing the Performance of Subcontractors using Markov Chains

7.1 Introduction

The previous chapter examined the validity of modelling performance data using Markov chains. The results show that Markov chain cannot be accurately implemented on the data for subcontractors that worked for the same contractors whereas the data of subcontractors worked for multiple contractors did fit the model. Therefore this chapter will use Markov chains to model change in each performance question over time to study how the performance levels of subcontractors is changing over time until they reach their maximum limits. As stated in Chapter 3, which describes the data collection for this research, there are six different performance factors used to assess subcontractors in Saudi Arabia. Using Markov chains to model performance will help in studying the change in subcontractor performance for each performance question. Questions that will be asked include how stable are performance ratings. That is do subcontractors tend to get the same ratings over time or do they tend to change. How often should records be updated? It is often said that past data does not guarantee future performance. The question is how reliable are future estimates made on past data? Furthermore, are there identifiable factors that affect the likelihood that a particular subcontractor will change performance level? It will also show if there is a common basis of some changes in all factors or if there are any specific characteristics for any performance factors

To implement Markov chains, the two collected data sets are studied in chapter 6 to examine their eligibility to be analysed using Markov chains if they are first order Markov chain. They were divided into two groups where group 1 was about subcontractors that worked for the same contractors and group 2 was about subcontractors that worked for multiple contractors. The results show that Markov chains method is to be implemented only on group 2 whereas group 1 is not because it is not a first order Markov chains. For further details, refer to Chapter 6 to have a better understanding of why group 1 is not a first order Markov Chain.

7.1.1 Markov chains:

A Markov chain is a dynamical system that represent a mathematical model of sequences of random variables that evolve with time where the future is only affected by the past through the present (Konstantopoulos, 2009). A finite Markov chain is a stochastic process where it moves within a finite number of states with respect to that the probability which is entering any specific state depends only on the last state that is occupied (Kemeny and Snell, 1960).

Markov Chains have been widely used as a tool for analysing different types of stochastic systems over time, (Ching and Ng, 2006). Moreover, they have also been used as a decision making tool to forecast the future outcomes of systems based on the current position or status. (Brémaud, 1999) claims that, despite the limitation that future states are only based on the current, or most recent, state and a transition matrix, Markov chains provide diverse behaviours. This has resulted in Markov Chains being implemented in widely different areas, including engineering, where they are capable of providing quantitative and qualitative outcomes as answers to questions. The time in Markov Chains can be discrete, continuous or a totally ordered set (Konstantopoulos, 2009). According to (Kemeny and Snell, 1960), a Markov process is a finite stochastic process where the past data should not affect the forecasting of the future except in so far as it causes the present.

7.1.2 Markov Property:

(Riordan, 2014) lists the properties that a process should have if it is to be modelled by a Markov chain. The first property is that there should be a Transition Matrix that has entries representing probabilities for transitioning from the current state (row) to the succeeding state (columns), see Equation 7.1.

$$P = \begin{bmatrix} p_{11} p_{12} \cdots p_{1r} \\ p_{21} p_{22} \ldots p_{2r} \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ p_{r1} p_{r2} \cdots p_{rr} \end{bmatrix}$$

Where $0 \le p_{ij} \le 1$, and for all *i* there is

$$\sum_{j=1}^{r} p_{ij} = \sum_{j=1}^{r} P(X_{m+1} = j | X_m = i) = 1 \qquad \text{Equation 7.1}$$

The second property is that there are initial state probabilities that specify the probabilities that the Markov chain starts in any particular state as shown in Equation 7.2

$$S_0(i) = \frac{x(i)}{\sum x(i)} \qquad \text{Equation 7.2}$$

Where:

i represents the different states.

$S_0(i)$: the initial probability of being in state *i*.

$x(i)$: the number of sequences where the first observation is of state *i*.

$\sum x(i)$: the total number of sequences in the data set.

The probabilities of being in any particular state at time t_1 can be obtained by multiplying the initial probabilities by the transition matrix as shown in Equation 7.3

$$S_1 = S_0 P \qquad \text{Equation 7.3}$$

Where:

S_0: the initial state probability vector

S_1: the state probability vector at time t_1

P: the transition matrix.

This process could be repeated until the required time is reached. To repeat this process based on specific number of time steps, Equation 7.4 is used as shown

$$S_k = S_{k-1} P = S_0 P^k \qquad \text{Equation 7.4}$$

Where ***k*** refers to the number of time steps experienced by the system to achieve the required outcomes. As ***k*** becomes large the repetitions of this process lead to the probabilities stabilising. This is the third property of Markov chains which is the stationary probability vector.

The stationary probability vector gives the long run probabilities that the Markov chain is in any particular state. This vector can be obtained either through continually iterating Equation 5.3 until the difference between S_k and S_{k-1} is as small as desired, or by solving Equation 7.5 for S.

$$SP = S \qquad \text{Equation 7.5}$$

Where S is the final result as a stationary state and P is the transition matrix. Markov chains with regular transition Matrices always have a stationary state (Riordan, 2014). If all entries in the transition matrix are positive (i.e. not zero) then the transition matrix is regular.

7.2 Methodology:

Implementing Markov chains involves two steps: determination of the initial probabilities and determination of the transition matrices for each question. These two are calculated as follows:

1- The initial probability vector for each of the performance questions will be calculated. An example will be given using data set 2. This includes subcontractors that worked for multiple contractors, which were demonstrated in the previous chapter to fit a first order Markov process. That data set has 36 subcontractors, which means that the initial probabilities of the three states, representing the ratings and labelled (1-3), 4 and 5, will be derived from the first 36 projects done by those subcontractors. Calculating these initial probabilities will be through Equation 7.2. In this case:

i: the states (1-3), 4 and 5.

$x(i)$: the number of subcontractors that started in state i.

$\sum x(i)$: the total number of subcontractors in the data set (i.e. 36).

As an example of identifying the initial probability vector the results for question 1 are:

$$S_0(1-3) = \frac{14}{36} = 0.389$$

$$S_0(4) = \frac{18}{36} = 0.500$$

$$S_0(5) = \frac{4}{36} = 0.111$$

$$S_0 = [0.389 \quad 0.500 \quad 0.111]$$

2- The transition matrix for each performance question will show how the performance of subcontractors changes from project to project through identifying the probabilities of staying in the same state or changing from one state to another. Calculating the transition matrices will utilise the performance data from all projects. Following the example above, the 36 subcontractors conducted 141 projects. All of this data will be used. Establishing the transition matrix for each performance question will use Equation 4.1.

As an example of identifying the transition matrix, the results for question 1 are shown below:

- Firstly, the process starts by identifying how many transitions involve staying in the same state and how many involve moving to each other state as shown in Table 7.1:

Table 7.1. The number of transitions for each state between each pair of states for performance question 1.

	(1-3)	**4**	**5**	**Total**
(1-3)	25	12	0	37
4	6	25	13	44
5	0	9	15	24

- Next, each number in Table 7.1 is divided by the total for that row. This gives the transition matrix as shown below.

$$P = \begin{bmatrix} 0.676 & 0.324 & 0 \\ 0.136 & 0.568 & 0.295 \\ 0 & 0.375 & 0.625 \end{bmatrix}$$

After acquiring the transition matrix for performance question 1, the next step is to determine the limit state probabilities using Equation 7.3. As an example, applying Equation 7.3 to the results of the first question gives

$$S_1 = [0.389 \quad 0.500 \quad 0.111] \times \begin{bmatrix} 0.676 & 0.324 & 0 \\ 0.136 & 0.568 & 0.295 \\ 0 & 0.375 & 0.625 \end{bmatrix}$$

$$S_1 = [0.331 \quad 0.452 \quad 0.217]$$

Continuing this process according to Equation 7.3 gives the results provided in Table 7.2

Table 7.2. The limiting matrices for all the states of performance question 1.

Q1	P(1-3)	P(4)	P(5)
S_0	0.389	0.500	0.111
S_1	0.331	0.452	0.217
S_2	0.285	0.446	0.269
S_3	0.253	0.447	0.300
S_4	0.232	0.448	0.319
S_5	0.218	0.450	0.332
S_6	0.209	0.451	0.340
S_7	0.202	0.452	0.346
S_8	0.198	0.452	0.350
S_9	0.196	0.452	0.352
S_{10}	0.194	0.452	0.354
S_{11}	0.193	0.453	0.355
S_{12}	0.192	0.453	0.355
S_{13}	0.191	0.453	0.356

Rearranging Equation 7.3 gives Equation 7.4. Equation 7.4 is repeated until the error is sufficiently small. In this case the error being sufficiently small is taken as being less than 0.001.

$$error = S_k - S_{k-1}P \qquad \text{Equation 7.6}$$

After determining the initial and the final probability vectors they can be compared in order to help in identifying in what way the performance of those subcontractors is changing, for example whether it is improving or not.

The final thing to be determined from the Markov chain model is an estimate of the appropriate updating time to use for the performance records. In the interviews conducted in Saudi Arabia with 5 large organisations that have historical records about the performance of subcontractors (discussed in chapter 3), the results show that the updating time used by those companies for their records ranges over at least from 1 to 5 years, see Table 7.3.

Table 7.3. Updating time of historical records in Saudi Arabia

Organisations	Updating Frequency for Historical Records (Years)
O1	1
O2	2
O3	- 2-3 for medium size projects - 5years for Large Projects
O4	2-3 years
O5	2 years

The reasons for these different updating times of the historical records are discussed in more detail in chapter 3. In short, they are all about the strategy that each organisation has to assess subcontractors to accept their bids for their projects.

The data will also be used to see how much error results from using historical performance measurements. The point of a Markov chain is that the state changes over time and so data that is very old is not very useful. The question is how much change occurs and over what time frame. The data in set 2 is based on annual subcontractor ratings. The total number of projects is 4 for each subcontractor. Thus, the updating time will be identified by testing the data of performance over these 4 years.

To do this the identity matrix will be repetitively multiplied by the transition matrix for each performance question and the number of repetitions required before the probability drops to a given level. The identity matrix is used as the starting point in this calculation because for any subcontractor that has a particular rating in hand their probability of getting that rating is 1, since they already have it. As time progresses they are less and less likely to stay in that particular rating state since they will eventually move up or down to another state. Counting the repetitions gives a measure of how long they stay in each state.

7.3 Results and Discussion:

As mentioned in the methodology section, this chapter has two main goals. The first one is to identify the change in performance for subcontractors that worked for multiple contractors. That is, whether there is an improvement or not and the mechanism behind it. The results of this show that, generally, the performance of subcontractors is improving. The second goal is to determine an appropriate updating time for the historical records of subcontractors. The general result shows that the historical records should be updated annually.

7.3.1 Recognising the Change in Performance and its Implications

Table 7.4 shows the initial probability vectors. It clearly shows that subcontractors are most likely to start either in state (1-3) or in state 4 and unlikely to start in state 5. The transition matrices for each performance question are shown in Table 7.5.

Table 7.4. The initial matrices for performance question for subcontractor that worked for multiple contractors

	Initial Matrices for performance questions					
	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**
P(1-3)	0.389	0.417	0.472	0.389	0.417	0.500
P(4)	0.500	0.361	0.417	0.472	0.500	0.333
P(5)	0.111	0.222	0.111	0.139	0.083	0.167

Table 7.5. The transition matrices for performance questions for subcontractors that worked for multiple contractors

Transition Matrices for performance questions

Q1	**(1-3)**	**4**	**5**	**Q4**	**(1-3)**	**4**	**5**
(1-3)	0.676	0.324	0	**(1-3)**	0.412	0.500	0.088
4	0.136	0.568	0.295	**4**	0.259	0.537	0.204
5	0	0.375	0.625	**5**	0.118	0.529	0.353

Q2	**(1-3)**	**4**	**5**	**Q5**	**(1-3)**	**4**	**5**
(1-3)	0.382	0.529	0.088	**(1-3)**	0.471	0.529	0
4	0.298	0.468	0.234	**4**	0.275	0.529	0.196
5	0.125	0.375	0.500	**5**	0	0.400	0.600

Q3	**(1-3)**	**4**	**5**	**Q6**	**(1-3)**	**4**	**5**
(1-3)	0.735	0.441	0	**(1-3)**	0.436	0.487	0.077
4	0.191	0.596	0.191	**4**	0.250	0.521	0.229
5	0	0.333	0.458	**5**	0.222	0.333	0.444

The results in Table 7.5 for subcontractors that worked for multiple contractors show that the questions generally fit two patterns in a similar way that two factors were found in the factor analysis discussion in Chapter 4. Specifically the odd numbered questions, which are related to management attributes, behave differently to the even numbered question, which are related to technical abilities.

The most obvious difference is that subcontractors never moved directly from state (1-3) to state 5 and also never moved from state 5 directly to state (1-3) for the management questions. However, this did happen, although with only low probability, for the technical ability questions.

The second noticeable difference between the groups is that subcontractors are more likely to have the same rating in questions about management attributes than to have the same rating in questions about technical abilities. In other words the diagonal elements are larger for the management attributes than for the technical abilities. Specifically, the probability of remaining in state (1-3) is higher for each of the management attribute questions than for any of the technical ability questions. The same is true for state 4 and for state 5 (although minor overlap exists if different states are compared).

The next step after generating the initial probability vectors and the transition matrices for each performance question is to use the Markov chains to determine the limiting steady state probabilities. This will answer the question of whether the performance of subcontractors that worked for multiple contractors is improving over time or not. Table 7.6 compares the limiting matrices for each performance question with their initial probabilities. Table 7.7 shows the number of iterations required of Equation 7.4 for each performance question needed to reduce the error in Equation 7.6 to be reduced below 0.001.

Table 7.6. Comparing the initial and limiting matrices for each performance question

	The initial and limiting matrices for performance questions											
Questions	**Q1**		**Q2**		**Q3**		**Q4**		**Q5**		**Q6**	
Matrix	L	I	L	I	L	I	L	I	L	I	L	I
P(1-3)	0.389	0.191	0.417	0.276	0.472	0.263	0.389	0.272	0.417	0.258	0.500	0.299
P(4)	0.500	0.453	0.361	0.460	0.417	0.503	0.472	0.525	0.500	0.498	0.333	0.467
P(5)	0.111	0.356	0.222	0.264	0.111	0.234	0.139	0.202	0.083	0.244	0.167	0.234

L: *refers to the limiting matrices*, **I**: *refers to the initial matrices*

Table 7.7.The number of conversions that each performance question needs to form the limiting matrices

	Number of conversions for performance questions					
	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**
No. of Iterations	13	6	12	5	10	5

Table 7.7 shows that subcontractors took much longer to reach the limiting values in the management questions, 1, 3 and 5, compared to the technical ability questions, 2, 4 and 6, taking into account that each single iteration represents one project taking one year. Part of the reason for this is because of the fact that the transition matrix does not allow subcontractors to jump immediately from state (1-3) to state 5 or vice versa. That is, the probability for such a transition is 0%. Thus at least two transitions were required to make this jump instead of the single jump that was possible in the technical ability questions.

Another part of the reason that management questions took longer to stabilise than technical ability questions is that there were larger improvements in the management questions than in the technical ability questions. This can be shown by taking the average for all performance questions. To do this the probability of being in a state was multiplied by the state itself and these values summed separately for the initial and limiting probabilities. For state (1-3) the value 3 was used in the calculation. Next the differences between the averages of the limiting values and the initial values were determined for each performance question. The results in Table 7.8 show that the average differences for technical ability question are all less than 0.300 which is the lowest average for management questions. Thus, the performance of subcontractors in management questions is improving more than for technical ability questions.

Table 7.8. The averages of subcontractors in performance questions

	Performance Averages for Subcontractors											
Questions	**Q1**		**Q2**		**Q3**		**Q4**		**Q5**		**Q6**	
Matrix	I	L	I	L	I	L	I	L	I	L	I	L
Average	3.722	4.164	3.806	4.013	3.639	3.970	3.750	3.930	3.667	3.985	3.667	3.934
Difference	0.442		0.207		0.332		0.180		0.318		0.268	

L: refers to the limiting matrices, **I**: refers to the initial matrices

Therefore, the results of the performance of subcontractors that worked for multiple contractors show that they are improving over time. Specifically, their improvement in management questions is notably greater than their improvement in technical ability questions.

7.3.2 Updating Time of Performance Historical Records:

The final thing that is examined in this chapter is the ideal time between updates of the historical records of performance for subcontractors. This will be examined through starting with the identity matrix and repeatedly multiplying by the transition matrix for each performance questions as described in Section 7.2. Results for the diagonal elements are presented in Table 7.9. These represent the probability that the subcontractor that started in a particular state is in that state after the given number of years.

The first column in Table 7.9 contains all values equal to 1.000 because the subcontractor must be in that state since they were just measured as being in that state. The second column shows the probability that they will still be in that state the next year. It is noticeable that for all questions the probability that the subcontractors will remain in a particular state drops very quickly. For example, unless the limiting probability for the state is greater than 0.5, or at least close to 0.5, the probability that the subcontractor will still be in the state will be less than 0.5 within two years, mostly within one year. (Note that the value of 0.5 used here is arbitrary. It has simply been selected so that the time for different questions to cross the same threshold can be compared.)

Table 7.9. The probability that a subcontractor will be in the same state after a number of years

Questions	Years										
Q1	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1-3)	1.000	0.676	0.501	0.393	0.324	0.279	0.249	0.229	0.216	0.207	0.202
4	1.000	0.568	0.478	0.459	0.455	0.453	0.453	0.453	0.453	0.453	0.453
5	1.000	0.625	0.501	0.446	0.414	0.394	0.382	0.373	0.368	0.364	0.361
Q2	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1-3)	1.000	0.382	0.315	0.290	0.281	0.277	0.276	0.276	0.276	0.276	0.275
4	1.000	0.468	0.465	0.462	0.461	0.461	0.460	0.460	0.460	0.460	0.460
5	1.000	0.500	0.349	0.294	0.275	0.268	0.266	0.265	0.264	0.264	0.264
Q3	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1-3)	1.000	0.625	0.464	0.381	0.333	0.305	0.288	0.278	0.272	0.268	0.266
4	1.000	0.609	0.526	0.509	0.505	0.504	0.504	0.504	0.503	0.503	0.503
5	1.000	0.579	0.418	0.340	0.297	0.272	0.257	0.248	0.242	0.239	0.237
Q4	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1-3)	1.000	0.412	0.310	0.282	0.275	0.273	0.272	0.272	0.272	0.272	0.272
4	1.000	0.537	0.528	0.526	0.526	0.526	0.526	0.526	0.526	0.526	0.526
5	1.000	0.579	0.418	0.340	0.297	0.272	0.257	0.248	0.242	0.239	0.237
Q5	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1-3)	1.000	0.471	0.367	0.318	0.291	0.276	0.268	0.264	0.261	0.260	0.259
4	1.000	0.529	0.504	0.501	0.499	0.499	0.498	0.498	0.498	0.498	0.498
5	1.000	0.600	0.438	0.352	0.304	0.277	0.262	0.254	0.250	0.247	0.246
Q6	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1-3)	1.000	0.436	0.329	0.306	0.301	0.300	0.299	0.299	0.299	0.299	0.299
4	1.000	0.521	0.469	0.466	0.466	0.467	0.467	0.467	0.467	0.467	0.467
5	1.000	0.444	0.291	0.249	0.238	0.235	0.234	0.234	0.234	0.234	0.234

Two general patterns appear in this data. Firstly the technical questions converge to their limiting values faster than the management questions. Secondly, state 4 takes longer to converge to the limit than the other two states. However, examining the probabilities in Table 7.9 to detect these trends is difficult because each probability has a different end state.

Alternative measures to examine the decreasing value of past performance measurements that become progressively older are to look at the variance of the expected outcomes and to examine the entropy of the information.

The variance for a discrete probability distribution is given by Equation 7.7.

$$\sigma_x^2 = \sum[x^2 \times P(x)] - \mu_x^2 \qquad \text{Equation 7.7}$$

Table 7.10 shows the growth in the variance of the expected outcome of performance for subcontractors starting in each state. The table has been shaded to show where the variance has converged to within 0.001 of the limiting variance. As can been seen in Table 7.10 the variance of the performance of subcontractors that start in sate 4 grows faster than the variance of the performance of subcontractors that start in state (1-3) or in state 5. State 5 in question 6 provides a single exception, which occurs because the variance overshoots the limiting variance and then decreases until it converges with the limiting variance. However, this exception still takes longer to converge to the limiting variance than subcontractors that started in state 4. It is just that it is now converging from above instead of from below. The reason that the variance grows more quickly for subcontractors that start in state 4 is most likely because the performance for subcontractors can either increase or decrease, whereas for the other two states it can only change in one direction (i.e. for state (1-3) a change in performance must be an increase, while for state 5 a change in performance must be a decrease). In all three cases for the same question the limiting variance is the same because in any Markov chain the long run probabilities are the same no matter what state the system starts in. The faster growth of the variance for subcontractors that start in state 4 means that the maximum variance will be reached sooner. This is actually the opposite of what appeared to be happening in Table 7.9. A closer look at Table 7.9 shows that in fact subcontractors that start in state 4 do converge more quickly. It is just that they reach the threshold of 0.5 later because that threshold is closer to the limiting probability than it is for the other states. Table 7.10 makes this clear because all rows converge to the same limiting variance.

Table 7.10. The variances of the data of performance for subcontractors

Questions	Performance Data Variances										
Q1	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1-3)	0.000	0.219	0.433	0.527	0.555	0.557	0.550	0.543	0.536	0.531	0.527
4	0.000	0.407	0.489	0.508	0.514	0.516	0.517	0.518	0.519	0.519	0.519
5	0.000	0.234	0.350	0.419	0.459	0.482	0.496	0.505	0.510	0.513	0.515
Q2	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1-3)	0.000	0.384	0.504	0.529	0.536	0.538	0.539	0.539	0.539	0.539	0.539
4	0.000	0.528	0.535	0.538	0.539	0.539	0.539	0.539	0.539	0.539	0.539
5	0.000	0.484	0.555	0.549	0.544	0.541	0.540	0.540	0.540	0.540	0.539
Q3	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1-3)	0.000	0.234	0.385	0.452	0.479	0.490	0.494	0.495	0.495	0.496	0.496
4	0.000	0.391	0.474	0.491	0.495	0.496	0.496	0.496	0.496	0.496	0.496
5	0.000	0.244	0.388	0.453	0.479	0.490	0.493	0.495	0.495	0.496	0.496
Q4	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1-3)	0.000	0.395	0.459	0.468	0.469	0.470	0.470	0.470	0.470	0.470	0.470
4	0.000	0.460	0.470	0.470	0.470	0.470	0.470	0.470	0.470	0.470	0.470
5	0.000	0.415	0.470	0.471	0.470	0.470	0.470	0.470	0.470	0.470	0.470
Q5	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1-3)	0.000	0.249	0.401	0.461	0.484	0.493	0.498	0.500	0.501	0.501	0.502
4	0.000	0.464	0.493	0.498	0.500	0.501	0.501	0.502	0.502	0.502	0.502
5	0.000	0.240	0.440	0.496	0.508	0.508	0.506	0.505	0.503	0.503	0.502
Q6	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1-3)	0.000	0.384	0.486	0.516	0.525	0.528	0.529	0.529	0.529	0.529	0.529
4	0.000	0.479	0.528	0.530	0.529	0.529	0.529	0.529	0.529	0.529	0.529
5	0.000	0.617	0.570	0.541	0.532	0.530	0.529	0.529	0.529	0.529	0.529

More importantly is the observation that the technical performance questions converge to the limiting variance faster than the management performance questions. It can be seen that the technical questions (2, 4 and 6) all reach the limiting variance (at least for state 4) before any of the management questions (1, 3 and 5) do. This is similar to the pattern observed in Table 7.9.

The entropy measures the number of ways that a system can be arranged, where higher entropy means that there are more ways that the system can be arranged, which results in a greater likelihood that the system is disordered. Entropy here specifically refers to Shannon entropy. This is the expected value (average) of the information in a message. The message in this case is data about the expected performance of subcontractors. Entropy is defined as the negative logarithm of the probability distribution of possible events or messages. In this case the possible events are specific performance evaluations. It is given by Equation 7.8.

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log p(x_i) \qquad \text{Equation 7.8}$$

Entropy is zero when an outcome is certain. The more uncertain the outcome is the higher the entropy is. In this case if the most recent performance measurement of a subcontractor is sufficiently recent that we can predict with high probability what the performance is then the entropy is low. If the performance measurements are outdated then they have little predictive power and so the entropy is high. Since the performance is being modelled as a Markov chain the entropy will increase until the system reaches the long term steady state probabilities where it will remain until new data is collected.

The results of the entropy in Table 7.11 for all states of performance questions show similar patterns to the results of the variance. Subcontractors that start in state 4 still converge faster than subcontractors that start in the other two states. Also there seems to be a slight trend that subcontractors that start in state (1-3) are more likely to take longer to converge to the limiting entropy than subcontractors that start in state 5 for the management questions. This is probably due to the fact that there is a general increase in performance over time in the Markov chain and so subcontractors that start in state 5 have already reached the higher performance level.

Table 7.11. Measuring the Entropy of the past performance data for subcontractors.

Questions	**Past performance Data Entropy**										
Q1	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1-3)	0.000	0.909	1.352	1.498	1.542	1.548	1.541	1.532	1.524	1.518	1.513
4	0.000	1.375	1.474	1.491	1.497	1.500	1.501	1.502	1.503	1.503	1.503
5	0.000	0.954	1.238	1.359	1.421	1.454	1.473	1.485	1.491	1.496	1.499
Q2	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1-3)	0.000	1.325	1.498	1.525	1.532	1.534	1.535	1.535	1.535	1.535	1.535
4	0.000	1.523	1.531	1.533	1.534	1.535	1.535	1.535	1.535	1.535	1.535
5	0.000	1.406	1.536	1.541	1.538	1.536	1.535	1.535	1.535	1.535	1.535
Q3	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1-3)	0.000	0.954	1.305	1.423	1.468	1.484	1.491	1.493	1.494	1.495	1.495
4	0.000	1.357	1.472	1.491	1.494	1.495	1.495	1.495	1.495	1.495	1.495
5	0.000	0.982	1.323	1.434	1.474	1.488	1.493	1.495	1.495	1.495	1.495
Q4	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1-3)	0.000	1.336	1.448	1.462	1.464	1.465	1.465	1.465	1.465	1.465	1.465
4	0.000	1.454	1.466	1.465	1.465	1.465	1.465	1.465	1.465	1.465	1.465
5	0.000	1.379	1.467	1.468	1.466	1.466	1.465	1.465	1.465	1.465	1.465
Q5	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1-3)	0.000	0.998	1.356	1.448	1.480	1.492	1.497	1.499	1.501	1.501	1.501
4	0.000	1.459	1.492	1.497	1.500	1.501	1.501	1.502	1.502	1.502	1.502
5	0.000	0.971	1.389	1.482	1.504	1.507	1.506	1.504	1.503	1.503	1.502
Q6	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1-3)	0.000	1.312	1.476	1.511	1.521	1.523	1.524	1.524	1.524	1.524	1.524

4	0.000	1.477	1.524	1.526	1.525	1.524	1.524	1.524	1.524	1.524	1.524
5	0.000	1.530	1.556	1.535	1.527	1.525	1.524	1.524	1.524	1.524	1.524

The second observation, that technical questions converge faster than management questions, is also evident in Table 7.11. This is similar to the pattern that was found in both Table 7.9 and Table 7.10.

The overall result for all of the above mathematical measurements is that the performance level for the technical questions varies much more than the performance level of the management questions. Therefore decisions of how often to update should be based on technical performance rather than on management performance. The data shows that the variance and entropy grows very quickly and so these should be updated at least annually.

7.4 Conclusion:

In conclusion, this chapter studies the change in performance in six different questions for subcontractors that worked for multiple contractors. The data was analysed using Markov chain models to find out how the performance of subcontractors varies over time in these six questions and also to find out if there are any common patterns in these changes. The results show that the performance of subcontractors in management questions (1, 3 and 5) are more stable and so slower to change compared to their performance in technical questions (2, 4 and 6). The other result about the change in performance for the subcontractors that were studied is that their performance is improving over time in all performance questions.

The other point this chapter examines is the updating time of the historical records about the performance of subcontractors. Interview results show that companies have a wide variety of actual updating times for their records ranging from 1 to 5 years. In order to examine the speed at which information regarding performance becomes stale calculations were performed to determine the probability that a subcontractor would still be in the same state as it had been measured to be in previously. Initial calculations involved directly examining these probabilities. However, while patterns seemed to emerge it was difficult to validate these patterns. Therefore examination was made of the growth in the variance of the performance outcome and in the growth of the information entropy of the performance outcome. Variance refers to the expectation of the squared deviation from the mean. Entropy is the expected value (average) of the information in a message. Both of these showed that the technical questions varied more quickly than the management questions. The other pattern that emerged from both analyses is that subcontractors that start in state 4 more quickly reach a maximum entropy state where previous data becomes completely worthless than subcontractors that start in the other two states. The results also show that the historical records of subcontractors in general lose their information content quickly and should be updated at least annually.

The overall results drawn in this chapter is that more attention is required in tracking the change in performance for subcontractors in technical questions because they have more variability than their performance in management questions. This is may be the result of hiring labour by subcontractors for individual projects, which was found to be one issue encountered by contractors and clients with subcontractors in Saudi Arabia. More investigation is required to test the validity of this hypothesis.

Chapter 8 Analysing the Performance of Subcontractors using Hidden Markov Models

8.1 Introduction:

According to the results in chapter 6, the data that is suitable to be tested using Markov chain and Hidden Markov Models is data set 2, which is about subcontractors that worked for multiple contractors. This data is tested using Markov chain in chapter 6 and the results show that the performance of subcontractors that worked for multiple contractors are expected to be improving over time. Moreover, subcontractors are improving gradually in management questions of performance and significantly in technical questions.

All measurements are a combination of a measurement of the underlying phenomena and noise. Better measurement techniques are able to produce measurements with less noise. In this chapter, the data will be tested using hidden Markov model via the Baum Welch algorithm. The purpose of this is to separate any underlying trend in the performance of the subcontractors from the noise.

8.2 Baum Welch algorithm:

The Baum Welch algorithm is a method for determining the parameters of a discrete hidden Markov models. It consists of an iterative algorithm procedure that starts from an initial point, which contains an estimated set of parameters values, to build a sequence of re-estimates that improves the likelihood of data, (Merialdo, 1993). The hidden Markov model is a statistical model that was proposed by Baum and Petrie in 1966 to be used in describing a Markov Process that has a hidden parameter, (Zhang, 2007). (Eddy, 1996) states that a hidden Markov model “is a finite model that describes a probability distribution over an infinite number of possible sequences.” The assumption of a hidden Markov model is that there is another hidden sequence of data that exists behind the observed sequence of data where this hidden sequence occupies a series of states, (Zhang et al., 2014). There are three problems involved with hidden Markov models (Rabiner, 1989). The first problem is about evaluating the probability of that a particular observed sequence results from a particular model. The second problem is the learning problem where the most likely parameters of the model are estimated based on a specific set of observations. The third involves the prediction of the state sequence given the hidden Markov model and the sequence of observations. These three problems are linked to their corresponding algorithms by (Zhang et al., 2014), see Figure 8.1.

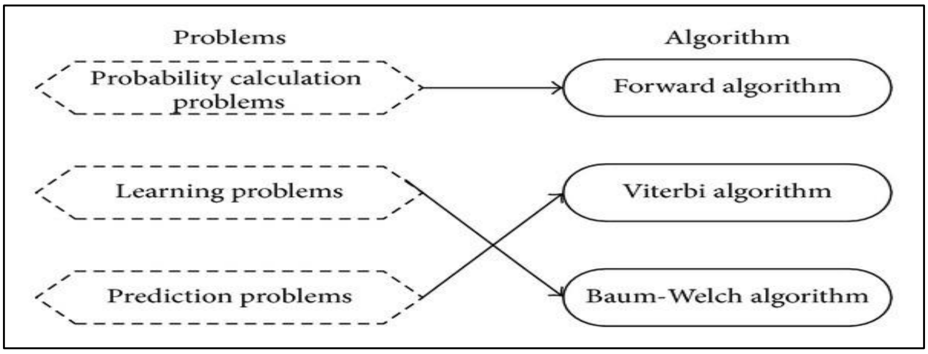

Figure 8.1: The algorithms for the three problems of hidden Markov models, (Zhang et al., 2014)

8.3 Justification for Using Baum Welch Algorithm:

This section will answer the question of choosing Baum Welch over Viterbi algorithm which is commonly chose in most research in maximising the best state sequence of hidden Markov models. Rodríguez and Torres (2003) conducted an important comparative study between Baum Welch algorithm and Viterbi training showing the advantages and disadvantages of each method. They states re-estimation of the hidden Markov models (HMM) is usually conducted in compliance with Maximum Likelihood

Estimation criterion (MLE) by maximising the probability of the training samples with respect to the model. Moreover, this to be conducted by using the Expectation Maximisation algorithm (EM) which works by maximising the log likelihood from incomplete data by maximising the expectation of log likelihood iteratively from complete data set which is done by Baum Welch Algorithm. On the other hand, they stated that the approximation of MLE by maximising the probability of the best hidden Markov model state sequence known as segmental k-means is usually done to be calculated using since it has less computational effort than Baum Welch does. Viterbi training provides the same or slightly worse performance. However, Baum Welch algorithm has three competent properties compared to Viterbi. Firstly, in discrete HMMs Baum Welch needs only non-zero random values to verify the stochastic constraints, so, does not need any model initialisation whereas Viterbi training need some reasonable initialisation either by using data obtained from another models or to train initial models on a handle-labelled subset of the trained database. Secondly, Baum Welch can have suitable initialisation for continuous HMMs by using the output distribution parameters of discrete HMMs and the means and variance which are obtained by the vector quantisation. Thirdly, Baum Welch algorithm uses all of the available data to generate a solid and optimal estimates of a given HMM state. In contrast, re-estimating the parameters of the HMM state results in a less solid models because Viterbi training uses only observations inside the segments corresponding to a given HMM state. As a result, the segmentation of Viterbi training will not precisely match the right one that is generated by an expert. Viterbi is able to avoid this limitation when the size of the data is large enough so the error in segmentation will cancel each other which will result in having more relevant features to identify an HMM state. The use of Viterbi is preferred over Baum Welch when the transition probabilities for the hidden part of the model and the emission probabilities of the visible part are known.

Since the aim of this chapter is that to remove noise from the predictions of the performance of subcontractors that worked for multiple contractors using hidden Markov models, there is a need to have a good estimate of the parameters so that the efficiency of the calculations could be improved. Taking into consideration that the determination of the observation is critically important to assure the efficiency of hidden Markov models, the use of Baum Welch algorithm is significant. Moreover, Baum Welch algorithm will solve the learning problem that is if the selected observation is inappropriate the result of that will affect the time of training the HMM which may also be not completed. Therefore, Baum Welch algorithm will help in improving the results of HMM where the parameters are estimated by providing an observation sequence and then generating a new hidden Markov model for the propose of detection, (Zhang et al., 2014).

Based on the properties of Baum Welch and Viterbi that are discussed above; and to achieve the aim of this chapter, Baum Welch is suitable and chosen to be implemented to the size of the data that will be tested in this chapter

8.4 Methodology:

To implement Baum Welch algorithm, there three main components need to be used and their values are arbitrary according to Moss (2008). These three components are:

1- The first component is the starting probabilities of the states.
2- The second component is the transition probabilities for all performance questions

3- The third component is the observations for all performance questions with the number of occurrence of each observation, see Table 8.1. To make the reading of this table easier, the performance assessments (1-3), 4 and 5, where (1-3) is the lowest and 5 is the highest, will be replaced as follows:
"**A**" replacing 5, "**B**" replacing 4 and "C" replacing (1-3) assessments.

So, based on the number of forms in Table 8.1 there are 145 observations found in all performance questions where the technical questions have the highest numbers of observations which are not less than 26 out of 27 compared to the maximum number of observations found in all questions whereas the management questions have not more than 23 observations. Table 8.2 shows the results of investigating the observations in performance questions has identified some unique forms that each performance question has. It can be seen that the technical question, specifically questions 2 and 6, have more unique forms of observation than what management questions do.

Table 8.1. the forms of the observation in each performance questions for subcontractors that worked for multiple contractors

No.	Q1		Q2		Q3		Q4		Q5		Q6	
	Observation	No. of occurrence	Observation	No. of occurrence	Observation	No. of occurrence	Observation	No. of occurrence	Observation	No. of occurrence	Observation	No. of occurrence
1	AAB	1	CBBB	2	CCBC	2	CBBB	4	CBBB	2	CBBB	5
2	BAB	1	CAAB	1	BABB	2	BAAB	1	BABB	2	BAAB	1
3	BBB	1	BBAB	3	BAAB	2	BBAB	3	BAAB	2	BBAB	1
4	CCBC	1	BCBC	1	CBBB	2	BBBB	1	BBBB	4	CCBC	4
5	BABB	1	ABAC	1	CCCC	5	CCBC	4	CCBC	3	CBAB	1
6	BAAB	1	CCBB	2	CBCB	1	BBAC	1	CBCB	2	BBAC	1
7	BBBB	4	CCBA	2	CCCB	2	CCBB	1	CCCC	2	CCBB	1
8	ABBA	1	CBBA	1	BBBB	5	CCCC	2	CCCB	2	CCCC	2
9	CBBB	2	CCBC	2	CCBB	2	BBCC	1	CCBB	1	BBCC	1
10	BAAA	3	CBCC	2	CBBC	1	CCBA	1	CBBC	1	CCBA	1
11	CBCB	2	BBCC	1	CBCC	2	BCBB	2	CBCC	2	BCBB	2
12	CCCC	5	CCCC	1	BABA	1	BCBA	1	BABA	1	BCBA	1
13	CCCB	1	BBAA	1	BBAB	1	BCAC	1	BBAB	1	CCAC	1
14	CCBB	2	BCBB	1	BCBC	1	ABBB	1	BCBB	1	BBBB	1
15	BCCC	2	BCAC	1	BAAA	1	BBBC	1	BCBC	1	BBBC	1
16	BBBA	1	ABBC	1	AAA	1	CBCB	1	BBAB	1	CBCB	1
17	AABA	1	BBAC	1	ABB	1	BBCA	1	BAAA	2	BBCA	1
18	BBAB	1	BBCA	1	BBCB	1	CBCA	1	AABB	1	CBCA	1
19	BABA	1	CBCB	2	AAB	1	BBBA	1	BBC	2	BBBA	1
20	AABB	1	BBBA	1	AAAA	1	ABAB	1	AAA	1	ABAA	1
21	BCCB	1	ABBB	2	BBAA	1	BBB	1	BBCB	1	ABBA	1
22	AAAA	1	AAA	2			ABA	1	AAAA	1	CCB	1
23	CBAA	1	BCCB	1			BCCB	1			ABA	1
24			AAB	1			AAB	1			ACCC	1
25			AAAA	1			AAAA	1			AAC	1
26			BAAA	1			BBAA	1			AAAA	1
27											BAAA	1
Total		36		36		36		36		36		36

Table 8.2. The number of the unique observations in performance questions

Performance Question	No. of Unique Observations	Observation Forms
Q1. Project Management (planning, organisation and follow up).	4	BAB, BCCC, AABA, CBAA
Q2. Work Quality and Compliance with Specifications	4	CAAB, ABAC, CBBA, ABBC
Q3. Compliance with Time Schedule	1	AAB
Q4. Project Staff Level	1	ABAB
Q5. Availability of Equipment and required Systems	1	BBC
Q6. Applications of Security and Safety Procedures.	6	CBAB, CCAC, ABAA, CCB, ACCC, AAC

After providing these three main components and then discussing the different forms of observations in performance questions, the Baum Welch algorithm will be implemented taking into consideration that it will find the local maximum for the parameters of the observaions $\lambda^* = argmax_0 P(Y|\theta)$ where the random initial conditions set is $\lambda = (A, B, \pi)$, where A, B and π are the three probability measures. More details of the following equations used in implementing Baum welch algorithm are all provided by Rabiner (1989). The implementation will be as follows:

1- The first step is to calculate the forward procedure (α) as given in Equation 4.1 where the intilaisation of the forward is given in part 1 of the equation and the induction is given in part 2 of the equation.

1- $\alpha_1(i) = \pi_i b_i(O_1), 1 \leq i \leq N.$

2- $\alpha_{t+1}(j) = \left[\sum_{j=1}^{N} \alpha_t(j) a_{ij}\right] b_j(O_t + 1), 1 \leq t \leq T - 1$

Equation 8.1

Where $\alpha_1(i) = P(O_1 = O_1, \dots, O_t = o_t, X_t = i|\lambda)$ which is the probability of observing O_1, O2, …, O_t and being in state *i* at time *t*. N is the possible number of states

This step is to be applied on each observation form for each performance question then to get summation of the total of each observation form.

1- 2- The second step is to apply the backward procedure (β) as given in Equation 8.2 where the initialisation of the backward is given in part 1 of the equation and the induction is given in part 2 of the equation.

1- $\beta_T(i) = 1, 1 \le i \le N$

2- $\beta_t(i) = \sum_{j=1}^{N} a_{ij} b_j(O_{t+1})\, \beta_{t+1}(j)$

Equation 8.2

Where $\beta_t(i) = P(O_{t+1}, O_{t+1}, \dots, O_T|q_t = i, \lambda)$ which is the probability of the ending partial sequence O_{t+1}, …, O_T given starting state *i* at time *t*.

This step will be restricted to $1 \le j \le 4$ for the observation sequences that form 4 digits and restricted to the condition $1 \le j \le 3$ for the observation sequences that form 3 digits.

3- The third step is to update the hidden Markov models to estimate the parameters using the results of the forward and backward procedures by calculating the temporary variables according to Bayes' theorem explained as follows:

a- The probability of being in state *i* at time *t* where the observed sequence O and the parametersλ are given, see Equation 8.3

$$\gamma_t(i) = P(q_t = i|O, \lambda) = \frac{\alpha_t(i)\beta_t(i)}{P(O|\lambda)}$$

Equation 8.3

$$= \frac{\alpha_t(i)\beta_t(i)}{\sum_{i=1}^{N} \alpha_t(j)\beta_t(i)}$$

Where the normalisation factor $P(O|\lambda) = \sum_{i=1}^{N} \alpha_t(i) \, , \beta_t i$ makes $\gamma_t(i)$ a probability measure given by Equation 8.4

$$\sum_{i=1}^{N} \gamma_t(i) = 1 \qquad \text{Equation 8.4}$$

b- The state sequence has to be then retrieved using Equation 8.5

$$\delta_t(i) = \pi_i b_i(O_i), 1 \leq i \leq N \qquad \text{Equation 8.5}$$

c- The probability of being in state *i* and *j* at time *t* and *t*+1 where the observed sequence O and the parameters λ are given, see Equation 8.6

$$\xi_t(i,j) = \frac{P(q_t = i, q_{t+1} = j|O,\lambda)}{P(O|\lambda)}$$

$$= \frac{\alpha_t(i) a_{ij} b_j(O_t + 1)\beta_{t+1}(j)}{\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_t(i) a_{ij} b_j(O_t + 1)\beta_{t+1}(j)} \qquad \text{Equation 8.6}$$

Where $P(q_t = i, q_{t+1} = j|O,\lambda)$ is the numerator term divided by $P(O|\lambda)$ to get the desired probability.

d- The parameters O can after that be updated where the expected frequency that is spent in state *i* at time 1 is given in Equation 5.1 :

$$\overline{\pi}_i = \gamma_1(i) \qquad \text{Equation 8.7}$$

Then, the expected number of transitions from state *i* to state *j* in comparison to the expected number of transitions away from state *i* is to be calculated using Equation 8.8

$$\overline{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(ij)}{\sum_{t=1}^{T-1} \gamma_t(i)} \qquad \text{Equation 8.8}$$

Where $\sum_{j=1}^{N} \overline{a}_{ij} = 1, 1 \leq i \leq N$

After that, the expected number of times that the output of observations have been equal to v_k while in state i over the expected total number of times in state i to be calculated using Equation 8.9

$$\overline{b}_j(k) = \frac{\sum_{t=1}^{T} 1_{O_t=v_k} \gamma_t(j)}{\sum_{t=1}^{T} \gamma_t(j)} \qquad \text{Equation 8.9}$$

Where k is the number of values of the observation O and the the indicator function is given as:

$$1_{O_t=v_k} = \begin{cases} 1 \; if \; O_t = v_k \\ 0 \; otherwise \end{cases}$$

Therefore, after applying all of the above steps from 1 to 3, they will be repeated iteratively until the desired convergence is achieved.

To assure the optimal results of implementing Baum Welch algorithm, the number of the hidden states has to be identified so the right model is then analysed. Identification of the number of the hidden states will be as follows:

1- Different numbers of hidden states will be tested using Baum Welch algorithm, which are: 1, 2, 3 and 4 hidden states.

2- The log likelihoods for each performance question in each set of hidden states will be generated depending on the first step and then Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) will be implemented on the log likelihoods of each performance question in each set of hidden states to find out the lowest value of each method. AIC and BIC are penalised likelihood criteria, Zucchini et al. (2016). The BIC and AIC equation are derived from Zucchini et al. (2016) as follow:

$$AIC = -2\ Log\ L + 2\ p^2 \qquad \text{Equation 8.10}$$

$$BIC = -2\ Log\ L + p^2\ \ln(n) \qquad \text{Equation 8.11}$$

Where:

$Log\ L$= Log Likelihood of each performance question.

p = number of hidden states.

n = number of different observed categories.

3- The lowest value will help in deciding what the suitable number of the hidden states is and then to starts discussing the results of the Baum Welch algorithm.

In order to ensure that the model is correct a cross validation process is used. The observations are divided into three groups randomly. The observations consist of three subcontractors with 3 rankings each and 33 subcontractors with 4 rankings each. Thus the three subcontractors with 3 ranking were each put in separate groups. 11 of the 33 subcontractors with 4 rankings each were randomly selected to add to each group. This random selection was carried out by assigning a randomly generated number to each of the 33 subcontractors and then placing them in the groups based on this random number.

Each of the three groups was then used separately as a training data set for the Baum Welch algorithm to determine the starting, transition and emission probabilities. In each case the remaining two groups were combined to create a testing data set. The log likelihood of the model created from the training data set was then determined based on the corresponding testing data set. These log likelihoods were then used to determine the AIC and BIC. The average of the three AICs and BICs were then taken for determining what the optimum number of hidden states is.

The full data set was then analysed using the Baum Welch algorithm to get the most accurate possible starting, transition and emission probabilities to use in the discussion that follows.

8.5 Results:

As stated in the methodology section, the Baum Welch algorithm is used to determine the starting probabilities, the transition probabilities and the emission probabilities. However, even before this can occur it is necessary to select the number of hidden states.

It was initially decided to confirm the number of hidden states by conducting a cross validation for the data by dividing the data into thirds where one third would be used to train Baum Welch and the resulting probabilities would be then be tested using the other two thirds as the test data. This would be performed three times so that each third took a turn as the training data. However, problems with cross validation emerged because some of the transitions in observed states that occurred in the testing data did not occur in the training data. In these cases it became impossible for the probabilities determined in the training stage to generate the testing data. Once the forward algorithm that determines the log likelihood encounters a probability of zero it is unable to continue.

Other fractions were also used to divide the data into training and testing sets, but this problem continually emerged. This problem only affected a small proportion of the results, but it was enough to prevent cross validation from giving a definitive result. The results that were obtained involved AIC generally indicating that the optimum number of hidden states is 2 and BIC indicating that the optimum number of hidden states is 3.

Since cross validation presented problems it was decided to use the whole data set for both training and testing the model so that definitive results could be obtained and then compared with the patchy results obtained from cross validation. The results of the Log Likelihood, AIC and BIC for each performance question using between 1 and 4 hidden states are shown in Table 8.3. The results of AIC and BIC are also shown in Figure 8.2 and Figure 8.3.

Table 8.3. The results of the log likelihoods, BIC and AIC of each performance question for subcontractors that worked for multiple contractors

Questions	Q1	Q2	Q3	Q4	Q5	Q6
H1						
Log Likelihood	-149.403	-150.697	-144.462	-142.879	-143.353	-148.068
BIC	299.905	302.492	290.021	286.856	287.804	297.234
AIC	300.807	303.393	290.923	287.758	288.706	298.136
H2						
Log Likelihood	-118.584	-142.806	-115.380	-135.116	-118.689	-140.798
BIC	241.563	290.006	235.154	274.625	241.772	285.991
AIC	245.169	293.612	238.760	278.231	245.378	289.597
H3						
Log Likelihood	-112.826	-140.425	-111.096	-132.267	-115.909	-135.597
BIC	235.540	290.737	232.080	274.422	241.706	281.080
AIC	243.653	298.850	240.193	282.535	249.819	289.193
H4						
Log Likelihood	-111.093	-137.954	-108.601	-129.116	-114.691	-132.657
BIC	239.762	293.485	234.780	275.808	246.958	282.891
AIC	254.185	307.907	249.202	290.231	261.381	297.314

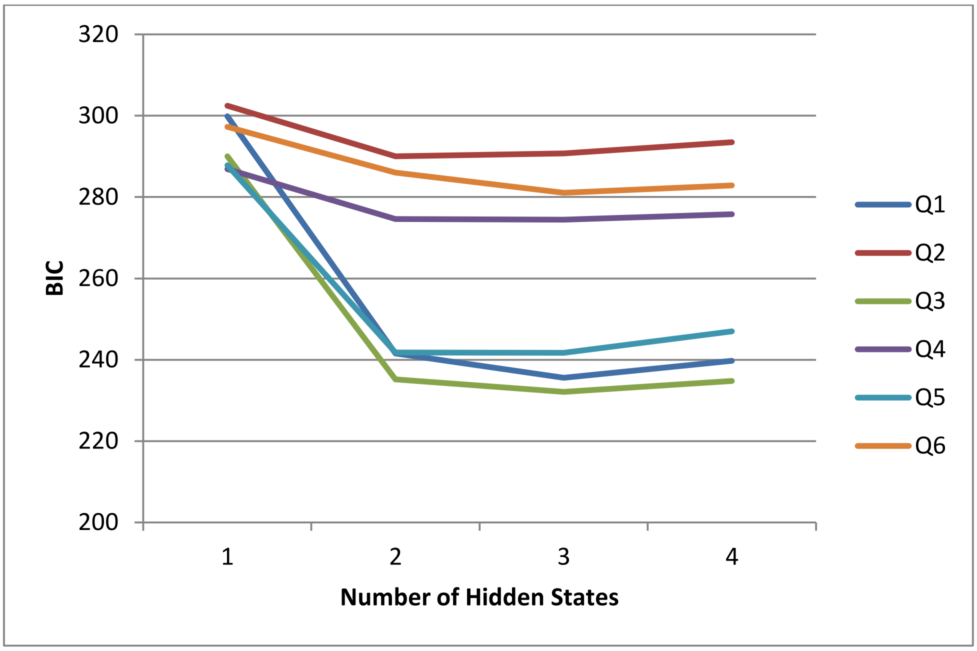

Figure 8.2. The BIC values for different number of hidden values for performance questions

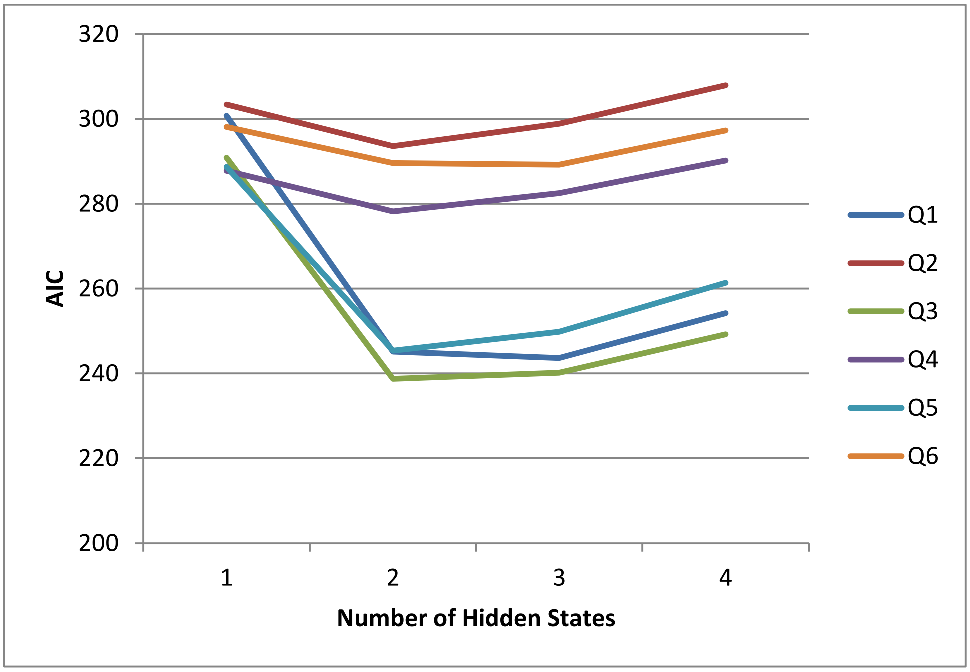

Figure 8.3. The AIC values for different number of hidden values for performance questions

Firstly it can be seen that using two hidden states is always better than using one hidden state. This result holds for both AIC and BIC. The importance of the result that two hidden states is better than one hidden state is that one hidden state is not actually a hidden Markov model. Instead it is equivalent to the observed states being randomly observed and having no relationship with previous observations of the same subcontractor. If one hidden state was the best result then there would be no point at all in tracking the performance of subcontractors. However, the results show that one hidden state is always outperformed by two hidden states.

The next most obvious thing to observe is that while all questions are clustered together for the one hidden state model there is a large separation in likelihoods between the management questions and the technical questions for hidden Markov models with more than one state, i.e. real hidden Markov models. This would tend to indicate that past data on management performance is much more useful in predicting future performance than past data on technical performance.

As can be seen in Table 8.3 and Figure 8.3, the using two hidden states results in the lowest AIC values, except for questions 1 and 6 where three hidden states give a lower AIC. Similarly Table 8.3 and Figure 8.2 show that using three hidden states results in the lowest BIC values, except for question 2 where two hidden states give a lower BIC. Therefore, the results of Baum Welch Algorithm will be discussed based on both two hidden states and three hidden states.

The results of the starting probabilities, transition probabilities and emission probabilities for the hidden Markov models with two hidden states are shown in Table 8.4, Table 8.5 and Table 8.6. The results for the hidden Markov models with three hidden states are shown in Table 8.7, Table 8.8 and Table 8.9.

Table 8.4. The results of starting probabilities for Baum Welch algorithm for two hidden states

Questions	Q1	Q2	Q3	Q4	Q5	Q6
S1	0.504	0.372	0.458	0.354	0.394	0.205
S2	0.496	0.628	0.542	0.646	0.606	0.795

Table 8.5. The results of transition probabilities for Baum Welch algorithm for two hidden states

Q1	S1	S2
S1	1.000	0.000
S2	0.113	0.887

Q4	S1	S2
S1	1.000	0.000
S2	0.156	0.844

Q2	S1	S2
S1	0.905	0.095
S2	0.174	0.826

Q5	S1	S2
S1	1.000	0.000
S2	0.000	1.000

Q3	S1	S2
S1	1.000	0.000
S2	0.000	1.000

Q6	S1	S2
S1	1.000	0.000
S2	0.070	0.930

Table 8.6. The results of emission probabilities for Baum Welch algorithm for two hidden states

Q1	A	B	C		Q4	A	B	C
S1	0.393	0.607	0.000		S1	0.354	0.578	0.068
S2	0.000	0.246	0.754		S2	0.018	0.450	0.532
Q2	**A**	**B**	**C**		**Q5**	**A**	**B**	**C**
S1	0.504	0.458	0.039		S1	0.448	0.552	0.000
S2	0.019	0.424	0.556		S2	0.000	0.472	0.528
Q3	**A**	**B**	**C**		**Q6**	**A**	**B**	**C**
S1	0.382	0.618	0.000		S1	0.554	0.394	0.051
S2	0.000	0.347	0.653		S2	0.066	0.457	0.478

Table 8.7. The results of starting probabilities for Baum Welch algorithm for three hidden states

SP	Q1	Q2	Q3	Q4	Q5	Q6
S1	0.309	0.269	0.359	0.285	0.083	0.246
S2	0.203	0.201	0.131	0.000	0.329	0.356
S3	0.488	0.530	0.511	0.715	0.587	0.398

Table 8.8. The results of transition probabilities for Baum Welch algorithm for three hidden states

Q1	S1	S2	S3
S1	1.000	0.000	0.000
S2	0.261	0.739	0.000
S3	0.000	0.164	0.836

Q4	S1	S2	S3
S1	1.000	0.000	0.000
S2	0.000	1.000	0.000
S3	0.000	0.446	0.554

Q2	S1	S2	S3
S1	0.645	0.355	0.000
S2	0.489	0.499	0.012
S3	0.000	0.238	0.762

Q5	S1	S2	S3
S1	0.600	0.400	0.000
S2	0.424	0.576	0.000
S3	0.000	0.000	1.000

Q3	S1	S2	S3
S1	1.000	0.000	0.000
S2	0.000	1.000	0.000
S3	0.000	0.121	0.879

Q6	S1	S2	S3
S1	0.822	0.000	0.178
S2	0.219	0.781	0.000
S3	0.000	0.622	0.378

Table 8.9. The results of emission probabilities for Baum Welch algorithm for two hidden states

Q1	A	B	C
S1	0.581	0.419	0.000
S2	0.000	1.000	0.000
S3	0.000	0.181	0.819

Q4	A	B	C
S1	0.446	0.554	0.000
S2	0.155	0.530	0.315
S3	0.000	0.461	0.539

Q2	A	B	C
S1	0.782	0.000	0.218
S2	0.000	1.000	0.000
S3	0.000	0.340	0.660

Q5	A	B	C
S1	1.000	0.000	0.000
S2	0.000	1.000	0.000
S3	0.000	0.455	0.545

Q3	A	B	C
S1	0.493	0.507	0.000
S2	0.000	0.917	0.083
S3	0.000	0.207	0.793

Q6	A	B	C
S1	0.629	0.371	0.000
S2	0.000	0.668	0.332
S3	0.000	0.000	1.000

8.6 Discussion

The following discussion will cover the implications of the probabilities derived above for both the two hidden state models and the three hidden state models. Firstly the emission probabilities matrix will be discussed since that is the one that relates to the actual observations and hence can aid in interpreting the meaning of each hidden state. Next the transition probabilities matrix will be discussed as this controls the evolution of the system. Finally the starting probabilities will be discussed. The discussion will also cover the difference between the behaviours of the management and technical questions in relation to their hidden states.

8.6.1 Two Hidden States

The hidden Markov models with two hidden states divide the subcontractors up into high and low performing subcontractors. The following discussion will explain this in more detail.

8.6.1.1 Emission Probabilities

Examination of Table 8.6 shows that for all of the performance questions hidden state 1 is most likely to emit A or B, while hidden state 2 is most likely to emit B or C. Thus it is seen that subcontractors in hidden state 1 generally perform better than subcontractors in state 2.

Note: the fact that all 6 questions gave the same pattern is deliberate. Changing the initial conditions provided to the Baum Welch algorithm can result in the hidden states being reordered. That is, hidden state 2 would be the better performing state than hidden state 1 under different starting assumptions for the probabilities. However, all of the probabilities would be the same; they would just be positioned differently in the relevant matrices consistent with a relabelling of the hidden states. This was tested and confirmed. To simplify the discussion initial assumptions were selected that would result in hidden state 1 always being the high performing state rather than hidden state 2.

Thus subcontractors in hidden state 1 will be referred to as high performance while subcontractors in hidden state 2 will be referred to as low performance. The emission matrix therefore says that high performance subcontractors will normally receive ratings of A or B, while low performance subcontractors will normally receive ratings of B or C. This is absolute for management questions: high performance subcontractors never receive C and low performance subcontractors never receive A. However, for the technical questions it is not quite absolute, but for all questions the probability of a high performance subcontractor receiving a C or a low performance subcontractor receiving an A is always less than 7%.

8.6.1.2 Transition Probabilities

The transition matrices in the hidden Markov models are very diagonally dominant. In fact, for questions 3 and 5 the transition matrices are identity matrices, so for these questions the performance of a subcontractor will never change. For questions 1, 4 and 6 the probability of moving from hidden state 1 to state hidden 2 is zero, while for the opposite direction it is still very low (less than 20%). So for these questions subcontractors will only ever improve. Technical performance question 2 is the only one where subcontractors drop from the high performance state, hidden state 1, to the low performance state, hidden state 2, and even this has a probability of less than 10% (compared to 17% for the opposite transition).

In terms of distinguishing between management and technical questions, the management questions (1, 3 and 5) are the ones with minimal or no change (two identity matrices and one "improve only" matrix). The technical questions (2, 4 and 6) involve greater probabilities of changing (two "improve only" matrices and one "bi-directional" matrix). This is similar to the pattern noted in Chapter 6 that changes in management performance happen more slowly than changes in technical performance.

8.6.1.3 Relation between Emission and Transition Probabilities Matrices and the Original Data

For the management questions (questions 1, 3 and 5) almost all subcontractors either only received As and Bs or only received Bs and Cs for any given question. The only exception is that one of the subcontractors received CBAA for management question 1. This is why the Baum Welch Algorithm produced models for questions 3 and 5 where the transition matrices are identity matrices (i.e. no transitions possible) and emission matrices where one hidden state only emits As and Bs and the other hidden state only emits Bs and Cs. For these two questions each subcontractor is in one of the hidden states or the other and never changes. This does not mean that their performance level never changes; just that it does not change enough for it to jump from one hidden state to the other.

The CBAA result mentioned above for one of the subcontractors in question 1 does not fit this pattern. Therefore, the transition matrix for management question 1 has a small probability that a subcontractor in hidden state 2 (where they emit the initial C) moves to hidden state 1 (where they emit the later As). It is possible to fit the data for question 1 with an identity matrix for the transition matrix by changing one of the zero entries in the emission matrix to be non-zero. However, the Baum Welch Algorithm found that the likelihood of this was less than the likelihood of the results given in Table 8.4, Table 8.5 and Table 8.6.

The original data in Table 8.1 for the technical questions (2, 4 and 6) is quite different. For each of these questions there are subcontractors that received an A rating, with a C rating coming later; and vice versa. Thus none of the technical questions have transition matrices that are identity matrices. However, the transition matrices generally have a structure that involves improvement: the probability of improving from hidden state 2, which indicates low performance, to hidden state 1, which indicates high performance, is always bigger than the opposite probability, which was zero in two cases. The subcontractors that received C ratings after A ratings would then be represented by the small probability of being in hidden state 1, but emitting a C.

8.6.1.4 Starting Probabilities

The results of the updated starting probabilities show that subcontractors usually start in hidden state 2 as low performing except for performance question 1 where they are close to being equal. This further supports the idea that the hidden Markov model has a structure that involves subcontractors generally improving over time.

If the starting probabilities had been such that most subcontractors started in the high performance hidden state then the emission and transition matrices that were found would instead represent a system that was highly stable at the high performance level, with any drops in performance level being quickly rectified.

However, the high probability of starting in the low performance hidden state 2 would be expected in an industry where many subcontractors are formed, improve over time, and later withdraw.

One problem with the two hidden state model is that the sequences of subcontractor ratings are sufficiently short that the sequences show very little transition from the low performing state to the high performing state. The subcontractors may be improving, but this improvement is not large enough to make a clear transition between the two states, particularly in the management questions. Therefore the next section will discuss the three hidden state models. These have more resolution and so more clearly indicate the improvements over even such a short time.

8.6.2 Three Hidden States

The hidden Markov models with three hidden states divide the subcontractors up into high, medium and low performing subcontractors. The following discussion will explain this in more detail.

8.6.2.1 Emission Probabilities

Generally looking at the emission probabilities it can be seen that subcontractors in hidden state 1 usually emit A and B, except for question 2, where hidden state 1 emits A or C. Similarly subcontractors in hidden state 3 always emit B or C and never emit A. In fact, for question 6, hidden state 3 only emits C. Subcontractors in hidden state 2 are most likely to emit B. in fact, for questions 1, 2, and 5 they only emit B.

Note: the fact that all questions have the same pattern of hidden state 1 having the highest performing subcontractors and state 3 having the lowest is deliberate for the same reasons discussed above for the two hidden state models.

Thus subcontractors in hidden state 1 will be referred to as high performance, subcontractors in hidden state 2 will be referred to as medium performance, and subcontractors in hidden state 3 will be referred to as low performance.

In the two hidden state models each hidden state always output more than one rating. However, in the three hidden state models there are 5 cases where a hidden state always outputs the same rating (i.e. the probability is 1.0). Three of these 5 cases involve hidden state 2, that is, the medium performance state.

Hidden state 1 for question 2 breaks the pattern of the highest performing hidden state only emitting As and Bs. Instead it emits As and Cs. This is probably a result of the individual observation sequences being fairly short. It is noticeable that the transition probabilities between states 1 and 2 for this question are high compared to most other of diagonal probabilities, such that sequences like BBAB involve hidden state changes when the rating changes for question 2, whereas when it occurred for questions 1, 3 and 4 it would more likely involve the subcontractor being in hidden state 1 the whole time.

8.6.2.2 Transition Probabilities

The results of the transition probabilities show the performance of subcontractors is expected to have very high probability to stay in the same hidden state between consecutive ratings. Once again, for most questions the matrix is diagonally dominant. There are two exceptions. The first one is hidden state 3 for performance question 6 where subcontractors are more likely to move to hidden state 2 than stay in hidden state 3. The second exception is hidden state 2 for question 2 where the probability of staying in hidden state 2 is slightly less than the probability of changing, but still greater than the probability of changing to any particular other hidden state. On the other hand, the probability of staying in the same state reaches 100% in some performance questions: hidden state 1 for question 1, hidden state 3 for question 6 and both hidden states 1 and 2 for questions 3 and 4. However, the 6 out of 18 times that this occurs for the three hidden state models is a much smaller proportion than the 7 out of 12 times that it occurs in the two hidden state models. Thus since the width of the states is smaller for the three hidden state models subcontractors are more likely to make a transition than occurred in the two hidden state models.

Generally all of the performance questions show transition matrices involving subcontractors that improve over time. Clearly for state 3 the only way is up. However, subcontractors can stay in this state, but there is only one question (question 5) where subcontractors do not ever improve from state 3. Subcontractors that leave state 2 (medium performance) almost always move to state 1 (high performance) except for question 2 where there is a 0.012 probability of dropping to hidden state 3. For questions 1, 3 and 4 subcontractors that reach hidden state 1 always stay there. For the other questions they are more likely to stay in hidden state 1 then to drop down to hidden state 2.

For questions 1 hidden state 1 is an absorbing state. Over time all subcontractors will move into this state and stay there. For questions 3 and 4 both hidden state 1 and hidden state 2 are absorbing states. Over time all subcontractors in hidden state 3 will move to hidden sate 2 and remain there.

Comparing management questions to technical questions it appears that improvement in technical questions is easier than management questions. However, this is not as clear cut as in the two hidden state models. In the three hidden state models the technical and management question improvement probabilities tend to overlap.

8.6.2.3 Starting Probabilities

The results of the starting probabilities show that subcontractors are most likely to start in hidden state 3, the lowest performing state, for all questions. This lends further support to the idea that the hidden Markov model has a structure that involves subcontractors generally improving over time, similar to the two state hidden Markov models.

8.7 Updating Time of Performance Historical Records with Baum Welch Algorithm:

The final thing that is examined in this chapter is the ideal time between updates of the historical records of performance for subcontractors. As was done in the previous chapter this will be examined through starting with the identity matrix and repeatedly multiplying by the transition matrix for each performance questions as was described in Section 7.3.2 in chapter 77.2. Results for the diagonal elements are presented in Table 8.10. These represent the probability that the subcontractor that started in a particular hidden state is in that hidden state after the given number of years.

The first column in Table 8.10 contains all values equal to 1.000 because the subcontractor must be in that state since they were just measured as being in that state. The second column shows the probability that they will still be in that state the next year. (Note that the value of 0.5 used here is arbitrary. It has simply been selected so that the time for different questions to cross the same threshold can be compared.)

Two general patterns appear in this data. Firstly subcontractors tend to improve away from state 3 (low performance) much faster in technical questions than in management questions. Secondly, hidden state 2 (medium performance) takes longer to converge to the limit than the other two hidden states. However, examining the probabilities in Table 8.10 to detect these trends is difficult because each probability has a different end state.

Alternative measures to examine the decreasing value of past performance measurements that become progressively older are to look at the variance of the expected outcomes and to examine the entropy of the information.

The variance for a discrete probability distribution is given by Equation 7.7 in previous chapter.

Table 8.10. The probability that a subcontractor will be in the same state after a number of years

Questions	**Years**										
Q1	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
3	1.000	0.836	0.699	0.584	0.488	0.408	0.341	0.285	0.239	0.199	0.167
2	1.000	0.739	0.547	0.404	0.299	0.221	0.163	0.121	0.089	0.066	0.049
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Q2	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
3	1.000	0.762	0.584	0.449	0.346	0.268	0.209	0.164	0.129	0.103	0.084
2	1.000	0.645	0.590	0.579	0.575	0.573	0.572	0.571	0.570	0.569	0.569
1	1.000	0.499	0.425	0.414	0.413	0.412	0.412	0.412	0.412	0.412	0.412
Q3	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
3	1.000	0.879	0.773	0.680	0.598	0.526	0.463	0.407	0.358	0.315	0.277
2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Q4	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
3	1.000	0.554	0.307	0.170	0.095	0.052	0.029	0.016	0.009	0.005	0.003
2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Q5	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	1.000	0.576	0.501	0.488	0.486	0.485	0.485	0.485	0.485	0.485	0.485
1	1.000	0.600	0.530	0.517	0.515	0.515	0.515	0.515	0.515	0.515	0.515
Q6	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
3	1.000	0.143	0.078	0.077	0.094	0.110	0.122	0.129	0.133	0.135	0.135
2	1.000	0.611	0.501	0.439	0.408	0.394	0.388	0.386	0.386	0.386	0.386
1	1.000	0.676	0.580	0.525	0.497	0.483	0.478	0.476	0.476	0.476	0.476

Table 8.11 shows the growth in the variance of the expected outcome of performance

for subcontractors starting in each state. The table has been shaded to show where the variance has converged to within 0.001 of the limiting variance. For most questions it takes longer for the hidden Markov model to converge than for the Markov chain model to converge in Table 7.10.

Table 8.11. The variances of the data of performance for subcontractors based on the results of Baum Welch algorithm

Questions	Performance Data Variances													
Q1	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
3	0.000	0.137	0.311	0.470	0.589	0.664	0.698	0.698	0.674	0.633	0.582	0.527	0.470	0.415
2	0.000	0.193	0.248	0.241	0.210	0.172	0.137	0.106	0.081	0.062	0.046	0.035	0.026	0.019
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Q2	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
3	0.000	0.725	0.803	0.764	0.696	0.623	0.557	0.499	0.452	0.414	0.383	0.360	0.341	0.327
2	0.000	0.229	0.249	0.257	0.263	0.267	0.270	0.273	0.274	0.276	0.277	0.278	0.278	0.279
1	0.000	0.274	0.273	0.273	0.275	0.276	0.277	0.278	0.278	0.279	0.279	0.279	0.280	0.280
Q3	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
3	0.000	0.106	0.175	0.218	0.240	0.249	0.249	0.241	0.230	0.216	0.200	0.184	0.168	0.153
2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Q4	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
3	0.000	0.247	0.213	0.141	0.086	0.050	0.028	0.016	0.009	0.005	0.003	0.002	0.001	0.000
2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Q5	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	0.244	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250
1	0.000	0.240	0.249	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250
Q6	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
3	0.000	0.235	0.279	0.311	0.359	0.408	0.446	0.470	0.485	0.492	0.495	0.497	0.497	0.497
2	0.000	0.171	0.292	0.380	0.436	0.469	0.485	0.493	0.496	0.497	0.497	0.497	0.497	0.497
1	0.000	0.584	0.675	0.637	0.584	0.545	0.521	0.507	0.501	0.498	0.497	0.497	0.497	0.497

However, much more interesting is that the variances generally tend to increase and then decrease. This is more clearly shown in Figure 8.4. The increase occurs as more and more subcontractors move away from any particular state. However, the decrease is a result of most of the transition matrices having absorbing states. If a transition matrix has only one absorbing state, such as occurs in question 1 where the highest performance hidden state is an absorbing state, then eventually all subcontractors will move into this state and stay there. When this happens the variance becomes zero because all subcontractors are in this one single state. If the subcontractor starts in one of these absorbing states then they will never leave it so the variance in that case is always zero.

The variances for questions 3 and 4 converge to zero more quickly than the variances for question 1 because for these two questions the two highest performing states are absorbing states so subcontractors will always reach an absorbing state after at most one transition.

The entropy measures the number of ways that a system can be arranged, where higher entropy means that there are more ways that the system can be arranged, which results in a greater likelihood that the system is disordered. Entropy here specifically refers to Shannon entropy. This is the expected value (average) of the information in a message. The message in this case is data about the expected performance of subcontractors. Entropy is defined as the negative logarithm of the probability distribution of possible events or messages. In this case the possible events are specific performance evaluations. It is given by Equation 7.8, see previous chapter.

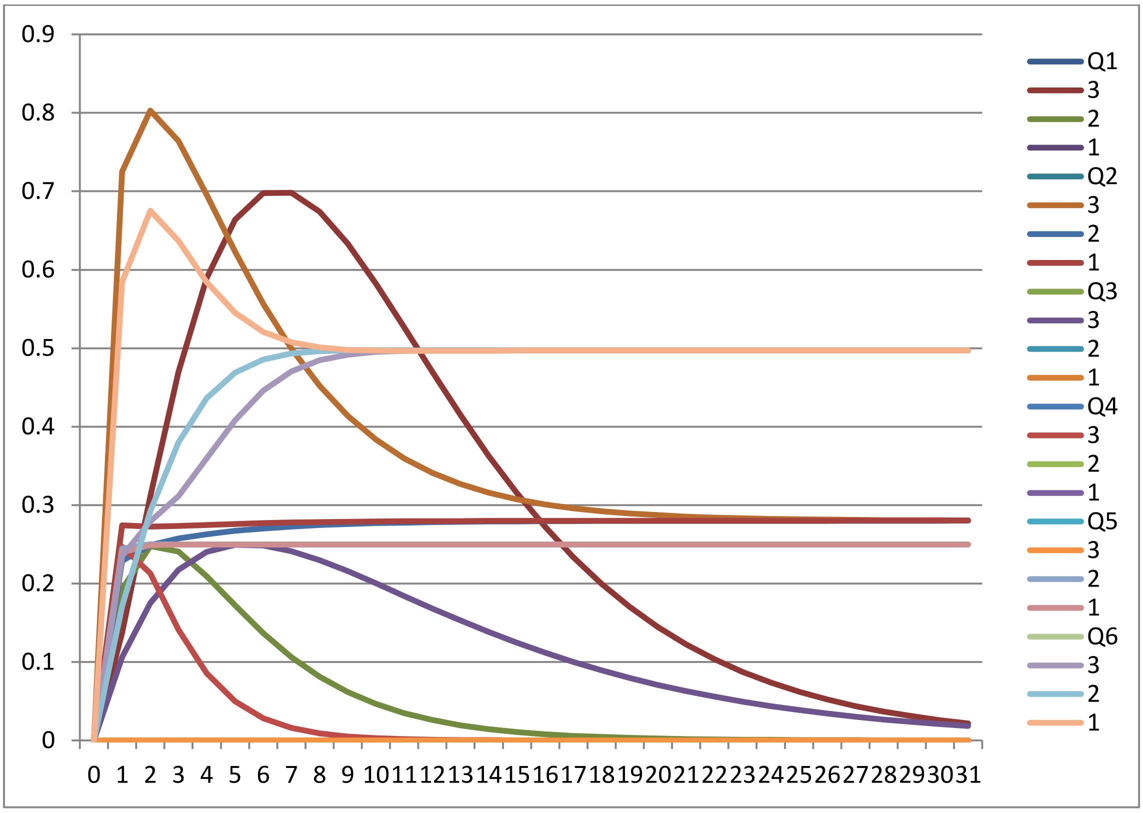

Figure 8.4. The increase and decrease in variance growth in performance levels

Entropy is zero when an outcome is certain. The more uncertain the outcome is the higher the entropy is. In this case if the most recent performance measurement of a subcontractor is sufficiently recent that we can predict with high probability what the performance is then the entropy is low. If the performance measurements are outdated then they have little predictive power and so the entropy is high.

The results of the entropy in Table 8.12 for all states of performance questions show similar patterns to the results of the variance. Once again the hidden Markov models take longer to converge than the Markov chain, see Table 7.11. Similarly the entropy increases as the subcontractors leave their initial states and then decrease if the particular question has one or two absorbing states.

The overall result for all of the above mathematical measurements is that the hidden state representing the performance level is much more stable than the actual observed values. However, while this might indicate that older data is more accurate for this model than for the Markov chain model there is still a necessity to gather enough data so that the hidden state can be determined. Thus there is still a need for frequent data collection.

Table 8.12. Measuring the Entropy of the past performance data for subcontractors based on the results Baum Welch algorithm.

Questions	**Past performance Data Entropy**													
Q1	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
3	0.000	0.644	1.060	1.326	1.486	1.565	1.582	1.553	1.491	1.409	1.313	1.210	1.106	1.002
2	0.000	0.828	0.994	0.973	0.880	0.762	0.642	0.531	0.434	0.351	0.281	0.224	0.177	0.140
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Q2	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
3	0.000	0.791	1.335	1.528	1.582	1.571	1.529	1.475	1.420	1.367	1.320	1.280	1.245	1.217
2	0.000	0.938	1.011	1.040	1.059	1.072	1.081	1.088	1.093	1.097	1.099	1.102	1.103	1.104
1	0.000	1.083	1.086	1.089	1.093	1.097	1.100	1.102	1.103	1.105	1.105	1.106	1.107	1.107
Q3	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
3	0.000	0.531	0.772	0.904	0.972	0.998	0.996	0.975	0.941	0.899	0.851	0.801	0.749	0.698
2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Q4	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
3	0.000	0.991	0.890	0.659	0.451	0.297	0.190	0.119	0.074	0.045	0.027	0.016	0.010	0.006
2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Q5	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	0.983	1.000	1.000	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
1	0.000	0.971	0.997	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
Q6	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
3	0.000	0.957	1.133	1.200	1.286	1.352	1.391	1.411	1.421	1.427	1.429	1.431	1.431	1.431
2	0.000	0.758	1.147	1.309	1.377	1.407	1.420	1.427	1.430	1.431	1.431	1.431	1.431	1.431
1	0.000	0.675	1.208	1.400	1.452	1.455	1.447	1.440	1.435	1.433	1.432	1.431	1.431	1.431

8.8 Conclusion:

In conclusion, this chapter examines the use of hidden Markov models via Baum Welch algorithm to predict the future performance of subcontractors working with multiple contractors. The concept of predicting the future performance of subcontractors using this method require identifying the optimum number of hidden states which could be specified either by using experts opinion about the expected scenarios for future performance of subcontractors or by using some mathematical method to decide the number of the hidden states. In this study, there are three numbers of hidden states were tested: 2, 3 and 4 hidden states. The decision of the number of hidden states was taken by using Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) to find the lowest value out of the three different proposed numbers of hidden states. For the case study used in the chapter the ideal number of hidden states was two or three. This corresponds to the same number as for number of rating levels or less. The AIC and BIC methods are also implemented on one hidden state to assure that the use of hidden Markov models actually did fit the data better than a Markov chain and was not simply just taking advantage of the extra available parameters. In both cases it was clear that the hidden states corresponded to different levels of subcontractor performance. The advantages of favouring the use of three hidden state for forecasting future performance of subcontractors is that two hidden states show minimal transitions because each of the two hidden states covered a wide range of performance.

The results also show that the hidden Markov models have a structure whereby there is generally improvement in subcontractor performance over time. The results of technical performance questions show that the performance of subcontractors is expected to be improving faster than in management questions.

The other point this chapter examines is the updating time of the historical records. This was found to be different to the case for Markov chains.

Markov chain models indicated that it was necessary to update frequently because the historical records became outdated quickly. With hidden Markov models the issue was that a fair amount of data was required to actually identify which hidden state the subcontractor is in. However, once identified this information was valid for longer.

The next chapter will discuss the overall findings of the thesis, including a more detailed comparison of the Markov chain model with the hidden Markov model.

Chapter 9 Discussion

This chapter discusses all the results from the previous chapters to come to an overall result that achieves the aims and explains the contributions of the thesis.

The thesis presents a methodology for analysing the performance of subcontractors. The methodology is tested by applying to a set of case study data obtained from Saudi Arabian construction industry. The case study data was first examined for consistency. Next the validity of the methodology was confirmed through application to the case study data.

9.1 Different performance measures and their effects on assessing the performance of subcontractors

The data collected based on the aims of the thesis can be divided into two types. The first type was exploratory data. This exploratory data was gathered by conducting interviews with different organisations and contractors in Saudi Arabia since it was the source of the case study data. The aim of collecting the exploratory data was to identify the different performance measures used by different organisations and contractors in assessing the performance of subcontractors. Those that were interviewed were selected as they worked for organisations that are known to keep historical records of the performance of their subcontractors. The other major purpose of collecting this data was to compare the performance measures used by these organisations to their selection criteria for subcontractors; and also to compare their performance factors to a common historical performance record.

The results of the first comparison (between performance factors used by different organisations) showed that most of the selection factors used by 5 interviewed organisations that have internal systems were based on previous performance measurements of the subcontractor. These results validate the findings of some previous studies in the literature reviews conducted by Alfeld (1988), Maturana et al. (2007), Derek Lavelle et al. (2007) and Costa and Tavares (2013) about the importance of past performance in selecting subcontractor to work for construction projects. All performance data collected by an organisation for its own purposes would be used in future selection of subcontractors. However, each internal system of assessing the performance of subcontractors was different from one organisation to another, as shown in Chapter 3. These 5 organisations have different numbers of performance factors range from 3 to 6 factors. On the other hand, they mostly agreed on the four most important factors, i.e. compliance with time scheduled, safety adherence, work quality; and the availability of equipment and resources. Each of these four factors was collected by four out of the 5 organisations, although in general it was a different four organisations each time. In line with the later distinction that was found between management performance factors and technical performance factors, two of the four factors fall into each group. It is interesting that the four factors that were considered most important by the collection of 5 organisations are evenly divided between these two groups. Two are management performance factors: compliance with time schedule and availability of equipment and resources, and two are technical performance factors: work quality and application of safety procedures.

The first comparison was based on binary data: did the organisation collect data on a particular performance factor or not? The second comparison involved ranking the specific performance factors used by the governmental records. Out of the four most important factors it was found that availability of equipment was generally ranked below the other three performance factors. The rankings to the performance factors used in the governmental record showed that the performance factors were ranked from the less to the most important as shown in Table 9.1. More details regarding this were discussed in Chapter 4.

Table 9.1: The ranking given to the performance factors used in the governmental record by other in Saudi Arabia

Performance Factors	Mean scores	Ranking
Work Quality and Compliance with the specifications	3.8	1
Application of safety procedures	3.8	1
Compliance with time schedule	3.5	3
Availability of equipment	1.7	4
Project Management (planning, organisation and follow up)	1.6	5
Project Staff Level	1	6

The next issue considered was whether the importance of a factor would affect how critically the contractors would assess that item. It was assumed that if an item was considered important that the contractor would assess this item more critically. On the other hand it was assumed that if an item was not considered important then the assessment would be less critical and so higher scores would be given. This was formulated into the hypothesis that contractors more harshly assess the performance of subcontractors for factors that they consider more important. The results of testing this hypothesis are presented in Chapter 4 and showed that it is accepted. There were strong negative correlations between the importance rankings of each organisation and the mean performance scores given to subcontractors for particular performance questions.

However, correlation does necessarily imply causation. It could be that importance was assigned to different factors based on how much the contractor perceived that each factor needed to be improved across the industry (or at least their small part of it). Thus high importance ranking was given to performance factors that typically only achieved low scores instead of low scores being given for important factors. On the other hand there could be a third thing that causes both high importance and low scores. The negative correlations could be linked to the finding of Meng (2012) which found that poor performance in construction projects resulted from the deterioration between project parties where each has different concerns. This study suggested that more strategic relationship needs to be formed between different project parties to improve the performance. This could be a source of further investigation.

One of the discoveries during the interviews was that 3 out of 5 organisations that have their own internal records do not consider external data sources at all when making subcontractor selections. Thus they are contributing to an external data source, but not using it. This leads to the question of whether the data that they contribute to the governmental records is reliable. This overall result leads to the fact that more investigations are recommended to review the relations between the most important factors used by organisations and contractors and the assessments of the performance of subcontractors.

It is noticeable that the number of KPIs recorded in either the 5 interviewed organisations records or in the government records is much less than the number recommended by various studies in the literature, such as Ali et al. (2013). This result leads to the recommendation for more studies to compare the key performance indicators recommended by the various studies in the literature with the actual ones that are used by organisations and why there is a difference.

9.2 The consistency of different performance ratings given for subcontractors

The strong negative correlations between importance of a performance question and the scores given for performance related to that question by each interviewed organisation led to the concern that performance ratings from different contractors might be inconsistent. This might occur since different organisations had different importance rankings for different performance measures. Thus data was collected where 30 subcontractors were all rated by the same 4 contractors. Chapter 5 presented an analysis of this data to test the consistency between the contractors. The Cronbach's alpha values were all above 0.7 indicating that the consistency was good. In fact, for three of the contractors the alpha value was above 0.9. This indicates excellent consistency.

9.3 Internal structure of the performance data

The next issue tested was to determine if there was any internal structure within the data. Factor analysis was used to test for this. This analysis is also described in Chapter 5.

The result of the factor analysis was that the performance measurements were separated into two groups. The first group was labelled management performance questions and included project management, compliance with time scheduled and the availability of equipment. The second group was labelled technical performance questions and included work quality, project staff Level and safety procedures.

The distinction between these two groups of performance questions was found to be useful in the analyses performed in subsequent chapters where the behaviour of the two groups was often different. For example when modelling the data using Markov chains it was found that the technical performance questions varied more than the management performance questions. This will be discussed below.

9.4 The Fitting of Performance Data to Markov Models

The overall aim of this thesis is to show that hidden Markov models are useful for identifying the performance level of individual subcontractors. To do this it is necessary to show that the data possesses the Markovity property. This means that the only data needed to predict the future performance of subcontractors is their current level of performance. If the data has this property then it fits a first order Markov chain.

However, the question arose of whether subcontractors that worked exclusively for a single contractor would have different results to subcontractors that worked for multiple contractors. This is important to know because if they behave differently then they will not fit the same Markov chain. Therefore these two groups were tested for Markovity separately.

Testing for Markovity requires that each subcontractor has at least three data points. This is because Markovity means that knowing the past does not help with predicting the future once the present is known. Thus, testing the Markovity requires measurements from at least three different times so that there can be a past present and future to compare. The small set of data that contained subcontractors that worked for more than one subcontractor, but for multiple times for any of them, was examined for situations where they had either worked for the same contractor at least three times. Those cases were added to the same contractor group. Cases where a subcontractor had worked for three different contractors were added to the different contractors group.

As a result, the validation was conducted separately on the two data sets of subcontractors worked for the same and multiple contractors. The result would also determine if the two groups could be combined if they were first order Markov chain. The validation results show that Markov chain cannot be accurately implemented on the data for subcontractors that worked for the same contractors whereas the data of subcontractors worked for multiple contractors did fit the model. More details about the nature of the lack of fit, and about reasons why this occurred were discussed in Chapter 6. The reasons can be summarised as that subcontractors were more likely to receive the same rating on consecutive jobs with the same contractors because the contractors had already formed an opinion of their performance level and so it tended to change less than would be expected from modelling the larger data set.

An attempt was made to investigate whether the stability of subcontractors' performance in management questions is related to the type of work performed. Table 9.2 below shows the p-values for the Markovity test for individual subcontractors. Just under half of the results are significant, showing that breaking the data into these small subsets creates a problem of too small a sample size, and therefore the analysis could not be continued. As a result, this also prevents implementing hidden Markov models to analyse the causes of the optimal frequency result found in Chapter 8.

Table 9.2: The results of the Markovity for converting each type of subcontractors to a first order Markov chains after extracting third states of 0 and 2; and removing second states of 1 and 2

Performance Question	**Services**	**Structure**	**Earthmoving**
1	0.0559	0.623	0.0144
2	0.5744	0.4445	0.0482
3	5E-05	0.5018	0.391
4	0.9784	0.0285	0.0326
5	0.0068	0.5407	0.7892
6	0.9828	0.1021	0.0017

Therefore the following chapters only analysed the data obtained from subcontractors that worked for multiple contractors.

9.5 Comparing the results of the performance change of subcontractors using Markov chain and the hidden Markov models

The results of the performance of subcontractors that worked for multiple contractors using Markov chain showed that they are expected to perform better over time in all performance questions. This can be seen by the transition probabilities from lower to higher performance levels being generally higher than the transition probabilities from higher to lower performance levels. Also the limiting probabilities had a higher probability for the subcontractors to be in higher performing states than the initial probabilities.

However, their improvement in the management performance questions was slower to change compared to their performance in technical questions. In other words the management performance was more stable than the technical performance.

Hidden Markov models were tested using two, three and four hidden states. AIC and BIC analysis results showed that two or three hidden states gave the best fit.

Hidden Markov models with two states showed a clear structure of the two states representing high and low performance. These models fit the data well, but tended to obscure the improvement of the subcontractor performance. The models tended to divide the subcontractors into two groups where the high performing group always received one of the top two performance ratings and the low group always received one of the bottom two performance ratings. This was a result of the transition matrices for many of the questions being the identity matrix, which models no change between hidden states. The usefulness of this result is that it shows that the measured performance of subcontractors can be visualised as consisting of two components. Firstly there is the actual performance of the subcontractor. Secondly part of the measurement is an error term. The hidden state represents the underlying performance level, while the changing between emitted performance measurements represents the error in the measurement.

The difficulty with the two hidden states models is that having only two hidden states means that each state covers a wide range of performance levels. Therefore, subcontractors may be changing their performance level but staying within the same hidden state. As a result, the three hidden states model was examined.

Hidden Markov models with three hidden states also showed a clear division of subcontractors into hidden states representing high, medium and low performance levels. These models tended to have transition matrices that showed improvement in the same way that the Markov chains did. Thus, attempting to narrow the range of performance level represented by each hidden state by increasing the number of hidden states was successful. However, for some of the questions the probabilities in transition matrices showed the subcontractors staying in particular hidden states, similar to what was seen in the two hidden states models.

These instances where the probabilities indicated no change in hidden states may have occurred because the sequences of the performance ratings for the individual subcontractors were too short to show sustained improvement. This is not a problem if the aim is simply to determine the current level of the subcontractor. In that case having a short sequence of data is sufficient since it places the subcontractor within a particular hidden state that represents a certain performance level and different subcontractors can be compared.

These results of Markov chains and hidden Markov models lead to the question of how often performance records should be updated.

On the other hand, the common result of analysing the performance of subcontractors that worked for multiple contractors is that their performance is more stable in management performance question compared to their performance in technical performance questions. There are three possible causes of this result, which are: the difficulty of measuring the performance questions in these two different groups, the importance of performance questions in each group and variability in the qualification of workers.

According to Cox et al. (2003), qualitative performance factors, such as management performance factors, are difficult to measure compared to quantitative performance factors, such as technical performance questions. This is because the quantitative indicators can be measured in terms of dollars, units or man-hours, such as cost, on time completion and percent complete. On the other hand, the qualitative indicators are not as reliable for measuring performance because of their perceived difficulty or inability to be measured. Chan and Chan (2004) found the same results where the performance factors were divided into two groups, qualitative and quantitative factors. It states that the quantitative indicators are measured mathematically but the other group is about the qualitative indicators that are hard to measure where asking the stakeholders about their opinion is suggested to resolve this difficulty. The findings of these two studies could help in understanding the difficulties perceived in assessing the management performance questions compared to technical performance questions.

The second possibility relates to the importance of the performance questions. Chapter 4 studies how contractors assess subcontractors in the most important factors. According to the results in Table 4.2, the three factors that were generally assigned the highest importance (time, safety and quality) are output factors. These things are received by the contractors. On the other hand the three factors that were generally assigned the lowest importance (staff level, project management, and equipment availability) are input factors. These are the things that enable the subcontractors to do their work. Thus more importance is attached to factors regarding the actual work of the subcontractors than to the factors that enable them to do work. This increased importance of the technical factors may result in increased scrutiny by the assessors and therefore greater variability as more details are included.

The third possibility is that the skill levels of the workers is highly variable and this is flowing through into the technical question performance results. This appears to be a likely cause since at the time the data was collected there was no centralised scheme for certification of the workers. Since the data was collected such a scheme has been implemented in Saudi Arabia for this purpose. On the other hand, a scheme was in place for certification of engineers and managers in Saudi Arabia at the time data was collected. Therefore the skills related to management questions had been certified and this seems to be the main reason that the management performance is more stable than the technical performance.

9.6 Frequency of updating historical records

The other aim of using Markov chains and hidden models was to determine the optimum frequency for updating historical records of subcontractors based on their performance. This was one of the questions that had been asked of the interviewed organisations that have historical records of subcontractors. The answers ranged from 1 to 5 years based on different considerations that were explained in Chapter 3. In terms of determining the optimum updating time from the collected performance data, the results of Markov chains and hidden Markov models were used and then compared.

To do this it was assumed that a subcontractor started in a particular state and then the probability that it would still be in that state after different amounts of time were determined. Also two methods were used to measure the uncertainty of which state it would be in at any time after starting in a particular state. These two methods were calculation of the variance of the performance level and the Shannon entropy.

For the Markov chains it was found that the probability of remaining in a particular state quickly dropped, especially in the technical performance questions. Similarly the variance of outcome and the entropy of the outcome grew quickly. Based on these results it is recommended that when simply using the measured performance ratings that these ratings need to be updated at least annually since performance levels change quite quickly, especially for technical questions.

On the other hand, with hidden Markov models the probability of leaving each particular state at any given time was much lower. In other words subcontractors would stay in any particular state for longer. Interestingly both the variance and the entropy would initially increase and then later decrease. This did not happen for the Markov chains. The reason for this is that the increase represents the uncertainty of when the subcontractor would leave the state, while the decrease occurs because subcontractors improve their performance over time and thus in the long term future they all converge on the hidden state representing the highest performance level due to improvement. This is not evident in the Markov chains because the level of error in any particular performance measurement is large enough to create a high variance or entropy of outcome.

This means that the age of performance data can be much older when using hidden Markov models then when using a Markov chain. Thus hidden Markov models are better than Markov chains. Simply using the most recent performance data for predicting future performance is equivalent to using a Markov chain.

Thus when predicting future performance, it is necessary to either have recent performance data or to have a sequence of performance data so that the better hidden Markov model can be used.

This application of this methodology to other countries is necessary because it is unknown how the required frequency for updating performance varies between countries. Thus while the method can be replicated the results may be different due to the procedures, culture and other factors that differ between countries

9.7 The Inaccuracies and Sensitivity of the Subjective Performance Scale

One point to be discussed regarding the implementation of the proposed methodology is that the bias encountered in the data collected about the performance of subcontractors in Saudi Arabia. The issue here is that the results of thesis are limited to the case of subcontractors in the construction industry of Saudi Arabia; and also the scale used to assess the performance of subcontractors is subjective. In this regard, the subjectivity in scale used should be discussed in terms of bias and any inaccuracies that might affect the validity of the performance evaluation.

The proposed methodology uses different tests before being implemented. The aim of these tests is to minimize or remove the biases or inaccuracies that affect the applicability of the methodology. Briefly, there are different tests conducted starting by identifying the most important performance factors for different organisations using mean score method to find out the correlations between this result and the performance assessments given by them for the same subcontractors. Following to this test, Cronbach's alpha analysis is also conducted to identify the level of agreement between multiple contractors on the performance ratings given for subcontractors. The idea of conducting these two tests prior to implementing the other analysis methods is that to find out whether the performance ratings are consistent before their use. For instance, if the results of identifying the most important factors have significant differences to the results of consistency analysis it means that the reliability of the performance rating is questionable. The results of the important performance factors show that there are more important performance factors than others. However, the results of Cronbach's alpha analysis show that the level of agreement for the performance ratings given by multiple contractors is high, which means that the most important factors are output factors and the less important performance factors are input factors.

The other test that is used is Kullback's algorithms to validate the Markovity of the performance data. That means the performance data have to fit first order Markov chain to be continue using this methodology. Another analysis that is used to minimise or remove bias from data is hidden Markov models via Baum Welch algorithm using Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) where they help in deciding the optimum number of the hidden states that represents different levels of performance of subcontractors. The number of the hidden states would provide more accurate results to be analysed and discussed as shown in Chapter 8. In addition to hidden Markov models, Shannon entropy analysis is also implemented to examine the uncertainty of the performance data while their frequent updating time is analysed to help in analysing the optimum updating time with regard to variance growth analysis results.

All of these tests helped in minimising the biases or inaccuracies in the data to ensure more valid results are produced. Those tests lead to the conclusion that Markov chains is preferable with most recent data of performance but hidden Markov models is more preferable for older or longer data set of performance.

On the other hand, these results are limited to the case study of Saudi Arabia and it is recommended that this methodology be applied to the performance data of subcontractors in other countries.

Chapter 10 Conclusion

10.1 Research Overview

This thesis introduces a methodology for analysing the performance of subcontractors. The methodology is tested by applying to a set of case study data that was obtained from the construction industry of Saudi Arabia.

Exploratory data was gathered by conducting interviews with different organisations and contractors in Saudi Arabia. This data was gathered to identify three aspects. The first aspect was comparing selection factors used by contractors where most of them use past information to assess the performance of subcontractors. The other aspect involved identifying the most common performance factors between the interviewed organisations and the performance factors used in the governmental performance records. The results of this comparison was expected to provide a base for comparing the actual key performance indicators (KPIs) with a study stating 10 KPIs identified in Saudi Arabia. The third aspect involved comparing the ranking given by these organisations of the most important factors used in the government records and their actual assessments to test the hypothesis that they were critical in assessing the performance of subcontractors in the most important factors.

Historical performance data was collected for testing the consistency of ratings given by different contractors and the consistency was found to be good. In fact for most of the contractors testing the consistency between them was excellent. This is a very important result because it tells contractors that they can trust the performance ratings of other contractors. This data was also tested to determine the internal structure within the six performance questions used in the case study of Saudi Arabia. This result was helpful in analysing the performance of subcontractors.

The historical performance data was collected in two sets and were examined to confirm the validity of implementing Markov chains on them. That is, whether they could be modelled as first order Markov chains. Subcontractors that had worked for multiple contractors were tested separately to the subcontractors that had each worked for a single contractor to determine if the two data sets could be combined. However, subcontractors that had each only worked for a single contractor did not fit the first order Markovity condition. Subcontractors that worked for multiple contractors did fit this condition. Therefore, only subcontractors that worked for multiple contractors were analysed using Markov chains and hidden Markov models.

The performance of subcontractors that worked for multiple contractors was analysed using Markov chains to provide a baseline of results using transition matrices directly generated from the raw data. The two previously identified groups of performance questions were then discussed in terms of their differences in behaviour.

Their performance was also analysed using hidden Markov models, utilising the Baum Welch algorithm to determine the emission matrices, transition matrices and starting probabilities. The emission matrices were then used to interpret the meaning of the underlying the hidden states.

The results of analysing the performance of subcontractors using Markov chians and hidden Markov models were used to determine the optimum frequency for updating their historical records based on their performance. This is important because the interviewed organisations had widely divergent opinions on how often this frequency should be.

10.2 Main Research Findings

- The initial exploratory research of the Saudi Arabian case study showed that Saudi construction organisations are interested in 8 different selection factors that require past information to assess the performance of subcontractors. Five of the 8 different performance factors were similar to the 6 performance question used in the governmental performance records. Four of these performance factors were used by 4 of the 5 interviewed organisations: compliance with time scheduled, safety adherence, work quality; and the availability of equipment and resources. Ranking of the performance measures used in the governmental records by the 5 organisations indicated that compliance with time scheduled, work quality and safety were all ranked as highly important, with the others being generally considered unimportant. Comparing the ranking given by each organisation with their actual assessments of performance found that the importance of a performance factor affects how critically the contractor assesses subcontractors in that factor. Results showed that contractors more critically assess subcontractors in the most important performance factors by rating them lower while rating subcontractors higher in the less important factors.
- Comparing the result of the most important performance factors used by 5 organisations in Saudi Arabia with the results of the study that was conducted by Ali et al. (2013) showed differences. Their result showed that 10 key performance indicators should be tracked, while in actuality the number of performance factors used in the governmental records was only 6 and those used by the interviewed organisations averaged 4.25. This result will be discussed in the following section on future research and recommendations.

- The results of the consistency of performance ratings given by different contractors indicate that their assessments are trustworthy. This result is important because three of the interviewed organisations stated that they do not use performance ratings given by other contractors or organisations.
- The result of determining the internal structure of the performance questions used in the governmental records in Saudi Arabia showed they involved two different groups. The first group was labelled management performance questions and included project management, compliance with time scheduled and the availability of equipment. The second group was labelled technical performance questions and included work quality, project staff Level and safety procedures. This result showed that the 4 most important performance factors identified by the interviewed organisation were evenly divided between these two groups. This structure was found useful when discussing the results of the different performance questions and determining the optimal frequency of updating the performance records.
- Part of the proposed methodology is to ensure that the model is valid. The case study data was tested for validity of applying Markov chains using Kullback's algorithm. This showed that the data that involved subcontractors that worked for the same contractors did not fit the Markovity condition. Interpretations for the lack of fit were then proposed. The lack of fit can be summarised as that subcontractors were more likely to receive the same rating on consecutive jobs with the same contractors because the contractors had already formed an opinion of their performance level and so it tended to change less than would be expected from modelling the larger data set. This methodology can be extended

to other data sets. Since only the data involving subcontractors that worked for multiple contractors fit the Markovity condition only this data was analysed using Markov chain and hidden Markov models.

- The results of applying Markov chains to the data showed that the performance of subcontractors is generally improving over time. However, improvement in the management performance questions was slower compared to improvement in technical performance questions. Examination of the growth of the variance in performance and its Shannon entropy showed that the performance of subcontractors in management questions is more stable than in technical questions.
- The most important result of the thesis is that when the hidden Markov models were applied to the case study data the hidden states did in fact represent different performance levels. This was clearly shown through the structure of the emission matrices.
- The transition matrices generated showed a similar trend of improvement in performance over time to the Markov chains. However, the actual results were more stable and due to the short lengths of the sequences analysed improvement was not always indicated, especially in the two hidden state models. This is not a problem if the aim of the exercise is to determine the current level of performance, since the hidden Markov models indicated that this performance level was quite persistent.
- Comparison of the Markov chain results with the hidden Markov models highlighted the difference in this persistence. Using Markov chains records should ideally be a maximum of one year old. However, the greater stability of

the hidden Markov models indicated that an equivalent reliability of results could still be obtained for two year old records, or even longer for management performance.

10.3 Statement of Contributions and Research Novelty

The contributions of this research involve introducing a new methodology to track the performance of subcontractors using different analysis methods compared to those stated in the literature and showing that this methodology works by applying to case study data. This methodology involves the following:

1 Testing the consistency of different ratings of performance by different contractors using Cronbach’s alpha.

2 Validating that the model is appropriate for the data by testing for the Markovity condition.

3 Applying the hidden Markov model to the data.

4 Determining the number of hidden states to use in the model using AIC and BIC.

5 Examining the emission matrix to ensure that the hidden states represent different performance levels. The hidden states can be reordered to make the connection between performance level and hidden states clearer.

6 The growth in the variance of the performance level arising from repeated application of the transition matrix and its Shannon entropy can be used to determine an appropriate trade-off between updating frequency and accuracy.

In terms of practical contribution of assisting with the selection of subcontractors the research does not specify how this should be done. Instead it gives clear indication of how long data on performance is valid for so that erroneous decisions are not made based on outdated data.

10.4 Limitations and Future Directions

This research involved some limitations that need to be addressed. These limitations are described as follows.

10.4.1 Limitations

The first limitation is that the methodology has only been applied to one case study. Secondly, the sequences of subcontractor performance records were not very long for individual subcontractors since the governmental records have a limit on how long data is retained.

10.4.2 Recommendations

Several recommendations have been derived from the limitations and the results of this research.

Investigate why researchers that recommend which KPIs need to be tracked seem to have longer lists (e.g. (Ali et al., 2013)) than are used in the systems actually used by organisations. Is this because the researchers interviewed people who are not actually dealing with the performance assessments as a system?

The methodology introduced in this research needs to be validated by examining the performance of subcontractors in different countries to test the generality of this method.

Subcontractors that worked for a single contractor were found to fail the Markovity condition. One reason for a subcontractor to restrict its customers to a single contractor could be the presence of a strategic relationship. More investigation into the effect of strategic partnerships on performance improvement is required using this methodology.

More studies are recommended to be conducted involving longer sequences of performance records to more accurately include the improvement of subcontractors over time. This will allow a distinction to be made between how frequent the data needs to be updated during the sequence vs how long the results are valid after the last record in the sequence

Bibliography

ALFELD, L.E. (1988) Construction Productivity. (New York, McGraw-Hill).

AKSORN, T. & HADIKUSUMO, B. 2008. Critical success factors influencing safety program performance in Thai construction projects. *Safety Science,* 46**,** 709-727.

ALENCAR, L. & DE ALMEIDA, A. Selecting subcontractors in projects using a multicriteria group decision model. Industrial Engineering and Engineering Management (IEEM), 2010 IEEE International Conference on, 2010. IEEE, 1563-1567.

ALI, H. A. E. M., AL-SULAIHI, I. A. & AL-GAHTANI, K. S. 2013. Indicators for measuring performance of building construction companies in Kingdom of Saudi Arabia. *Journal of King Saud University-Engineering Sciences,* 25**,** 125-134.

ARSLAN, G., KIVRAK, S., BIRGONUL, M. T. & DIKMEN, I. 2008. Improving sub-contractor selection process in construction projects: Web-based sub-contractor evaluation system (WEBSES). *Automation in Construction,* 17**,** 480-488.

ARTTO, K., ELORANTA, K. & KUJALA, J. 2008. Subcontractors' business relationships as risk sources in project networks. *International Journal of managing projects in business,* 1**,** 88-105.

ATKINSON, R. 1999. Project management: cost, time and quality, two best guesses and a phenomenon, its time to accept other success criteria. *International journal of project management,* 17**,** 337-342.

AUTRY, C. W. & GOLICIC, S. L. 2010. Evaluating buyer–supplier relationship–performance spirals: A longitudinal study. *Journal of Operations Management,* 28**,** 87-100.

BASSIONI, H. A., PRICE, A. D. & HASSAN, T. M. 2004. Theoretical formulation of a framework for measuring business performance in construction.

BRÉMAUD, P. 1999. *Markov Chains*, Springer-Verlag New York.

CHAN, A. P. & CHAN, A. P. 2004. Key performance indicators for measuring construction success. *Benchmarking: an international journal,* 11**,** 203-221.

CHAN, A. P., CHAN, D. W., CHIANG, Y. H., TANG, B.-S., CHAN, E. H. & HO, K. S. 2004. Exploring critical success factors for partnering in construction projects. *Journal of Construction Engineering and Management,* 130**,** 188-198.

CHAN, D. W. & KUMARASWAMY, M. M. 1997. A comparative study of causes of time overruns in Hong Kong construction projects. *International Journal of project management,* 15**,** 55-63.

CHENG, M.-Y., TSAI, H.-C. & SUDJONO, E. 2011. Evaluating subcontractor performance using evolutionary fuzzy hybrid neural network. *International Journal of Project Management,* 29**,** 349-356.

CHING, W.-K. & NG, M. K. 2006. *Markov Chains: Models, Algorithms and Applications* [Online]. US: Springer. [Accessed 83].

CHUA, D. K. H., KOG, Y.-C. & LOH, P. K. 1999. Critical success factors for different project objectives. *Journal of construction engineering and management,* 125**,** 142-150.

COOKE-DAVIES, T. 2002. The "real" success factors on projects. *International journal of project management,* 20**,** 185-190.

COOKE, B. & WILLIAMS, P. 2013. *Construction planning, programming and control*, John Wiley & Sons.

COSTA, A. A. & TAVARES, L. V. 2013. Advanced multicriteria models to promote quality and reputation in public construction e-marketplaces. *Automation in Construction,* 30**,** 205-215.

COX, R. F., ISSA, R. R. A. & AHRENS, D. 2003. Management's perception of key performance indicators for construction. *Journal of Construction Engineering and Management-Asce,* 129**,** 142-151.

DEPUTY MINISTRY FOR CONTRACTOR CLASSIFICATION IN SAUDI ARABIA. 2015. *Classifications Forms for Contractors* [Online]. Available: https://contractors.momra.gov.sa/#CForms [Accessed 31 March 2015].

DEREK LAVELLE, JASON HENDRY & STEEL, G. 2007. THE SELECTION OF SUBCONTRACTORS: IS PRICE THE MAJOR FACTOR? *23rd Annual ARCOM Conference,*.

DULUNG, A. Z. A. & PHENG, L. S. 2005. Factors Influencing the Selection of Subcontractors in Refurbishment Works. *Architectural Science Review,* 48**,** 93-103.

DURBIN, R., EDDY, S. R., KROGH, A. & MITCHISON, G. 1998. *Biological sequence analysis: probabilistic models of proteins and nucleic acids*, Cambridge university press.

EDDY, S. R. 1996. Hidden markov models. *Current opinion in structural biology,* 6**,** 361-365.

EMUZE, F. & JULIAN SMALLWOOD, J. 2014. Collaborative working in South African construction: contractors' perspectives. *Journal of Engineering, Design and Technology,* 12**,** 294-306.

EOM, C. S. J., YUN, H. & PAEK, J. H. 2008. Subcontractor Evaluation and Management Framework for Strategic Partnering. *Journal of Construction Engineering and Management,* 134**,** 10.

EOM, S.-J., KIM, S.-C. & JANG, W.-S. 2015. Paradigm shift in main contractor-subcontractor partnerships with an e-procurement framework. *KSCE Journal of Civil Engineering,* 19**,** 1951.

FAMAKIN, I., AJE, I. & OGUNSEMI, D. 2012. Assessment of success factors for joint venture construction projects in Nigeria. *Journal of Financial Management of Property and Construction,* 17**,** 153-165.

GAMBS, S., KILLIJIAN, M. O., & DEL PRADO CORTEZ, M. N. (2012, April). Next place prediction using mobility markov chains. In *Proceedings of the First Workshop on Measurement, Privacy, and Mobility* (p. 3). ACM

GEORGE, D. & MALLERY, M. 2003. Using SPSS for Windows step by step: a simple guide and reference.

GHAHRAMANI, Z. 2001. An introduction to hidden Markov models and Bayesian networks. *International journal of pattern recognition and artificial intelligence,* 15**,** 9-42.

HAPONAVA, T. & AL-JIBOURI, S. 2009. Identifying key performance indicators for use in control of pre-project stage process in construction. *International Journal of Productivity and Performance*

Management, 58**,** 160-173.

HARTMANN, A. & CAERTELING, J. 2010. Subcontractor procurement in construction: the interplay of price and trust. *Supply Chain Management: An International Journal,* 15**,** 354-362.

HORTA, I. M., CAMANHO, A. S. & DA COSTA, J. M. 2009. Performance assessment of construction companies integrating key performance indicators and data envelopment analysis. *Journal of Construction Engineering and Management,* 136**,** 581-594.

HWANG, B.-G. & LIM, E.-S. J. 2012. Critical success factors for key project players and objectives: Case study of Singapore. *Journal of Construction Engineering and Management,* 139**,** 204-215.

KEMENY, J. G. & SNELL, J. L. 1960. *Finite markov chains,* Van Nostrand.

KIM, S.-Y. & HUYNH, T.-A. 2008. Improving project management performance of large contractors using benchmarking approach. *International Journal of Project Management,* 26**,** 758-769.

KONSTANTOPOULOS, T. 2009. Markov Chains and Random Walks. Uppsala Universitet.

KULLBACK, S., KUPPERMAN, M. & KU, H. 1962. Tests for Contingency Tables and Marltov Chains. *Technometrics,* 4**,** 573-608.

LAU, E. & ROWLINSON, S. 2009. Interpersonal trust and inter-firm trust in construction projects. *Construction Management and Economics,* 27**,** 539-554.

LEE, J. K., HAN, S. H., JANG, W. & JUNG, W. 2017. “Win-win strategy” for sustainable relationship between general contractors and subcontractors in international construction projects. *KSCE Journal of Civil Engineering***,** 1-12.

LEVINSON, S. E., RABINER, L. R. & SONDHI, M. M. 1983. An introduction to the application of the theory of probabilistic functions of a Markov process to automatic speech recognition. *The Bell System Technical Journal,* 62**,** 1035-1074.

LI, B., AKINTOYE, A., EDWARDS, P. J. & HARDCASTLE, C. 2005. Critical success factors for PPP/PFI projects in the UK construction industry. *Construction management and economics,* 23**,** 459-471.

LU, W., SHEN, L. & YAM, M. C. 2008. Critical success factors for competitiveness of contractors: China study. *Journal of construction engineering and management,* 134**,** 972-982.

MATURANA, S., ALARCÓN, L. F., GAZMURI, P. & VRSALOVIC, M. 2007. On-site subcontractor evaluation method based on lean principles and partnering practices. *Journal of Management in Engineering,* 23**,** 67-74.

MATHEW, W., RAPOSO, R., & MARTINS, B. (2012, September). Predicting future locations with hidden Markov models. In *Proceedings of the 2012 ACM conference on ubiquitous computing* (pp. 911-918). ACM

MENG, X. 2012. The effect of relationship management on project performance in construction. *International journal of project management,* 30**,** 188-198.

MOSS, L. 2008. Example of the Baum-Welch Algorithm. *Q520, Spring*.

NG, S. T. 2007. Using Balanced Scorecard for Subcontractor Performance Appraisal.

NG, S. T. & TANG, Z. 2010. Labour-intensive construction sub-contractors: Their critical success factors. *International Journal of Project Management,* 28**,** 732-740.

NG, S. T., TANG, Z. & PALANEESWARAN, E. 2009. Factors contributing to the success of equipment-intensive subcontractors in construction. *International Journal of Project Management,* 27**,** 736-744.

NG, T. S., SKITMORE, M. & CHUNG, W. F. Ten basic factors to identify suitable subcontractors for construction projects. Proceedings of the CIB TG 23 International Conference on Professionalism in Construction: Culture of High Performance, 2003. CIB-International Council for Research and Innovation in Building and Construction, 1-8.

NORUSIS, M. 1993. SPSS for Windows Professional Statistics Release 6.0. SPSS. *Inc, Chicago*.

OAK, Y. N. 2012. *The Effects of Uncertainty and Reputation in Sub-Contracting Networks.*

OCHIENG, E. G. & PRICE, A. 2010. Managing cross-cultural communication in multicultural construction project teams: The case of Kenya and UK. *International Journal of Project Management,* 28**,** 449-460.

OGUNLANA, S. O. & TOOR, S.-U.-R. 2010. Beyond the 'iron triangle': Stakeholder perception of key performance indicators (KPIs) for large-scale public sector development projects. *International journal of project management,* 28**,** 228-236.

OLUWOYE, J. O., MACLENNAN, H. M. & CAROLAN, C. G. 1996. MODELLING BUILDING SUB-CONTRACTOR PERFORMANCE, A CASE STUDY OF THE NSW CONSTRUCTION INDUSTRY. *CIB TG 23 International Conference.* Hong Kong.

RABINER, L. R. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE,* 77**,** 257-286.

RADUJKOVIĆ, M., VUKOMANOVIĆ, M. & DUNOVIĆ, I. B. 2010. Application of key performance indicators in South-Eastern European construction. *Journal of Civil Engineering and Management,* 16**,** 521-530.

RASHID, K. M., DATTA, S., & BEHZADAN, A. H. (2017, December). Coupling risk attitude and motion data mining in a preemtive construction safety framework. In *Simulation Conference (WSC), 2017 Winter* (pp. 2413-2424). IEEE.

RIORDAN, K. 2014. Properties of Markov Chains. June 2014 ed. US: Governers State University.

RODRÍGUEZ, L. & TORRES, I. 2003. Comparative study of the baum-welch and viterbi training algorithms applied to read and spontaneous speech recognition. *Pattern Recognition and Image Analysis***,** 847-857.

SANVIDO, V., GROBLER, F., PARFITT, K., GUVENIS, M. & COYLE, M. 1992. Critical success factors for construction projects. *Journal of construction engineering and management,* 118**,** 94-111.

SINGAPORE CONTRACTORS ASSOCIATION LIMITED. 2017. *SLOTS REGISTERED CONTRACTORS* [Online]. Available: https://www.scal.com.sg/slotsRegistered [Accessed 31 July 2017].

SIRL, D. Markov chains: an introduction/review. MASCOS Workshop on Algebraic Dynamics, ed, University of New South Wales, Sydney, Australia, 2005.

SKIBNIEWSKI, M. J. & GHOSH, S. 2009. Determination of key performance indicators with enterprise resource planning systems in engineering construction firms. *Journal of construction engineering and management,* 135**,** 965-978.

SPAGNOLO, G. 2012. Reputation, competition, and entry in procurement. *International Journal of Industrial Organization,* 30**,** 291-296.

TAVAKOL, M. & DENNICK, R. 2011. Making sense of Cronbach's alpha. *International journal of medical education,* 2**,** 53.

ULUBEYLI, S., MANISALI, E. & KAZAZ, A. 2010. Subcontractor selection practices in international construction projects. *Journal of Civil Engineering and Management,* 16**,** 47-56.

WANG, J. & YUAN, H. 2011. Factors affecting contractors' risk attitudes in construction projects: Case study from China. *International Journal of Project Management,* 29**,** 209-219.

WILLIAMS, T. 2016. Identifying success factors in construction projects: A case study. *Project Management Journal,* 47**,** 97-112.

YEON, J., ELEFTERIADOU, L., & LAWPHONGPANICH, S. (2008). Travel time estimation on a freeway using Discrete Time Markov Chains. *Transportation Research Part B: Methodological*, *42*(4), 325-338

YU, S.-Z. & KOBAYASHI, H. 2003. A hidden semi-Markov model with missing data and multiple observation sequences for mobility tracking. *Signal Processing,* 83**,** 235-250.

YUAN, J., ZENG, A. Y., SKIBNIEWSKI, M. J. & LI, Q. 2009. Selection of performance objectives and key performance indicators in public–private partnership projects to achieve value for money. *Construction Management and Economics,* 27**,** 253-270.

ZHANG, Y., ZHAO, D. & LIU, J. 2014. The Application of Baum-Welch Algorithm in Multistep Attack. *The Scientific World Journal,* 2014**,** 7.

ZUCCHINI, W., MACDONALD, I. L. & LANGROCK, R. 2016. *Hidden Markov models for time series: an introduction using R*, CRC press.

<u>APPENDIX A</u>: THE RESULTS OF HOW CRITICAL THAT DIFFERENT CONTRACTORS ASSESSED SUBCONTRACTOR PERFROAMCNE.

Appendix A1. Correlation results for all performance questions for all 5 interviewed organisations

Organisation 1	*Mean Score*	*Rankings*
Mean Score	1	
Rankings	-0.3622645	1

Organisation 2	*Mean Score*	*Rankings*
Mean Score	1	
Rankings	-0.8558362	1

Organisation 3	*Mean Score*	*Rankings*
Mean Score	1	
Rankings	-0.2322371	1

Organisation 4	*Mean Score*	*Rankings*
Mean Score	1	
Rankings	-0.7527483	1

Organisation 5	*Mean Score*	*Rankings*
Mean Score	1	
Rankings	-0.982658	1

Appendix A2. P-values result of the regression analysis for all performance questions for organisation 1

SUMMARY OUTPUT FOR O1								
Regression Statistics								
Multiple R	0.362264464							
R Square	0.131235542							
Adjusted R Square	-0.085955573							
Standard Error	1.701728982							
Observations	6							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	1.749807221	1.749807221	0.60423992	0.480374291			
Residual	4	11.58352611	2.895881528					
Total	5	13.33333333						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	13.52231426	13.98259395	0.967081953	0.388267101	-25.29959027	52.34421879	-25.29959027	52.34421879
X Variable 1	-2.932637801	3.77271259	-0.777328708	0.480374291	-13.4073672	7.542091602	-13.4073672	7.542091602

Appendix A3. P-values result of the regression analysis for all performance questions for organisation 2

SUMMARY OUTPUT FOR O2								
Regression Statistics								
Multiple R	0.85583622							
R Square	0.732455635							
Adjusted R Square	0.665569543							
Standard Error	0.944359334							
Observations	6							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	9.766075129	9.766075129	10.95079141	0.029676701			
Residual	4	3.567258204	0.891814551					
Total	5	13.33333333						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	9.352833761	2.056933062	4.54698013	0.01044154	3.641872029	15.06379549	3.641872029	15.06379549
X Variable 1	-2.74379335	0.829141486	-3.309198002	0.029676701	-5.045859171	-0.441727528	-5.045859171	-0.441727528

Appendix A4. P-values result of the regression analysis for all performance questions for organisation 3

SUMMARY OUTPUT FOR O3								
Regression Statistics								
Multiple R	0.232237088							
R Square	0.053934065							
Adjusted R Square	-0.182582419							
Standard Error	1.77582463							
Observations	6							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	0.719120865	0.719120865	0.228035121	0.657907114			
Residual	4	12.61421247	3.153553117					
Total	5	13.33333333						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	14.43919475	24.66360584	0.585445407	0.589685322	-54.03795297	82.91634247	-54.03795297	82.91634247
X Variable 1	-2.734299868	5.725919945	-0.47753023	0.657907114	-18.63200227	13.16340253	-18.63200227	13.16340253

Appendix A1. P-values result of the regression analysis for all performance questions for organisation 4

SUMMARY OUTPUT FOR O4								
Regression Statistics								
Multiple R	0.752748301							
R Square	0.566630004							
Adjusted R Square	0.458287505							
Standard Error	1.201901266							
Observations	6							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	7.555066719	7.555066719	5.229988315	0.084142435			
Residual	4	5.778266614	1.444566654					
Total	5	13.33333333						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	16.27383941	5.97020557	2.725842388	0.052668313	-0.302108624	32.84978744	-0.302108624	32.84978744
X Variable 1	-3.605663404	1.576648285	-2.286916771	0.084142435	-7.983140817	0.771814009	-7.983140817	0.771814009

Appendix A2. P-values result of the regression analysis for all performance questions for organisation 5

SUMMARY OUTPUT FOR O5								
Regression Statistics								
Multiple R	0.982657992							
R Square	0.965616728							
Adjusted R Square	0.957020911							
Standard Error	0.338542324							
Observations	6							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	12.87488971	12.87488971	112.3356428	0.00044851			
Residual	4	0.458443621	0.114610905					
Total	5	13.33333333						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	25.3712532	2.146628208	11.81911852	0.000293335	19.41125782	31.33124858	19.41125782	31.33124858
X Variable 1	-5.46549726	0.515668845	-10.59885101	0.00044851	-6.897223501	-4.03377102	-6.897223501	-4.03377102

APPENDIX B: FACTOR ANALYSIS RESULTS IN SPSS

Communalities

	Initial	Extraction
Q1	1.000	.812
Q2	1.000	.732
Q3	1.000	.905
Q4	1.000	.925
Q5	1.000	.884
Q6	1.000	.915

Extraction Method: Principal Component Analysis.

Total Variance Explained

	Initial Eigenvalues		
Component	Total	% of Variance	Cumulative %
1	4.143	69.044	69.044
2	1.031	17.181	86.225
3	.356	5.929	92.154
4	.288	4.802	96.957
5	.114	1.908	98.864
6	.068	1.136	100.000

Extraction Method: Principal Component Analysis.

Total Variance Explained

Component	Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.143	69.044	69.044	2.594	43.234	43.234
2	1.031	17.181	86.225	2.579	42.991	86.225
3						
4						
5						
6						

Extraction Method: Principal Component Analysis.

Component Matrix[a]

	Component	
	1	2
Q3	.868	-.389
Q5	.861	-.378
Q6	.840	.457
Q4	.836	.476
Q2	.802	.297
Q1	.774	-.462

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

Rotated Component Matrix[a]

	Component	
	1	2
Q3	.890	.337
Q5	.877	.340
Q1	.874	.219
Q4	.257	.927
Q6	.273	.917
Q2	.359	.777

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

Component Transformation Matrix

Component	1	2
1	.709	.705
2	-.705	.709

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

Oneway

```
[DataSet1] D:\DATA\KKA97\Learning\Abdulaziz\Data.sav
```

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean	
						Lower Bound	Upper Bound
Factor1	1	30	10.4667	1.75643	.32068	9.8108	11.1225
	2	30	11.5333	2.33021	.42544	10.6632	12.4034
	3	30	11.5667	1.94197	.35455	10.8415	12.2918
	4	30	11.4000	1.69380	.30924	10.7675	12.0325
	Total	120	11.2417	1.97461	.18026	10.8847	11.5986
Factor2	1	30	10.6667	1.88155	.34352	9.9641	11.3692
	2	30	10.9333	1.65952	.30299	10.3137	11.5530
	3	30	12.0333	1.95613	.35714	11.3029	12.7638
	4	30	11.5667	1.94197	.35455	10.8415	12.2918
	Total	120	11.3000	1.91675	.17497	10.9535	11.6465

Descriptives

		Minimum	Maximum
Factor1	1	7.00	14.00
	2	7.00	15.00
	3	9.00	15.00
	4	9.00	15.00
	Total	7.00	15.00
Factor2	1	8.00	15.00
	2	8.00	15.00
	3	8.00	15.00
	4	9.00	15.00
	Total	8.00	15.00

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
Factor1	1.042	3	116	.377
Factor2	.293	3	116	.830

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Factor1	Between Groups	24.492	3	8.164	2.155	.097
	Within Groups	439.500	116	3.789		
	Total	463.992	119			
Factor2	Between Groups	34.333	3	11.444	3.295	.023
	Within Groups	402.867	116	3.473		
	Total	437.200	119			

Homogeneous Subsets

Factor1

	Contractor	N	Subset for alpha = .05
			1
Tukey HSD[a]	1	30	10.4667
	4	30	11.4000
	2	30	11.5333
	3	30	11.5667
	Sig.		.132

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 30.000.

Factor2

	Contractor	N	Subset for alpha = .05	
			1	2
Tukey HSD[a]	1	30	10.6667	
	2	30	10.9333	10.9333
	4	30	11.5667	11.5667
	3	30		12.0333
	Sig.		.246	.107

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 30.000.

APPENDIX C: THE ARBITRARY VALUED FOR BAUW WELCH ALGORITHM

Appendix C1. The starting probabilities for performance question for two hidden states for subcontractor that worked for multiple contractors

	Q1	Q2	Q3	Q4	Q5	Q6
S1	0.444	0.520	0.528	0.458	0.458	0.583
S2	0.556	0.480	0.472	0.542	0.542	0.417

Appendix C2. The transition probabilities for performance questions for two hidden states for subcontractors that worked for multiple contractors

Q1	S1	S2
S1	0.676	0.324
S2	0.432	0.568
Q2	S1	S2
S1	0.434	0.566
S2	0.364	0.636
Q3	S1	S2
S1	0.625	0.375
S2	0.281	0.719

Q4	S1	S2
S1	0.625	0.375
S2	0.281	0.719
	S1	S2
S1	0.471	0.529
S2	0.365	0.636
	S1	S2
S1	0.466	0.534
S2	0.250	0.521

Appendix C3. The emission probabilities of performance questions with two hidden states for subcontractors worked for multiple contractors

Q1	A	B	C
S1	**0.450**	**0.300**	**0.250**
S2	**0.415**	**0.360**	**0.225**
Q2	A	B	C
S1	**0.368**	**0.421**	**0.211**
S2	**0.143**	**0.500**	**0.357**
Q3	A	B	C
S1	**0.500**	**0.364**	**0.136**
S2	**0.333**	**0.429**	**0.238**

Q4	A	B	C
S1	**0.381**	**0.524**	**0.095**
S2	**0.214**	**0.643**	**0.143**
Q5	A	B	C
S1	**0.400**	**0.450**	**0.150**
S2	**0.235**	**0.412**	**0.353**
Q6	A	B	C
S1	**0.526**	**0.368**	**0.105**
S2	**0.200**	**0.600**	**0.200**

Appendix C4. The starting probabilities for performance question for subcontractor that worked for multiple contractors

	The staring probabilities for performance questions for two hidden states					
	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**
S1	0.389	0.417	0.472	0.389	0.417	0.500
S2	0.500	0.361	0.417	0.472	0.500	0.333
S3	0.111	0.222	0.111	0.139	0.083	0.167
	The staring probabilities for performance questions for four hidden states					
	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**
S1	0.300	0.400	0.450	0.350	0.350	0.500
S2	0.500	0.300	0.300	0.400	0.400	0.250
S3	0.100	0.200	0.150	0.150	0.150	0.150
S4	0.100	0.100	0.100	0.100	0.100	0.100

Appendix C5. The transition probabilities for performance questions for subcontractors that worked for multiple contractors

The transition probabilities for performance questions for three hidden states							
Q1	**S1**	**S2**	**S3**	**Q4**	**S1**	**S2**	**S3**
S1	0.676	0.324	0	**S1**	0.412	0.500	0.088
S2	0.136	0.568	0.295	**S2**	0.259	0.537	0.204
S3	0	0.375	0.625	**S3**	0.118	0.529	0.353
Q2	**S1**	**S2**	**S3**	**Q5**	**S1**	**S2**	**S3**
S1	0.382	0.529	0.088	**S1**	0.471	0.529	0
S2	0.298	0.468	0.234	**S2**	0.275	0.529	0.196
S3	0.125	0.375	0.500	**S3**	0	0.400	0.600
Q3	**S1**	**S2**	**S3**	**Q6**	**S1**	**S2**	**S3**
S1	0.735	0.441	0	**S1**	0.436	0.487	0.077
S2	0.191	0.596	0.191	**S2**	0.250	0.521	0.229
S3	0	0.333	0.458	**S3**	0.222	0.333	0.444

Appendix C6. The transition probabilities for performance questions for subcontractors that worked for multiple contractors

The transition probabilities for performance questions for four hidden states									
Q1	S1	S2	S3	S4	Q4	S1	S2	S3	S4
S1	0.550	0.250	0.150	0.050	S1	0.600	0.200	0.150	0.050
S2	0.150	0.550	0.200	0.100	S2	0.100	0.600	0.200	0.100
S3	0.100	0.200	0.550	0.150	S3	0.050	0.150	0.600	0.200
S4	0.100	0.150	0.200	0.550	S4	0.050	0.100	0.250	0.600
Q2	S1	S2	S3	S4	Q5	S1	S2	S3	S4
S1	0.450	0.300	0.150	0.100	S1	0.400	0.350	0.150	0.100
S2	0.250	0.450	0.200	0.100	S2	0.200	0.400	0.250	0.150
S3	0.100	0.300	0.450	0.150	S3	0.100	0.200	0.400	0.300
S4	0.100	0.150	0.300	0.450	S4	0.100	0.150	0.350	0.400
Q3	S1	S2	S3	S4	Q6				
S1	0.500	0.300	0.150	0.050	S1	0.250	0.350	0.250	0.150
S2	0.200	0.500	0.200	0.100	S2	0.250	0.250	0.350	0.150
S3	0.100	0.200	0.500	0.200	S3	0.150	0.250	0.250	0.350
S4	0.100	0.100	0.300	0.500	S4	0.150	0.250	0.350	0.250

Appendix C7. The emission probabilities of performance questions with three hidden states for subcontractors worked for multiple contractors

Q1	A	B	C	Q4	A	B	C
S1	0.200	0.350	0.450	S1	0.095	0.524	0.381
S2	0.261	0.348	0.391	S2	0.143	0.643	0.214
S3	0.227	0.455	0.318	S3	0.143	0.643	0.214
Q2	A	B	C	Q5	A	B	C
S1	0.211	0.421	0.368	S1	0.150	0.450	0.400
S2	0.357	0.500	0.143	S2	0.353	0.412	0.235
S3	0.214	0.500	0.286	S3	0.167	0.611	0.222

Q3	A	B	C	Q6	A	B	C
S1	0.136	0.364	0.500	S1	0.105	0.368	0.526
S2	0.238	0.429	0.333	S2	0.200	0.600	0.200
S3	0.143	0.524	0.333	S3	0.188	0.563	0.250

Appendix C8. The emission probabilities of performance questions with four hidden states for subcontractors worked for multiple contractors

Q1	A	B	C	Q4	A	B	C
S1	0.450	0.350	0.200	S1	0.381	0.524	0.095
S2	0.391	0.348	0.261	S2	0.214	0.643	0.143
S3	0.318	0.455	0.227	S3	0.214	0.643	0.143
S4	0.348	0.261	0.391	S4	0.350	0.450	0.200
Q2	A	B	C	Q5	A	B	C
S1	0.368	0.421	0.211	S1	0.400	0.450	0.150
S2	0.143	0.500	0.357	S2	0.235	0.412	0.353
S3	0.286	0.500	0.214	S3	0.222	0.611	0.167
S4	0.214	0.286	0.500	S4	0.353	0.235	0.412
Q3	A	B	C	Q6	A	B	C
S1	0.500	0.364	0.136	S1	0.526	0.368	0.105
S2	0.333	0.429	0.238	S2	0.200	0.600	0.200
S3	0.333	0.524	0.143	S3	0.250	0.563	0.188
S4	0.238	0.333	0.429	S4	0.105	0.368	0.526