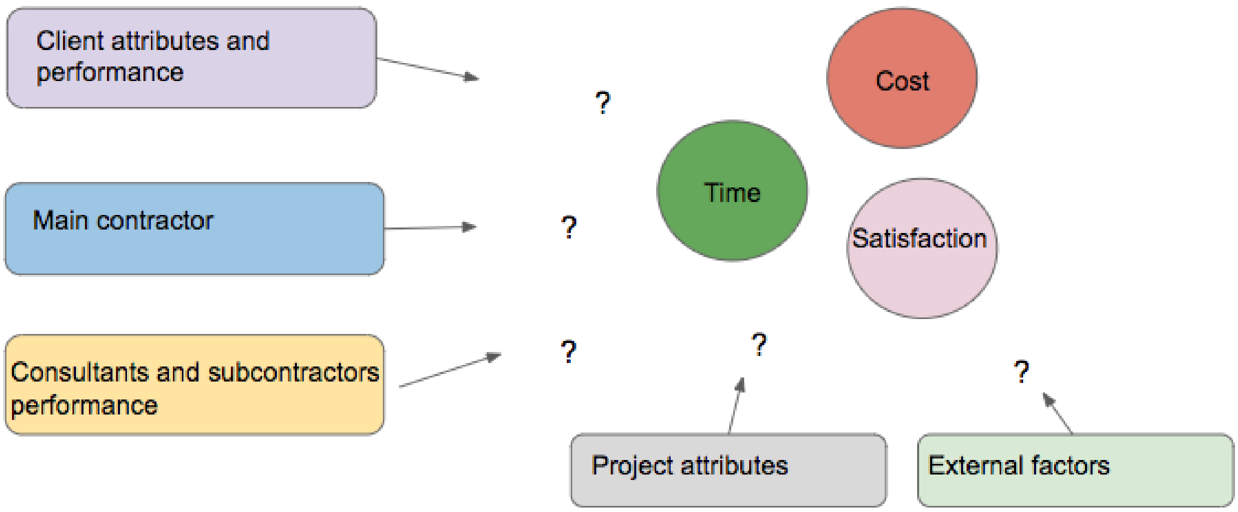

Performance prediction model for Swedish construction projects

Master's thesis in the master's Programme Design and construction project management

MAY SHAYBOUN
SEBASTIAN SCHENSTRÖM

Department of Architecture and Civil Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2018

MASTER’S THESIS ACEX30-18-75

Performance prediction model for Swedish construction projects

Master’s Thesis in the Master’s Programme Design and construction project management

MAY SHAYBOUN

SEBASTIAN SCHENSTRÖM

Department of Architecture and Civil Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY

Göteborg, Sweden

Performance prediction model for Swedish construction projects
Master's Thesis in the Master's Programme Design and construction project management

MAY SHAYBOUN

SEBASTIAN SCHENSTRÖM

Examensarbete ACEX30-18-75/ Institutionen för bygg- och miljöteknik,
Chalmers tekniska högskola 2018

Department of Architecture and Civil Engineering
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone: + 46 (0)31-772 1000

Front page image
Visualisation of factors behind project performance, authors' own creation.

Department of Architecture and Civil Engineering
Göteborg, Sweden, 2018

Performance prediction model for Swedish construction project

Master's thesis in the Master's Programme Design and construction project management

MAY SHAYBOUN
SEBASTIAN SCHENSTRÖM
Department of Architecture and Civil Engineering
Chalmers University of Technology

ABSTRACT

Construction projects tend to be influenced by interrelated issues that can result in unexpected costs or time overrun. Therefore, a need is highlighted by the industry as well as in the literature, to develop predictive models, that can aid in decision-making and guidance for corrective action throughout the development and implementation of construction projects. Performance in construction projects are often measured by key performance indicators. It is assumed that project performance is influenced by project attributes, external factors and the project organization. In this report, a previously made survey, that includes data from Swedish construction projects is analysed to find the most influential factors behind project performance, so that performance not only can be improved, but also predicted. The study that is analysed in this research includes answers from 324 main contractor representatives and 256 clients that participated in the survey in 2014. It includes answers about project attributes, external factors, the project organisation, cost and time for the projects as well as satisfaction of the client and the contractor.

First, a literature review, that covers theory about project management success in construction projects is carried out. Thereafter, a statistical correlation method is used to extract the features that are strongly correlated with three performance indicators: cost variance, time variance and client- and contractor satisfaction. A regression analysis is thereafter done in order to develop a model for predicting project cost, time and satisfaction. The result is an identification of the most correlating factors for project performance. The conclusion that can be made is that, although external factors and technical aspects of a building are important for project success, for example, project technical complexity, the use of blasting work or amount of prefabrication, the most recurring factors behind project performance can be derived to human related factors. Human related factors in the project life cycle are of high impact, such as the client role, the architect performance and collaboration throughout the stages of projects. These are also the factors that are most suitable for predicting if a construction project will be successful or not.

Keywords: performance, prediction, regression, success factors, key performance indicators

Contents

Preface

This report is part of the last stage of a master's programme, conducted at the Department of Architecture and Civil Engineering, Construction management at Chalmers University of Technology, Sweden. In this study, predictive models for key performance indicators have been developed and discussed. This report has been carried out with Professor Christian Koch as a supervisor and Prolog Bygglogistik AB as a hosting company. We would like to thank all people who made this report possible. We thank Christian Koch for complete support and supervision. We also thank Prolog Bygglogistik AB for their engagement and provision of resources in terms of professional consultancy and a great environment for conducting the work.

Göteborg May 2018

May Shayboun

Sebastian Schenström

Notations

All variables occurring in the report are alphabetically listed in the following table.

Roman upper case letters

N	Number of instances in a set of data
Z	Number of standard deviations from the mean value

Roman lower case letters

a	Level of significance
s	Standard deviation of the input data
t	Value taken of a t-distribution
w_a	Correlation coefficient in the regression model
w_0	Constant in the regression model
	$x_{\square}$ Attribute value in the regression model
x	Instance in the dataset
$\underline{x}$	Mean value

1. Introduction

Prolog is a Swedish consultancy company specialized in project management and with an aim on making the construction industry more efficient. The company has a variety of clients that range from municipalities to big construction firms. Tobias Nordlund who is the head of the office in Gothenburg expressed that his vision is to be able to point out what the most critical factors are to achieve a successful project and to tell his clients that ''if you pay attention to these factors, the outcome of the project will be like this''.

Construction projects are generally circumstantial and are frequently influenced by "multiple interrelated issues" whether it is a positive or negative, external or internal. Therefore, construction projects tend to differ from what is actually planned for (Grau, Back and Mejia-Aguilar, 2017). Mbugua et al. (1999) state that the characteristics of construction projects ''(..) increase the need to get things right first time every time, because the consequences of getting it wrong with even a single client or project, can seriously impact the business.'' (p.261). Therefore, there is a pressure on project teams to keep the planned schedule and planned cost regardless of the vulnerability of construction industry to be affected by uncontrolled events (Grau, Back and Mejia-Aguilar, 2017).

The need for a model to predict performance is highlighted in literature and driven first from industry actors' perspective. Grau, Back and Mejia-Aguilar (2017) state that in general it is difficult for project teams to accurately predict project outcome. One issue is that social aspects, in contrast to technical ones, are less documented and investigated in regards of how they affect the ability to achieve desired goals. El-Gohary, Aziz and Abdel-Khalek (2017) elaborate in a predictive model for productivity and argue that such models are important for decision-making, planning, estimation, scheduling and controlling. If relationships between "relevant influential factors" and the result are established in the models, it is easier to make changes and correct undesired outcome. Moreover, the need for prediction models is not a new trend. Al-Momani (2000) argues that adequate prediction of time for a construction project is important for the management and decision-making process and will benefit both the client and the contractor and ''decrease the risk of disputes''. Sanvido et al. (1992) call for a development of a predictive model to evaluate the cost performance when achieving or not achieving a project's critical success factors. The common perception among different authors is that critical success factors play an important role when trying to predict performance. Alias et al. (2014) suggest that if critical success factors for a construction project are identified, involved parties can use them to improve, but also forecast the possibility of project success. Chua, Kog and Loh (1999) state that there are a lot of factors that can affect the performance and outcome of a project. If the relationships between critical success factors and single objectives are identified, resources can be allocated more efficiently. Chan, Scott and Chan (2004) also claim that identified relationships between outcome and critical success factors, can improve performance. Furthermore, the authors claim that this knowledge also can be used for predicting project performance. Chan, Scott and Chan (2004) call for future studies in this area. The need for establishing a clear link between critical success factors and project outcome is also highlighted by Ogunlana and Toor (2009). This report can be seen as a response to these claims.

1.1 Problem formulation

The report is based on the following assumptions:
Project management success is when desired project plans are achieved, especially in terms of keeping budgets and schedules in the expected range as well as meeting or exceeding the participants expectations. Success in construction projects is highly affected by project performance. Project performance, the outcome of a construction project, can be measured by using key performance indicators. Key performance indicators are used to evaluate the success or failure of a project.

A construction project consists of a series of main processes such as project brief, design, contracting and construction. In these processes and among the project attributes and external factors; performance influencers can be identified and thereafter linked to each performance objective; cost, time and satisfaction. Performance influencers derive from two different sources: the general project conditions and the project organisation, see figure 2.1. The general project conditions consist of project attributes which consider for instance the type of contract, size of building and amount of prefabrication and external factors which consider market conditions, weather, regulations and requirements. The second source, project organisation, is a process which is embodied by the characteristics and performance of the client, main contractor, engineers, subcontractors and consultants.

The need for establishing clear links between critical success factors and project outcome is highlighted by Ogunlana and Toor (2009). The common perception among different authors, for instance Chua, Kog and Loh (1999), is that if the links are established, resources can be allocated more efficiently. According to Chan, Scott and Chan (2004), identified relationships between project outcome and critical success factors can not only improve but also be used for predicting project performance.

The aim of the study is firstly to identify what factors are influential to project performance and to answer what the most important performance influencers are behind cost variance, time variance client- and contractor satisfaction respectively. In order to reach the aim of the study, two sets of data, from two surveys answered by clients and contractor representatives from swedish construction projects, are analysed using a statistical analysis tool. Secondly, this paper aims to build a predictive model by using the most influential factors on projects performance that are found in the statistical analysis. The model is expected to be used as a guidance to improve performance so that undesired outcomes can be avoided, and project management success achieved.

RQ1: *What are the most important factors affecting the performance in Swedish construction projects?*

RQ2: *How can these factors be used in practice to help practitioners improve their performance?*

2. Methodology

The aim of the study is to build a model for predicting cost, time and satisfaction in construction projects. This would require a research design consisting of a statistical analysis to find the relationships between features and do the regression model, also a theoretical review to discuss the empirical results. The workflow includes finding the performance influencers correlated with the performance measures and using the regression analysis to develop predictive model. It is also necessary to elaborate on the variables that are found to be influential for project performance. Therefore, it is important to compare with theory about the subject to discuss and reach conclusions.

2.1 A quantitative and qualitative study

The research methodology that is adopted in this report is a mixed method which here is a combination of quantitative and qualitative study (Sachdeva, 2009). According to Sachdeva (2009), mixed method is beneficial, since it avoids the weaknesses that can come from using only one approach, quantitative or qualitative, which might make the research biased towards one observation or theory. Mixed method is therefore suitable for studies in social science and business management (Sachdeva, 2009). When dealing with business and social contexts, findings are more easily supported when using a combination of research methods to describe or explain the phenomenons (Bryman and Bell, 2011). The first part of this research is the survey performed by Josephson (2013), which gathered data about the opinions of clients and main contractors about clients, main contractors, consultants and subcontractors as well as opinions and facts about different project processes. The quantitative part of this research is the statistical analysis of the collected data and the development of the predictive model for construction projects success.

The research also adopts a pragmatic abductive research strategy. The report's data gathering is through a survey that is developed according to the framework presented by Josephson (2013). Mixed approach is the most suitable for making connections between theoretical and empirical results. This requires the combination of theoretical search and empirical analysis to come up with a rich discussion and reach a solid conclusion in a broad study that includes a considerable number of factors and more than one perspective.

Josephson (2013), who constructed a model of factors leading to productivity. It is the basis for forming the survey that was used for collecting the data in 2014. In other words, a survey for gathering data was built based on the model, *figure 2.1.* Both contractor representatives and clients were asked about specific construction projects. Contractor representatives were asked questions about general project conditions, such as project size, percentage of prefabrication, site-conditions and project comlexity. Questions were also asked about disturbances and how well the client, consultants documents, suppliers, piping and ventilation subcontractors perform. They were also questioned about their own company priorities, administrative support and how they have chosen subcontractors. Planned and actual *costs for construction* for the projects, planned and actual *time for construction* were declared by the contractors. The questions in the client survey differs from the contractor representatives survey. The clients were asked about type of contract and procurement, project program, pre-design, design and planning of the project. Very brief questions were asked about collaboration within the project team and the performance of the architect, structural engineer, different consultants and the main contractor. Josephson put more focus on examining how well the main contractors performed according to the clients. Planned and actual *total project costs* for the projects were declared by the clients. The respondents were also questioned about their own *satisfaction* of the projects; if the project exceeded their expectations, if they were happy with the result and if the project could be considered to be a success. In this report, satisfaction and cost-variance from the client survey are used as performance indicators and in the contractor survey, satisfaction and cost-variance are also complemented by a third performance indicator; time-variance. These performance measures have been seen to be the most common key performance indicators in the literature and they are represented very well in the data set.

2.1.1 The literature review

The theory research has been carried out at online libraries such as Google Scholar and Chalmers Library, using two main search strings and multiple simple ones when calling for a broader view. The replicability of the theoretical study can't be ensured since the selection of literature was according to what is considered fit with the subject of the research. Narrowing the search string resulted in irrelevant literature that was out of the scope of the research. Therefore, it was decided to use a main search string that resulted in 31700 articles, then scan the titles and abstract of the articles and therefore choose the literature that is valuable for the enrichment of a sufficient discussion.

It is worth mentioning that although causes of delays and problems with the construction industry might be exclusive for one country, studies that have been conducted in different countries can provide useful general knowledge about the reasons and influence of factors concerning delays or cost overrun in construction projects (Sambasivan and Soon, 2007). Furthermore, this report does not make distinctions by the type of facility, type of project or geographical location. Since it is assumed that buildings have similar characteristics; they consist of structures that are made by similar materials, require craftsmanship, project management and a set of equipment for conducting the work. Construction projects also share similarities in the project organizations that consist of clients, consultants, contractors and supply chains of subcontractors and material providers. In general, they have a set of contractual forms that organize the relationships between different actors in the project organization. Therefore, the authors of this report choose to make a general analysis where the features of building type, structural material, nationally specific contracts and power distribution among actors are not taken in consideration. Instead, the focus is towards the managerial aspects that exist in the survey questions.

Since the survey was conducted earlier and not for purpose of this report, it is necessary to investigate the representation of the questions that have been used and what aspects it did cover or leave out. This study examines the correlations between the model presented by Josephson (2013) and general research that has been conducted within the area of performance and critical success factors. Thus, to ensure that the survey and model are sufficient to cover the aspects that affect projects performance, it was necessary to find if previous literature on the subject is compatible with Josephson's survey model. The conclusion is that the model is enough to cover a considerable part of the success factors that are mentioned and discussed in the body of literature. The model does not cover all possible factors that can be identified in the literature. This can be explained by country context, success and failure factors can exist in one place and not exist in another. The literature that have been covered in this report is generally exploring the topic. However, major similarities were found to be sufficient for discussing the empirical results to come up with conclusions. Major aspect that are lacking in the model are *absence of disputes* and *predictability of time and cost* as key performance idicators, *regulations and requirements* as an external factor, *competence and style of monitoring*, *client organization size*, *top management competence*, *construction management effectiveness*, *late deleveries* as project organization factors also the *financial situations* of the client nor the contractor are considered.

2.1.2 Data analysis

The data analysis was done to conclude what factors are highly important for project performance and build the predictive model. After that, discussion should provide an understanding about the results and how the features that are found to be influential on performance measures are practiced. However, the survey questions are of subjective nature and rely on the respondent perception, for example, about the organization of consultants or an individual consultant. It is difficult to understand the exact practice in the perception of the respondent such as the client or the main contractor when they answer the questions. This limits the ability of the authors to build full overview about the practice behind each answer, which in turn make it more difficult to draw conclusions at the end of the research. Also, since the interpretation of Josephson questions in is limited, the comparison of the statistical results with previous theory becomes more difficult, especially in the discussion part where the authors need to explain how the regression analysis resulted factors are practised to provide useful recommendations. Since the model aims to establish links between performance influencers and outcome of the projects in terms of cost variance, time variance and satisfaction, it can also have the potential to be used as advisory practice to generate more satisfactory results for construction projects. Examples are the terms good collaboration and architect's performance, which are broad and neither described in detail or defined. It is not clear what good collaboration is and it is also not known if the respondents have the same perception on good collaboration. Another example of this limitation is the perception about the client. It is not clear whether this perception include the client organization or the client representative. Furthermore, there is also a probability that entry mistakes have been made during the collection and handling of data, which is also outside the control of the authors of this report. Altogether these limitations can have impact on the quality of the statistical analysis. The statistical analysis on the other hand, is replicable because it is following a systematic procedure, see chapter 2.2.

2.1.3 The survey

The data used in this research was collected for the report ‘’Produktivitetsläget i svenskt byggande 2014’’ and was based on the model presented in the report made by Josephsson (2013). The model visualises the relationship behind input and output, in Josephson’s report productivity is the output. In this report, the broader term performance will be used as an output. The literature study shows that the factors behind performance can be categorized into the model for factors behind productivity. Hence, it is assumed that this model also can be used to visualise the factors behind performance and that the questions in the survey made by Josephson (2013) are sufficient to represent the factors behind performance. The data was collected in a large survey about construction projects, sent to 1000 people and resulted in 580 responses. The survey was answered by 324 contractor representatives giving a 72% answering rate and 256 clients with 62% answering rate. The questions from the survey have been used in this report as a set of input variables that are compared with three performance outputs for the construction projects that are represented in the collected data.

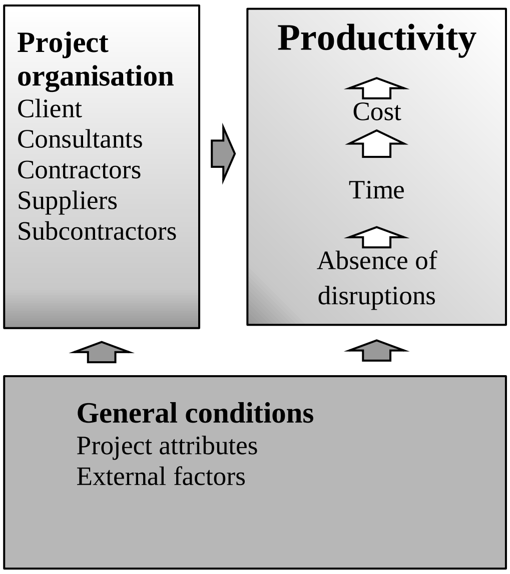

Figure 2.1 The ecosystem of factors and variables in a construction project that affects productivity. Presented by Josephsson (2013).

Findings made in the literature study show that several authors argue that the ‘’iron triangle’’ is outdated and call for a broader view on performance measurement that goes beyond the traditional measures: cost, time and quality, see chapter 3.2. Due to the limitation of not designing the survey questions, authors of this paper decided to follow the availability of data that already exists, hence this report does not take other performance objectives in consideration. Previously made literature reviews about KPI:s, for instance by Chan and Chan (2004), concludes that cost- and time variance as well as participants’ satisfaction, are suitable measures for project performance. Time variance and cost variance are considered to be objective measures and satisfaction as a subjective one. Therefore, these three performance indicators are considered to be suitable as performance measures in this report. Also, in the data set used in this report, cost- and time variance and satisfaction of the client and the contractor are the only variables that can be identified as relevant ‘’outputs’’. According to Chan and Chan (2004) time variance is defined by the difference in percentage between actual time for construction and planned time for construction. Cost variance is defined by the difference in percentage between actual cost and planned cost.

2.1.4 Availability of data - Pre-construction and construction phase

The data set was split into two categories, “pre-construction’’ and ‘’construction phase”. If the prediction model is to be used, it is logical to use information that are available *at the time of prediction*. For example, it is not possible to know if there was good collaboration on site or if the main contractor managed the subcontractors properly when the project is still in design phase. Therefore, this distinction was necessary for the prediction to be logical and useful at earlier stages of a project’s life cycle. The pre-construction phase consists of the project program, pre-design phase, design phase, procurement and contracting of main contractor as well as subcontractors and consultants. The construction phase consists of external factors, construction site aspects and performance of the main contractor and the consultants.

2.2 The statistical analysis model

The quantitative part of the study is a statistical analysis of the data, which ensures the credibility of this paper beside the mixed method approach. The statistical analysis follows the rules of feature selection, linear regression, cross validation (Witten et al., 2011). Validation using previous literature that supports or explains the findings of the data analysis is also made (Bryman and Bell, 2011). To analyse the data in this report, a regression analysis is used because it is, according to Montgomery, Peck and Vining (2012), suitable for problems where the relationship between variables is to be found and understood. Also, the prediction characteristics of regression analysis are sufficient to answer the research question regarding finding the most important variables and narrowing down the focus of practice.

Linear regression analysis is a common technique for numeric class predictions. It provides a linear equation consisting of attributes or inputs and predetermined weights that is used to predict a class or outcome of new instances (Witten et al., 2011). When the linear regression is performed, the result is a set of numeric weights that are used to predict future new instances. The formula of the regression model, where y is the output variable, w_0, (…), w_k are the weights and x_0, (…), x_k are the attributes values:

$$y = w_0 + w_1x_1 + w_2x_2 + (\dots) + w_kx_k$$

Formula 2.1 The formula of the regression model

There is also another method that has been considered, called neural networks, that is commonly used in prediction problems as well Montgomery, Peck and Vining (2012). Artificial neural networks are suitable for handling nonlinear problems. Another method for nonlinear regression which is called support vector regression has been also tried on the data set available for this research. The results have shown that the mean square error has not improved and then it was concluded that the nature of the research problem is more linear than nonlinear. However, the nonlinearity makes neural networks more prone to overfit and generate solutions that are bad in predicting future instances that are not used in the training set Montgomery, Peck and Vining (2012). Training set is the set of data that is used to learn the model in machine learning software. Comparisons between regression analysis technique and neural networks in the context of predicting cost of construction projects made by Lowe, Emsley and Harding (2006) as well as Kim, An and Kang (2004) show that both methods perform well. However, the use of a regression analysis is judged to be sufficient and suitable in this report. Since in this report authors are more interested in finding the quantifying the relationships between features and project performance and neural networks have the risk of overfitting.

2.2.1 Weka - a statistical analysis tool

The tool that is used to do the linear regression and attribute selection is Weka (Waikato Environment for Knowledge Analysis), which is a data mining software. According to Frank, Hall and Witten (2016), Weka provides the opportunity to easily implement learning algorithms on datasets because it provides tools for data processing, feeding data into learning schemes and analysing the results, without writing any code. The software contains a variety of methods for common data mining problems such as regression, classification, clustering, association rule mining and attribute selection (Frank, Hall and Witten, 2016). Also, Weka can be used in different ways and using preprocessing data filters such as normalization. It can either be used to apply learning schemes and analyse outputs, utilise learned models to generate predictions or to compare different learning schemes in order to find the most suitable. It was decided that Weka is an appropriate tool for the statistical analysis in this report, since the software provides sufficient tools for what is intended to be done in this report, such as linear regression and attribute selection.

2.2.2 Cross validation and attribute selection

Normally, if a *lot of data* is available, it is separated into three independent data sets; one for training, a second for testing and a third for validation (Witten et al., 2011). The error rate for the testing data gives a closer indicator for future performance of the model because it will be seen how the prediction model work on instances that haven't been used before. Thereafter it is clear whether the predictor is accurate or not. However, when the amount of *data is limited*, as in this report, a solution for this problem must be implemented (Witten et al., 2011). The fact is that prediction models' future performance, or error rate, is not ensured to be as good as the one at the time when the model is built. The result is that the model is useful for describing the data but not valid to predict future outputs (Witten et al., 2011; Olive, 2017; Montgomery, Peck and Vining, 2012). This is because the model is built by training on the same set of instances, which makes it over optimistic and prone to overfit. Overfitting can be reduced by using cross validation and avoiding unnecessary attributes by using attribute selection (Olive, 2017). Cross validation is a method to train and test the data by randomly dividing the instances in the data set to K number of folds. The model is thereafter tested for K times. Each time K is used to validate the K-1 other folds (Hall, 1999). In this report, the data is split into 10 folds for cross validation, which is a standard way to predict the error rate for a learning scheme. The 10 folds cross validation is, by a large number of experiments, agreed to be the most appropriate for a good error estimate, according to Witten et al. (2011). The final error estimate is the average of the 10 error estimates (Witten et al., 2011).

2.2.3 Understanding the data

The systematical steps for the analysis starts with understanding the data and getting to understand what it contains. The survey consists of two parts. The first part contains information regarding project number, location, structure type, client type, type of the construction, the area of the building and phone numbers of respondents. The second part of the survey consists of questions with yes/no questions and ranking questions from 1 to 5 asking about performance of clients, consultants, main contractor and subcontractors. The second step is to clean the data from noisy instances, such as projects that lack answers. Answers such as “don’t know” can not be used in the analysis and are therefore changed to blank spaces. This is the best way to deal with ‘’don’t know’’ answers so that the results are not affected. Data such as number of floors, type of facility, number of building blocks, as well as type of construction are deleted because a lot of missing values and therefore they are not useful in the analysis.

2.2.4 Detecting and deleting outliers

The next step is to detect outliers that exist in the output variable. Outliers are defined as “data object that deviates significantly from the rest of the objects, as if it were generated by a different mechanism” (Han, Pei and Kamber, 2011) and it is very important to distinguish them and exclude them from the analysis. Having extreme values that deviate highly from the mean value can cause the model to be unable to generalize especially if it is getting out of the pattern that is existing in the data sample (Yan and Su, 2009). Therefore, for the purpose of improving the model, Grubb test was used to detect and remove the outliers by first calculating z-scores and then compare it with the t-distribution (Han, Pei and Kamber, 2011). Following formula, as can be found on page 555 in the book by Han, Pei and Kamber (2011), is used for detecting the outliers:

$$z = \frac{|x-\underline{x}|}{s} \quad , \quad z \geq \frac{N-1}{\sqrt{N}} \sqrt{\frac{t^2_{a/(2N),N-2}}{N-2+t^2_{a/(2N),N-2}}}$$

Formula 2.2 Formula for detecting outliers in the data set

2.2.5 Dummy variables

After that, another necessary step in the data analysis was to process the nominal input variables and transform them to become what is known as "dummy variables". Dummy variables are a coding technique for nominal values to make them usable in the regression analysis, knowing that regression can only be applied on numerical values. The dummy variables take the values of (1) when the nominal value is true and the value (0) when the nominal value is not. Now the data set is ready for the regression analysis process (Montgomery, Peck and Vining, 2012).

2.3 Preparing the client survey data for analysis

Calculated cost variance, which is the percentage of the difference between actual and budgeted cost over the budgeted cost, is determined to be the output of the study together with satisfaction of the client. Information about cost of changes and cost of disturbance exist in the survey as an abstract number. This information is meaningless if they are not taken with proportion to actual cost. They are not useful in the analysis unless they are calculated as a percentage:

*Cost variance= ((actual cost - budgeted cost)/budgeted cost) *100*

Formula 2.3

*Percentage of cost of changes = (cost of changes/actual cost) *100*

Formula 2.4

*Percentage of cost of disturbance = (disturbance cost/actual cost) *100*

Formula 2.5

The survey also includes information about the amount of working months for own craftsmen, own office workers, subcontracted craftsmen and subcontracted office workers. The percentage of own and subcontracted employees working months in proportion to each other was thereafter calculated:

Percentage of own employees working months = (own employees working moths/

*(own employees working months + subcontracted employees working months+own craftsmen working months+subcontracted craftsmen months)) *100*

Formula 2.6

*Percentage of subcontracted employees working months = (external employees working months/ (own employees working months + subcontracted employees working months+own craftsmen working months+subcontracted craftsmen months))*100*

Formula 2.7

*Percentage of own craftsmen working months = (own employees working moths/ (own employees working months + subcontracted employees working months+own craftsmen working months+subcontracted craftsmen months)) *100*

Formula 2.8

*Percentage of subcontracted craftsmen working months = (own employees working moths/ (own employees working months + subcontracted employees working months+own craftsmen working months+subcontracted craftsmen months))*100*

Formula 2.9

2.3.1 Cost variance

The values of cost variance that were undefined, due to lack of entry of one of the budgeted or actual cost parameters were deleted, to make sure that data is clear from noise and in good quality for the analysis. It was also decided to delete the entries of cost of changes and cost of disturbaces due to missing many values. Deleting the projects with missing values would not solve this issue because it results in minimising the data-set size, hence reducing the quality of the analysis. Therefore, the decision to not consider cost of changes and cost of disturbances is justified in this way. Six outliers were found in the cost variance in the client's perspective. The model without the outliers was generating less mean square error and then it was decided to remove the outliers and proceed with the analysis.

2.3.2 Client satisfaction

The same process of cleaning and organizing the data for the client's perspective excel sheet for client's satisfaction is performed. However, client's satisfaction is concepyualised by the authors to be the sum of the three questions regarding happiness, success and expectations of the results of the project. To evaluate if the three questions are representatives of the satisfaction, a correlation test has been performed for the three attributes, see table 2.1. When the correlation coefficient is close the the value of +1 or -1, the two attributes are highly correlated and the more the value reaches 0 the correlation decreases. It is observed that there are strong correlations, therefore adding the three parameters together to represent the overall satisfaction can be done. Finally, after cleaning and organising both the output variables, the input variables were also organised and cleaned. This resulted in 58 input variables (of which 21 dummy variables) related to pre-construction and 11 input variables related to construction phase.

		Correlation coefficient/ client	Correlation coefficient/ contractor representatives
Successful project	Exceeded expectations	0.6057	0.6772
Successful project	Happy with the result	0.7852	0.7041
Happy with the result	Exceeded expectations	0.6099	0.6396

Table 2.1 Correlation coefficients of satisfaction parameters

2.4 Preparing the contractor survey data for analysis

The main contractor survey was processed the same way as the client survey. Data cleaning and understanding is crucial in both cases, since this results in more accurate observations. Data such as type of structure, geographical location and type of facility were removed, since they contained a lot of missing values.

By taking planned values in proportion to actual values, cost variance and time variance are calculated as they are the output variables in the analysis. The percentage of time for correction is taken in proportion to actual construction time, because the numbers representing the time taken for correction don't have a meaning otherwise. This is also done to the cost of disturbances, which was calculated as percentage of the actual cost of the project. However, cost of disturbances and time for correction are discarded afterwards and not included in the final model.

11 different types of contracts were represented among the projects in the survey. Three of the eleven categories, ''partnering contract'', ''samverkansentreprenad'' and ''direct labour contract'' were removed, since these types of contracts were lacking enough representation. The 8 remaining contract types were boiled down to two main categories; design-and build contract and traditional contract, using the categorization of common Swedish contract types according to Nordstrand (2008). The percentage of duration the main contractor's workers and subcontractor's workers have worked in proportion to each other was also calculated. The main contractor's survey was divided into two data sets, one that represents "pre-construction phase" and another that represents the "construction phase".

The next step is to detect and remove outliers from the data if they exist. With using Grubb test as the one performed on the data of the client survey. The results of the Grubb test show that there were 13 outliers in the cost variance and 4 outliers in the time variance. Deleting the outliers resulted in less mean square error and then it was decided to proceed with deleting the outliers.

The main contractor dataset finally consists of 17 input variables and 2 dummy variables in the pre-construction phase. The construction phase dataset consists of 25 input variables.

2.5 Forward and backward stepwise regression

In order to do the regression analysis, the number of attributes is supposed to be minimized in order to reach the set of inputs that produce the most accurate prediction. It was decided to scan the set of input variables with a stepwise backward elimination, where it starts with the full set of data and systematically deleting irrelevant attributes, one by one, until reaching the final set of attributes. The final set of input variables can not be changed any further and have the highest capability to produce accurate predictions. An attribute is irrelevant when it's value does not change systematically with the output class (Hall, 1999). It is also important to remove redundant attributes that are characterised as being correlated with one or more than other attributes.

A comparison between forward and backward stepwise regression models was made by Lowe, Emsley and Harding (2006) and Attalla and Hegazy (2003), because both methods deal with the significant variables differently. According to both authors, forward stepwise regression is not free from problems. The nature of forward stepward regression is that if a dependent variable already encloses another variable, that also has a significance on the output, other variables will appear to be less important when added to the analysis model. The backward stepwise regression can overcome this issue, since it allows all variables to be included in the beginning. The results show that backward selection method is more capable to extract more significant variables (Lowe, Emsley and Harding, 2006; Attalla and Hegazy, 2003).

In conjunction with the backward elimination, *CfsSubsetEval* "correlation-based feature selection" was conducted to find the attributes with the most individual predictive capability in relation to the output. In other words, the data set is searched to find correlations between input variables and respondent variables (output variables). The process of attributes selection was through Weka software as well. It is a Weka configured object and then the backward stepwise is is also configured within it as a search method (Witten et al., 2011). This method has the capability to find the input variables that are highly correlated with the output individually and to detect the degree of redundancy among them. In this case, it distinguishes the attributes that have correlation with the output but not intercorrelated.

The number of attributes at the beginning of the data analysis in the contractor data set were 19 in the pre-construction phase and 25 in the construction phase. In the client data set there are 58 input variables in the pre-construction phase and 11 in the construction phase. The distinction between pre-construction and construction, is according to the distinction presented in chapter 2.1.4. As a result, the separated data sets were analysed, attributes were selected, and regression was done four times for the client's survey, twice for every project phase and twice for each output, satisfaction and cost variance. The contractor survey was analysed six times, twice for each project phase and three for each output, satisfaction, time variance and cost variance.

2.6 Model adequacy

The numeric prediction evaluation can be evaluated using the root mean square error which is the same as the standard deviation of residuals, which defined as the root of the square value of the difference between the predicted value and the actual value (Witten et al., 2011). It gives an indicator of how much the actual values are deviating from the regression line generated by the model. The less the root mean square error the better the model is behaving in predicting. In other words, If the standard deviation, is low, this means that the data observations are closer to the ground truth value. But if the standard deviation is high, the the data are covering a large range of values. Since the root mean square error take the same measurement unit of the predicted output, it is easy to interpret (Han, Pei and Kamber, 2011).

p: numeric value of the prediction for the ith instance

a: actual value

n: number of observations

$$\sqrt{\frac{(p_1-a_1)^2+\ldots\ldots\ldots\ldots+(p_n-a_n)^2}{n}}$$

Formula 2.10 Numeric prediction evaluation, Witten et al., 2011.

Moreover, Standard deviation can give an indicator of how much the possible value of observations may vary. However, the normal distribution rule of thumb only applies if the errors are normally distributed. One way to test this assumption is to use a Quantile-Quantile plot analysis (Han, Pei and Kamber, 2011). This method is a powerful visualization for graphs and it can show if there is a change in the distribution.

2.7 Delimitation

The report covers construction projects in the Swedish construction industry. The study covers 580 respondents, 324 contractor representatives and 256 clients answered the survey in 2014. The projects are constructed between the year 2010 and 2014, but all the finishing dates are in the year 2013 and 2014. The study does not make any distinction between project type, building type, geographical location. It is aimed to make a general comparison between the projects. It only covers the Swedish construction market as the applicability to other contexts such as other countries or other construction industries is not known. In the report, only cost, time and satisfaction are considered as key performance indicators. The analysis only covers the perception of the client and the contractor.

2.8 Ethical considerations

The authors of this report have taken the ethical considerations when conducting the research. It was evaluated and made sure that the research study, analysis, discussion and conclusion would not cause harm for any of the participants in the survey. It is also important to make sure that participants have an informed consent and protected privacy, the survey questions are straight forward and do not require covert observation. Another ethical consideration regards the data collection and sharing. The data is confidential, and it is only used according to the consent of the owner. The authors also declare honesty and integrity when conducting the research.

3. Theoretical framework

The literature study in this report is based on previous articles and books about project success, performance, key performance indicators, factors behind success and failure as well as previous experiences of developing prediction models. This chapter aims to paint a picture to how performance is measured, what the factors are behind project performance and how performance can be predicted and why prediction of project performance is a good idea. The link between project performance and project success is also established, in order to justify why it is important to put effort in improving project performance.

3.1 Project success and performance

‘’Project success’’ is a term that does not have a single definition. As expressed by Ogunlana and Toor (2009) as well as Chan and Chan (2004), the perception of what project success is and what makes a project successful differs between stakeholders. According to Frödell, Josephson and Lindahl (2008), success in construction projects, when asking clients, is when a project is finished on budget and on time as well as keeping high quality. This definition is also used by Frimpong, Oluwoye and Crawford (2003) who specifies that a project is successful when the goals in the project plan are met, such as time schedule, budgeted cost and technical performance. Sanvido et al. (1992) refer to project success as when the expectations for all parties, for instance client, contractor, engineer or end user are met or exceeded. According to De Wit (1988), a distinction should be made between project success and success of project management effort. Frödell, Josephson and Lindahl (2008) elaborates in these terms and consider project management success as short term goals while project success considers long term goals. The term ‘’project success’’, according to De Wit (1988), is when the project objectives are met, judged from the perspective of all possible stakeholders in a project, whereas project management effort is usually measured using cost, time and quality. When project success is mentioned in this report, it is from the perspective of the client and the contractor and does not, in contradiction to Sanvido et al. (1992), include the judgement by engineers, end-users or other stakeholders. It is also important to clarify that it is project management success that is aimed at, according to the distinction made by De Wit (1988) and Frödell, Josephson and Lindahl (2008).

3.2 Performance indicators

Beatham et al. (2004) argue that for sustainable business success, it is important to measure future performance, performance of completed work and percepted performance. Measurement of performance is a tool to achieve continuous improvement in the construction industry (Mbugua et al., 1999). In order to measure performance, KPIs, key performance indicators, are often used. Frequently mentioned KPIs throughout literature are cost, quality and time. These three measures are commonly known as ''The iron triangle'', see Ogunlana and Toor (2009) as well as Chan and Chan (2004). Ogunlana and Toor (2009) argue that one should look beyond the iron triangle to search for new KPIs. Chua, Kog and Loh (1999) as well as Jha and Iyer (2006) present safety as one KPI that is often used in construction projects. Client satisfaction, the amount of defects and the accuracy of predicting project outputs are also presented as possible KPIs (Ogunlana and Toor, 2009; Alias et al., 2014). Jha and Iyer (2006) introduce a fifth parameter which is called ''no-dispute'', aiming to measure the relationships between parties in a project. According to Xiong et al. (2014) the satisfaction of participants in a construction project is important for the performance. The satisfaction of the contractor is especially important for obtaining good collaboration in the project (Xiong et al., 2014).

Chan and Chan (2004) present an extensive framework for measuring performance in construction projects. Key performance indicators are divided into *objective* and *subjective* measures. The objective KPIs presented by the authors are for instance time and speed of construction, rate of accidents, time variation and cost variation. The subjective performance indicators are for instance quality, client satisfaction and construction team satisfaction. A framework by Cox, Issa and Ahrens (2003) divides key performance indicators into *quantitative* and *qualitative* measures. Quantitative KPIs are for instance construction cost, units per man hour, cost per unit, on-time completion, amount of rework etcetera. Qualitative KPIs are for instance safety, turnover or workers motivation. Although different, the frameworks presented by Chan and Chan (2004) and Cox, Issa and Ahrens (2003) share a common denominator: KPIs are divided into groups depending on how quantifiable they are. For instance objective KPIs presented by Chan and Chan (2004) are easily measured and so are the quantitative indicators presented by Cox, Issa and Ahrens (2003). Khosravi and Afshari (2011) present a model for measuring success in construction projects. In the model, several KPIs are weighted and brought together to give a total score. The authors suggest that the most important performance measures are (1) cost, (2) time, (3) quality, (4) client satisfaction and (5) safety, health and environmental. Furthermore, Cox, Issa and Ahrens (2003) provide an example of that KPIs can be different depending on who is asked. The authors conclude that companies that use subcontractors to a high degree arranges ''on-time'' as the most important KPI, while companies that to a higher degree perform the work themselves rather value quality/rework, safety and units/man hour as the most important KPIs. This report will focus on following KPIs: construction time- and cost- variance as well as contractor and client satisfaction.

The analysis of the report is depending on the principle of evaluating the project according to the satisfaction of the contractor and the client, percentage of cost variance and time variance. In the article by Freeman and Beale (1992), the authors aim to develop performance measures to be applied in the project phase. Cost and time targets are found to be the highest success criteria for project participants. The argument of how to interpret cost and time over- and underrun is made with the assumption that the contract between the sponsor and the project manager is with fixed price. From the sponsor's perception, if the project is finished before the due date, it creates revenues because of finishing before expected dates, while it is considered as a failure if it exceeds the time plan. On the other hand, cost overrun will not affect the sponsor. From the project manager's view, cost overrun is regarded as a loss, whereas cost underrun is a gain. It is therefore assumed that schedule underrun, and cost underrun are success criterias in construction projects.

KPIs	
Cost	Walker (1995) Ogunlana and Toor (2009) Chan and Chan (2004) Chua, Kog and Loh (1999) Jha and Iyer (2006) Alias et al. (2014) Cox, Issa and Ahrens, (2003) Khosravi and Afshari (2011)
Time	Walker (1995) Ogunlana and Toor (2009) Chan and Chan (2004) Chua, Kog and Loh (1999) Jha and Iyer (2006) Alias et al. (2014) Cox, Issa and Ahrens, (2003) Khosravi and Afshari (2011)
Quality	Walker (1995) Ogunlana and Toor (2009) Chan and Chan (2004) Chua, Kog and Loh (1999) Jha and Iyer (2006) Alias et al. (2014) Cox, Issa, R. and Ahrens (2003) Khosravi and Afshari (2011)
Safety	Chua, Kog and Loh (1999) Jha and Iyer (2006) Cox, Issa and Ahrens (2003) Khosravi and Afshari (2011)

Client satisfaction	Ogunlana and Toor (2009) Jha and Iyer (2006) Alias et al. (2014) Khosravi and Afshari (2011)
Contractor satisfaction	Ali, Al-Sulaihi and Al-Gahtani (2013) Xiong et al. (2014)
Predictability of time and cost	Ogunlana and Toor (2009) Alias et al. (2014)
Defects	Ogunlana and Toor (2009) Alias et al. (2014)
Absence of disputes	Jha and Iyer (2006)

Table 3.1 Summary of identified and acknowledged KPIs. The purpose is to provide wide overview of the key performance indicators that are identified in the literature.

3.3 Factors behind success and failure

One early report made by Walker (1995) points out that there are four factors that affect construction performance in construction project, more specifically time performance. These are concerning the complexity of the project and the project scope, the effectiveness of the construction management, client's ability to create and maintain relationships with the design team and the construction management and the communication between design team, construction management and client's representative teams.

The model of the conditions and factors that affect the productivity in a construction project, as presented by Josephson (2013), is used as a lens to categorize the factors that are obtained from the literature in this chapter. According to Josephson (2013), productivity is the relationship between project input and output. Productivity should not, as argued by Tangen (2005), be confused with performance. However, productivity is part of the broader term performance, that is aimed to be measured in this report. Fig. 3.1 shows the model that is used to organise and visualise the factors behind project performance. The model visualises the relationship behind input and output, which is described by Koch and Lundholm (2018); The general conditions for the construction project are the inputs for the project process in which the

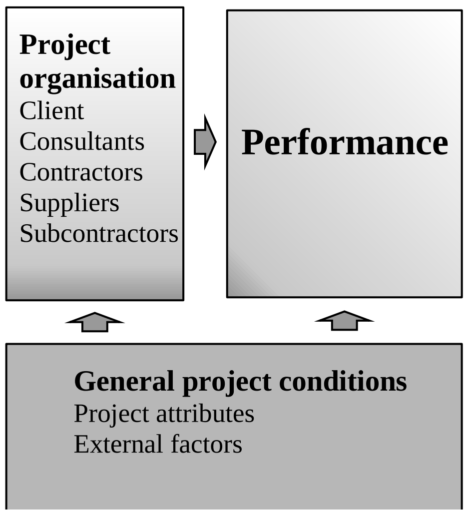

Figure 3.1 Based on figure 2.1 presented by Josephsson (2013). Revised to fit the objectives of this report; performance is used as a general output instead of productivity.

performance of the different parties have impact on the output. In the project process, interruptions and disturbances can occur, which have impact on time and cost. The output used by Josephson is productivity and is measured by cost per square meter. In this report, the broader term performance will be used as an output and time, cost and satisfaction will be the measures. The literature study shows that the factors behind performance can be categorized in to the model. Hence, it is assumed that this model can visualise the factors behind performance and that the questions in the survey made by Josephson (2013) is sufficient to represent the factors behind performance.

3.3.1 General project conditions

According to Josephsson (2013), general conditions for a project, as shown in fig. 3.1, consist of external factors and the project attributes. The project attributes, also called project-related factors, according to Josephson (2013) and Chan, Scott and Chan (2004), firstly refer to project characteristics: size, material and type of building, amount of prefabrication, type of structure and complexity, of which the last factor is also mentioned by Lam and Wong (2009), who state that project performance considering time, cost, quality and safety are highly affected by construction buildability. Josephson (2013) finds that construction projects with prefabricated structures generally have higher cost for construction than built in situ structures. Other factors behind project performance are also site-conditions, according to findings made by Josephson (2013), Al-Momani (2000) and Attalla and Hegazy (2003). Site-conditions, according to Josephson (2013), consider for instance lack of space in a construction site which in turn affects delivery of goods and the possibility to obtain adequate storage. Ground conditions are also mentioned by Josephson (2013). It considers the geological conditions and whether piling or extensive blasting work is needed. Project attributes also refers to conditions related to the type of contract and procurement (Josephson, 2013). Josephson (2013) also finds that the use of partnering and what type of partnering has impact on the project outcome. In the study made by Koch and Lundholm (2018), it could be observed that 24% of 430 investigated swedish construction projects use some type of partnering contract. The authors found that partnering contracts in swedish construction projects are usually between contractors and clients and more rarely between consultancies and suppliers. Moreover, it is found that contractors tend to perform better during partnering contracts considering collaboration, time and quality. Chan, Chan et al. (2004) present that partnering between two or more organisations in construction leads to improved communication, collaboration and trust between the organisations. Hence it can create more efficient projects and lower risk for time- and cost overruns.

Project attributes	
Project complexity	Walker (1995) Josephsson (2013) Koch and Lundholm (2018) Lam and Wong (2009) Chan, Scott and Chan (2004)
Amount of prefabrication	Josephsson (2013) Chan, Scott and Chan (2004)
Procurement and tendering method	Josephsson (2013) Koch and Lundholm (2018) Chan, Scott and Chan (2004)
Type of contract	Josephsson (2013) Koch and Lundholm (2018)
Size and number of floors	Josephsson (2013) Koch and Lundholm (2018) Chan, Scott and Chan (2004)
Type of project	Josephsson (2013) Koch and Lundholm (2018) Chan, Scott and Chan (2004)
Site conditions	Josephson (2013) Al-Momani (2000) Attalla and Hegazy (2003)
Ground conditions	Josephson (2013)
Partnering	Chan et al. (2004) Josephsson (2013) Koch and Lundholm (2018)

Table 3.2 Summary of identified and acknowledged project attribute related factors. The purpose is to provide wide overview of the performance influencers that are identified in the literature.

The external factors, that affect the outcome of a construction project, are presented in the framework presented by Chan, Scott and Chan (2004). In this framework, the category of external environment consists of factors belonging to the social, physical, economic or political surroundings of the project. Jha and Iyer (2006) argue that external factors are the ones that can have the most negative impact on project quality. Among the external factors, Josephsson (2013) concludes that weather conditions, in particular Swedish winters, is one of the most common causes behind delays in construction projects. Throughout the literature, weather is widely acknowledged as an important factor and mentioned by Frimpong, Oluwoye and Crawford (2003), Jha and Iyer (2006) and Al-Momani (2000). Furthermore, geographic location also has impact on a construction project. Josephsson (2013) identifies that construction projects in larger cities have a tendency to be more expensive and concludes that market conditions have an impact on the cost for production. Increase in material price during a project is one example of market conditions (Frimpong, Oluwoye and Crawford, 2003). Furthermore, market conditions affect the procurement of the main contractor and Jha and Iyer (2006) found that high competition with aggressive tendering in the early stages of a project can have negative impact on the quality performance. Furthermore, political decisions, regulations and requirements are also factors that can have impact on the project outcome (Attalla and Hegazy, 2003; Josephsson, 2013). For instance, incidents associated with collective pay agreement, as found by Josephsson (2013).

External factors	
Weather	Frimpong, Oluwoye and Crawford (2003) Jha and Iyer (2006) Al-Momani (2000) Josephsson (2013) Koch and Lundholm (2018)
Market conditions	Al-Momani (2000) Josephsson (2013) Chan, Scott and Chan (2004) Frimpong, Oluwoye and Crawford (2003)
High competition and aggressive tendering	Jha and Iyer (2006)
Regulations and requirements	Attalla and Hegazy (2003) Josephsson (2013)

Table 3.3 Summary of identified and acknowledged external factors. The purpose is to provide wide overview of the performance influencers that are identified in the literature.

3.3.2 Project organisation attributes

The project organisation refers to the team that is responsible for carrying out the project. The client, contractor, consultancies and subcontractors are all part of the project organisation and hence they all have impact on the project outcome (Josephson, 2013). Sambasivan and Soon (2007) argue that the contractor's and client's roles have a great impact on time overrun. In the framework presented by Chan, Scott and Chan (2004) and Alias et al. (2014), a distinction is made between human-related and project management-related factors. The project management-related factors in general refer to the organisational structure, systems for communication or for instance programs for quality assurance. In other words, the project management-related factors try to describe the project infrastructure. Human-related factors on the other hand refer to soft factors such as knowledge, skills, commitment and relationships within the project organisation. (Chan, Scott and Chan, 2004; Alias et al., 2014). One study made by Jha and Iyer (2006) shows that good communication and relationships between participants have positive impact on quality of a project while bad knowledge and conflicts seem to have the most negative impact on the quality. Sambasivan and Soon (2007) conclude that bad communication is one of the most common causes behind project delays. The conclusion made by Jha and Iyer (2006) is that human-related factors are very important to achieve good quality in construction project.

3.3.2.1 Client

In the study made by Josephson (2013), clients are evaluated and judged by their capabilities to make decisions, to give clear message on time, to collaborate and to plan the project. In the report made by Koch and Lundholm (2018), clients' ability to encourage innovative methods and procedures is also evaluated. The clients are considered to be good at creating a collaborative environment and provide explicit goals. On the other hand, ability to plan the project, give clear message and encourage innovation is something that clients lack, according to main contractors. Findings made by Josephson (2013) show that, according to site-managers, clients lack capabilities to give clear messages and to plan the projects, while they are generally better at creating a collaborating environment and to make decisions. Meanwhile, a survey made by Odeh and Battaineh (2002) shows that slow decision-making by clients is one of the common causes behind time overruns in projects. Regardless, the general view throughout the literature is that client decision-making is a factor behind project success, see Frödell, Josephson and Lindahl (2008), Chan, Scott and Chan (2004) and Alias et al. (2014). This is consistent with the findings made by Le-Hoai, Dai Lee and Lee (2008), which show that changes in design is one of the most significant causes behind cost and time overrun. Al-Momani (2000) found that changes made by the client is one of the most common causes behind time overruns. Attalla and Hegazy (2003) identified change in project scope or design by the client as one of five main reasons behind bad performance. Chan, Scott and Chan (2004) put forward the client's ability to contribute to the design and construction. Among others, the client's financial situation is also presented as a factor behind project performance and causes behind failure. Odeh and Battaineh (2002) conclude that client financing is one of the most common causes behind time overrun in project. This is also found by Sambasivan and Soon (2007) and Le-Hoai, Dai Lee and Lee (2008) who conclude that time and cost overrun can be caused by client financial problems and that the contractor is not paid for completed work. Le-Hoai, Dai Lee and Lee (2008) state that this can be a source for disputes.

Furthermore, the client's ability to communicate, collaborate and manage the relationships within the project organisation, is also highlighted in the literature. One early report made by Walker (1995) points out that one of the most important factors is the client's ability to create positive project team relationships with the other parties, for instance the construction management and design team. One of the most significant factors behind project success is also the client's commitment and ability to participate in the project (Frödell, Josephson and Lindahl, 2008). What is also important, is client competence and ability to monitor and give feedback. As stated by Jha and Iyer (2006), who claim that these factors have impact on the quality of a project. Odeh and Battaineh (2002) conclude that interference by client is one of the most important factors behind time overruns in construction project. Overall, the type of client and client's experience are important success factors according to Chan, Scott and Chan (2004) and Alias et al. (2014). Findings made by Koch and Lundholm (2018) show that the type of client have impact on the productivity. For instance, public clients have a tendency to spend more time on a project in relationship to the project size, compared to private clients. The procurement criteria that the client has is also presented by the authors as factors behind success; Is the emphasis on price, quality or time?

Project organisation: Client	
Decision-making	Frödell, Josephson and Lindahl (2008) Chan, Scott and Chan (2004) Odeh and Battaineh (2002) Josephson (2013) Koch and Lundholm (2018)
Commitment and participation	Frödell, Josephson and Lindahl (2008) Alias et al. (2014)
Collaboration with design and construction team	Walker (1995)
Client's ability to contribute to the design and construction	Chan, Scott and Chan (2004)
Competence and style of monitoring and feedback	Jha and Iyer (2006)
Relationships between participants	Jha and Iyer (2006) Chan, Scott and Chan (2004)
Changes made by the client	Al-Momani (2000) Le-Hoai, Dai Lee and Lee (2008)
Client experience	Chan, Scott and Chan (2004)
Client type	Chan, Scott and Chan (2004) Josephson (2013) Koch and Lundholm (2018)
Client organisation size	Chan, Scott and Chan (2004)

Procurement criteria (time, quality or price)	Chan, Scott and Chan (2004) Josephson (2013) Koch and Lundholm (2018)
Problems with financing and payment	Sambasivan and Soon (2007) Le-Hoai, Dai Lee and Lee (2008) Odeh and Battaineh (2002)

Table 3.4 Summary of identified and acknowledged client related factors. The purpose is to provide wide overview of the performance influencers that are identified in the literature.

3.3.2.1 Contractor, subcontractors and consultants

One of the main reasons behind bad performance in construction project is that the contractor is not performing well (Attalla and Hegazy, 2003). One common cause behind bad performance is that the contractor lacks relevant experience for the project. Findings made by Sambasivan and Soon (2007) and Odeh and Battaineh (2002) show that this factor has impact on both the project cost and time. Throughout the literature, contractor's managerial skills, are highlighted as an important factor behind project performance. Walker (1995) points out that construction management effectiveness is one of the main factors behind construction time performance. Problems with the schedule is one factor behind bad performance in construction according to Attalla and Hegazy (2003). Inadequate planning is one cause behind time overrun (Sambasivan and Soon, 2007; Odeh and Battaineh, 2002). Sambasivan and Soon (2007) state that inadequate planning can lead to cost overrun. Alias et al. (2014) highlight the importance of a well-planned design and construction. According to Alias et al. (2014) the skills of the designer is important for the project success. Walker (1995) suggests that design team's communication with construction management and client's representative teams are one of most important critical success factors. The project managers' skills and competence are presented as factors behind project performance. Frödell, Josephson and Lindahl (2008) argues that the project manager's competence is one of the most important success factors. Alias et al. (2014) put forward the skill of the project manager as a factor behind project success and the study conducted by Jha and Iyer (2006) show that the competence of the project manager has positive impact on quality of a project. Chan, Scott and Chan (2004) present several human related factors behind project success. The managerial skills of the team leaders are highlighted as success factors. These managerial skills refer to ability to organise, give feedback, plan, communicate and coordinate the project as well as motivating the staff. A committed management is one key to success according to Frödell, Josephson and Lindahl (2008). What is also presented by Chan, Scott and Chan (2004) as well as Frödell, Josephson and Lindahl (2008) is the commitment to cost, quality and time.

At the construction site, poor management and supervision are presented as causes behind time and cost overrun (Le-Hoai, Dai Lee and Lee, 2008; Sambasivan and Soon, 2007). Odeh and Battaineh (2002) state that contractor's productivity during construction is an important factor behind time overrun. Frödell, Josephson and Lindahl (2008) identify that the competence of the workforce as well as their commitment and participation is important for project success. Interaction and communication between workforces are also success factors that are put forward by the authors. Also, it is not only the project management that should emphasize on cost, quality and time. The workforce should also have that way of thinking (Frödell, Josephson and Lindahl, 2008). What is also highlighted is the project management's ability to support the production. Findings made by Le-Hoai, Dai Lee and Lee (2008) show that bad support from the project management is one of the most severe causes behind time and cost overruns in construction projects.

The support that the construction project get from the own organisation is also highlighted (Josephson, 2013; Koch and Lundholm, 2018). Chan, Scott and Chan (2004) suggest that one of the factors behind project success is the support and provision of resources that the project team leaders get from the own organisation. Findings made by Jha and Iyer (2006) show that top management competence and support have positive impact on quality of a project. In the report made by Koch and Lundholm (2018), the administrative support, providing of enough workforce and the priority of the project is evaluated. Also, to what degree the site-manager was involved in choosing subcontractors and suppliers were evaluated. The own organisation contains a lot of factors that can determine performance of a project. For instance, organisational structure or control mechanisms as presented as project management-related factors by Chan, Scott and Chan (2004) or financial situation as put forward by Le-Hoai, Dai Lee and Lee (2008).

In the literature, subcontractors and suppliers actions as influencers behind time and cost overruns are also announced. Delays caused by subcontractors are one of the ten most important causes behind time overruns as reported by Odeh and Battaineh (2002). Sambasivan and Soon (2007) present that problems with subcontractors, supply of material and labour, availability and failure of equipment are some of the most common causes behind delays in construction projects. Chan, Scott and Chan (2004) present control of subcontractors as a success factor. In the report made by Koch and Lundholm (2018), the ventilation and piping subcontractors where evaluated considering their ability to collaborate, to manage problems and interruptions and to keep the time schedule. The subcontractor's way of working and production methods were also graded as well as their ability to deliver a result according to the site-manager's expectations and requirements in the contract. The result of the study shows that the performance of ventilation and piping subcontractors are not satisfying when asking the site-managers. The other way around, the subcontractors are not fully satisfied either. In specific, it is the main contractors' time planning that the subcontractors are not satisfied with, according to Koch and Lundholm (2018).

Project organisation: Contractor, subcontractors and consultants	
Contractor experience	Sambasivan and Soon (2007) Odeh and Battaineh (2002)
Top management competence	Jha and Iyer (2006)
Management and workforce commitment and participation	Frödell, Josephson and Lindahl (2008) Alias et al. (2014) Chan, Scott and Chan (2004)
Project manager competence	Frödell, Josephson and Lindahl (2008) Chan, Scott and Chan (2004) Alias et al. (2014) Jha and Iyer (2006)
Project team ability to give feedback	Chan, Scott and Chan (2004)
Workforce competence	Frödell, Josephson and Lindahl (2008)
Communication between parties	Frödell, Josephson and Lindahl (2008) Sambasivan and Soon (2007) Walker (1995)
Workforce and team quality thinking	Frödell, Josephson and Lindahl (2008) Chan, Scott and Chan (2004)
Construction management effectiveness	Walker (1995)
Management of subcontractors	Chan, Scott and Chan (2004)
Top management support	Josepson (2013) Jha and Iyer (2006) Chan, Scott and Chan (2004) Le-Hoai, Dai Lee and Lee (2008) Alias et al. (2014)
People in the production involved in choosing subcontractors and suppliers	Koch and Lundholm (2018)
Team leaders’ skills to plan, organise coordinate and motivate	Chan, Scott and Chan (2004)
Design team competence	Alias et al. (2014)
Well planned design and construction	Alias et al. (2014)
Bad planning	Sambasivan and Soon (2007) Odeh and Battaineh (2002)

Bad site management	Sambasivan and Soon (2007) Le-Hoai, Dai Lee and Lee (2008)
Failure during construction phase	Sambasivan and Soon (2007)
Financial situation	Le-Hoai, Dai Lee and Lee (2008)
Productivity	Odeh and Battaineh (2002)
Subcontractors supply of labour and material	Sambasivan and Soon (2007)
Deliveries not on time	Al-Momani (2000)
Subcontractor delays	Odeh and Battaineh (2002)

Table 3.5 Summary of identified and acknowledged factors related to the project organisation. The purpose is to provide wide overview of the performance influencers that are identified in the literature.

3.4 Predicting performance

The benefits of predicting project performance are highlighted in the theory. For instance, by Sanvido et al. (1992) who declare that it gives the possibility to take necessary actions before a project fails and even the possibility to know on beforehand what projects to avoid. Lowe, Emsley and Harding (2006) assert that 15-20% in early stages and 13-18% in the design phase, is a decent accuracy for predictions of project cost. A forecast that indicates a project's future cost can help the client to predict the tender price and the budget as well as aiding the design management. As stated by previous authors, for instance Alias et al. (2014), Chua, Kog and Loh (1999), Chan, Scott and Chan (2004) it is possible to predict performance if the relationships between key performance indicators and factors are established. Earlier studies have found what factors are the most important for time, cost and quality. For instance, Jha and Iyer (2006) who attempted to identify what factors are affecting one of the KPI:s, namely quality. Data from a questionnaire survey covering the answers from 112 indian construction professionals was analysed using SPSS software. As a result, the factor model identified the most influential factors, that positively versus negatively affects quality in a project, see chapter 3.3.

In studies made by Lowe, Emsley and Harding (2006), Kim, An and Kang (2004) and Attalla and Hegazy (2003), attempts were made to predict project cost at early project stages. Comparisons between different prediction models such as neural networks, regression analysis and case-based reasoning were also made by Lowe, Emsley and Harding (2006) and Kim, An and Kang (2004). Since clients often complain about inaccurate early cost estimates and because accurate early estimates are important for control and follow up in construction projects, there was a need for improving the accuracy of prediction- and estimation models. It was concluded in both articles, that although the best neural network model resulted in the best prediction, compared to linear regression and case-based reasoning, the three methods are statistically similar.

This, due to the complexity of determining the number of hidden neurons as well as the learning rate, which actually rely on trial and error to obtain the right configurations for a certain data set., as well as it is time consuming to update with new instances. Case-based reasoning is more suitable to use when the explanation of cost estimation in construction project is important and it is very suitable to use for long term especially since the model is easy to update with new instances, even with new variables. Compared to the other two methods, regression analysis has the ability to lower the number of variables, showing which are having a higher contribution in the respondent variables, hence it this method gives better result.

As discussed in the introduction of this report, determining the factors of success and failure are crucial to predict project outcome. Cost has been the centre of attention for authors who have built different models for prediction. Such models have been developed by Attalla and Hegazy (2003), Lowe Emsley and Harding (2006), Kim, An and Kang (2004), using similar methods of regression analysis and neural networks. Focusing on measurable and straight forward facts about the project as input variables such as duration, building type, finishing grades, type of installation, floor area, total units, procurement type, building function and type of client.

Other authors, such as Trost and Oberlender (2003), were interested in the accuracy of estimates in projects, because of its importance for early decision making and since very early estimates tend to be imprecise. The aim of the research made by Trost and Oberlender (2003) was to develop a scoring system to evaluate early estimates in terms of quality, inclusiveness and accuracy of estimations. This is important throughout the project lifecycle and also to determine if the project actually went as it was planned or budgeted. In other words, cost variance can be a consequence of inaccurate early estimations according to Trost and Oberlender (2003). In the research, a factor analysis and a multivariate regression analysis were performed to find the most important factors that affects a project's ability to accurately predict cost. The results show that the process design, site requirements, team experience, quality of cost information, time to prepare cost estimation and market conditions are the most influential factors that have impact on cost estimation accuracy.

Martin, Burrows and Pegg (2006) developed a model for duration prediction using regression analysis relying on construction cost, procurement route, client type, contractor selection, building function and region as independent variables. The purpose of giving valuable advice to clients at an early stage, but the model was expected to be as planning or scheduling tool instead. The study excluded all projects that are newly built such as refurbishments or repair as well as infrastructure projects. The duration was the whole project timeline from the moment parties commit to invest until the project was ready to be used. Cost of the projects were considered to be the contractor fee and excludes the consultant one. However, it is not exactly clear why this distinction and exclusion was made and how it actually serves the regression analysis in terms of choosing variables. If the assumption of different types of construction projects is made, then the distinction between their differences is also a matter of investigation because assuming, they will also have different set of variables to be influencing their outcomes as well.

What previous prediction models have actually missed out is the management performance in the project lifes cycle, also, the project implementation phase. It is mainly focused on early stages and the information available at that point of a project life. However, problems tend to happen throughout projects and studying project management's impact on project success lacks mathematical evidence due to complexity of projects environment. Therefore, it is necessary to find the link between project management performance and project success (Mir and Pinnington, 2014). It has been concluded from Mir and Pinnington study that project management performance in terms of PM leadership, PM staff, PM life cycle management processes have great impact on project team. More significantly, PM KPIs were proven to have great correlation with project success. Therefore, it was recommended for organizations to develop KPIs performance management to reach higher project success and satisfaction.

3.5 Summary of literature review

What success is in construction projects has different definitions. One perspective is that a successful project is when the involved parties are satisfied. Construction projects have several stakeholders; however the primary stakeholders are the client and the contractor. It can be concluded that good performance is vital for success of the single construction project, as well as it is for the whole construction industry. Performance can be measured by using key performance indicators. The literature study identifies eight common key performance indicators; cost, time, quality, safety, client satisfaction, predictability of time and cost, defects and absence of disputes.

Factors behind performance in construction projects are also reviewed. They are identified and organised according to the model presented by Josephson (2013) into two main divisions; general conditions and project organisation. General conditions consist of the project attributes and external factors. Project attributes have to do with the type, size and complexity of a project as well as the contract type. External factors relate to weather conditions, market conditions, ground conditions and governmental issues. The second division, the project organisation consists of factors relating to the performance of the different actors. For instance, the performance of the main contractor, client, subcontractors, consultants and suppliers. The factors consider collaboration, communication, planning and management skills and quality of performed work.

The theoretical study also covers a literature review of earlier reports about performance predictions. It is concluded that it is important to identify the most influential factors and to establish clear links between these and each performance objective. Successful models have been developed by previous authors using data mining tools such as regression analysis and artificial neural networks. The both analysing tools are suitable for developing prediction models, even though they are both suffering from drawbacks and benefits.

4. Empirical study

In this chapter, the result from the regression analysis is presented. It is based on two different surveys conducted by Koch and Lundholm (2018), one for construction project clients and one for contractor representatives from different Swedish construction projects. In the two surveys, the client and the contractor representatives were asked different questions about project conditions, processes, the different actors and the project outcome. The answers about project conditions, processes and the different actors represent the input variables, while answers about project cost, time and satisfaction are used as output variables. Time performance is evaluated by comparing the planned time for construction with the actual time for construction, this is referred to as time variance. Time for construction is defined as the elapsed period between start of construction to final inspection. Cost performance is evaluated by comparing the planned cost with the actual cost. In the client survey, cost is referring to the total project cost, which includes expenses for acquisition, design, administration, project management, inspections, building permit, insurances, interests during the time of the project, the main contractor fee etcetera. In the contractor survey, cost is referring to the cost for construction, which includes salaries for workers, subcontractor fees, material costs etcetera. Both the clients and contractors are asked questions about their satisfaction. If they considered it to be a successful project, if it exceeded their expectations and if they were happy with the result.

The result from each survey is presented twice. First the result about identified factors in the pre-construction phase which includes project program, pre-design phase, design phase, procurement and contracting as well as procuring subcontractors and consultants, is presented. Thereafter, this chapter presents the result of the found factors in the construction phase which consists of, site production aspects and performance of consultant and subcontractors. As a result, there will be two different results for each of the two surveys. First, it will be presented, what factors behind cost variance and satisfaction can be obtained from the clients' answers. Thereafter, it will be presented what factors that have been identified to affect cost, time and satisfaction, derived from the answers of the answers of the contractor representatives. The coefficients of the factors are also exposed. A negative coefficient signifies that the factor has a decreasing effect on the output and a positive coefficient indicates that a factor has an increasing effect on the output. Factors that were found to have a correlation with the output but were not included in the final formulas will also be presented. The main reason for discarding some factors, is because the established correlations are not distinctive enough and adding these factors to the formulas result in less accurate prediction abilities.

4.1 Pre-construction phase

The survey that concerned the client was phase wise. It consists of questions about the project program, pre-design phase, design phase, main procurement criteria, type of contract, type of payment, market conditions, type of client and if partnering was used. These constitute the pre-construction phase. The survey that concerns the main contractor's view had a different setting. The pre-construction phase in this survey consists of questions about, client type, contract type, gross area, client performance, the contractor's own organization, site conditions and in- and outsourced workers.

4.1.1 Result from client survey

In the survey answered by the clients, 58 input variables are associated with the pre-construction phase. Here it is presented which of these 58 input variables are identified in the regression analysis, as significant factors for the two different outputs; cost variance and satisfaction.

4.1.1.1 Cost variance factors

The result from the clients' survey identifies that seven out of the 58 variables, are influential factors behind project cost variance. The cost variance is the difference in percentage between planned and actual total project cost. Cost variance ranges from -49% to 43% among the projects in the client survey. The identified factors can be divided into two groups, factors with negative coefficients and factors with positive coefficients. Six out of seven factors are weighted by negative coefficients and tend to be success factors behind cost underrun whereas one out of the seven factors, adversely have a positive coefficient, hence it is contributing to cost overrun in the model. This factor concerns the type of contract and according to the model, statistically, cost variance tends to be higher in *trades contract* projects. On average 5.22% cost variance for projects with this contract type compared to for instance design-and-build contracts that have -0,47% cost variance on average. Among the factors concerning the procurement criteria and the design phase, success factors for cost variance can be identified. Statistically, *''reference project''* and *''environmental aspects''* are procurement criteria that, according to the model, are more important for cost variance than other procurement criterias such as time, price or specific competence of the contractor. In the pre-design phase, well-made *''investigation of the plot''* can also be identified as a factor that statistically can decrease the cost variance in construction projects. In the survey questions, *''investigation of plot''* refers to the work that is done to make sure that the location is suitable for the intended construction considering property boundaries, infrastructure and and the detailed development plan. On average, the cost variance is -1,94% for projects with well-made investigation of plot, a score of 4 or 5 in the survey. In projects with no investigation of plot or a poor investigation of the plot, score 1 or 2 by the respondents, there is a 4.1% cost variance on average. In the design phase, *''performance of the architect''*, *''absence of disturbances and* *problems''* and *''follow time plan weekly''* are identified as significant factors behind cost variance.

Coefficient	*Factor*
w_1: -3.7664	x_1: *Main procurement criteria - reference project*
w_2: -10.436	x_2: *Main procurement criteria - environmental aspects*
w_3: -3.1159	x_3: *Pre-design phase - investigation of plot*
w_4: -9.6548	x_4: *Design phase - absence of disturbances and problems*
w_5: -8.1426	x_5: *Design phase - follow time plan weekly*
w_6: -8.605	x_6: *Design phase - performance of architect*
w_7: 5.2732	x_7: *Type of contract - trades contract*
w_0: 21.5643	

Table 4.1 List of the factors in pre-construction phase that have the highest correlation with cost variance

4.1.1.2 Client satisfaction factors

In the analysis of the client survey, nine out of 58 variables in the pre-construction stage are identified as factors correlated to client satisfaction. The pre-construction phase includes variables about the project program, pre-design phase and design phase. In the project program phase, *''follow time plan weekly''* is identified as a factor behind client satisfaction. The coefficient for this factor is positive, which discloses that it positively contributes to client satisfaction. Likewise does the perception of *''good result''* of the project program uphold the client satisfaction. Projects with good result in the project program, tend to have more satisfied clients. On average 12.98 out of 15, compared to 9.22 out of 15 for project where the result of the project program was bad. The factor *''absence of disturbances and problems''* in the project program was slightly less important for the client satisfaction, thus it was not included in the final formula. From the pre-design phase, three factors that positively affect client satisfaction are found: *''conducting risk analysis'', ''well thought-through description of goals''* and *''analyse of suitable type of contract''*. *''Investigation of ground conditions''* in the pre-design phase is also observed to have a minor positive correlation with client satisfaction. However, it is not selected for the final model since the model development rely only on the strongest correlations between output and input variables and the correlation between satisfaction and *''Investigation of ground conditions''* is weaker than the other correlations.

Among the identified factors in the design phase, the results show that *''good collaboration within project team''*, *''performance of architect''* and *''performance of ventilation consultant''* are contributing to the client satisfaction. According to the model, *''good result''* of the design phase is also one factor behind client satisfaction. Statistically it can be presented that projects, in which clients perceive that the result of the design phase was good (score 4 or 5), clients are more satisfied with the final result, on average 13.2 out of 15 maximum. This can be compared to 8.8 out of 15 maximum, on average satisfaction for the projects in which the clients perceived that the result of the design was not good (score 1 or 2).

Coefficient	*Factors*
w_1: *1.4652*	x_1: *Project program - follow time plan weekly*
w_2: *1.7597*	x_2: *Project program - good result*
w_3: *0.9827*	x_3: *Pre-design phase - conducting risk analysis*
w_4: *1.1433*	x_4: *Pre-design phase - well thought-through description of goals*
w_5: *1.9769*	x_5: *Pre-design phase - analyse of suitable type of contract*
w_6: *2.0534*	x_6: *Design phase - good collaboration within project team*
w_7: *1.1074*	x_7: *Design phase - good result*
w_8: *1.3236*	x_8: *Design phase - performance of ventilation consultant*
w_9: *0.8048*	x_9: *Design phase - performance of architect*
w_0: *2.9882*	

Table 4.2 List of the factors in pre-construction phase that have the highest correlation with client satisfaction

4.1.2 Result from contractor survey

In the survey answered by the contractor representatives, 19 input variables are associated with the pre-construction phase. Here it is presented which of these input variables are identified as significant factors for the three different outputs; cost variance, time variance and contractor satisfaction.

4.1.2.1 Cost variance factors

In the result from the statistical analysis of the contractor representatives study, eight of the 19 input variables in the pre-construction phase appear to be factors correlated to project cost variance. Two factors have the impact of causing cost savings and lower cost variance, which is a decreasing effect and six have the opposite impact as increasing effect on cost variance. Client *''planning the project well''* and showing *''good decision-making''* are two client attributes that have a decreasing effect on the cost-variance. The better the client plans the project and the better the client's decision-making is, the more cost efficient the project gets according to the statistical model. It is not found whether the type of client, public or private, have any correlation with cost variance. On the other hand, it is obtained that the type of contract has influence on the cost variance. Two different types of contracts are represented in the statistical model: design-and build contracts and traditional contracts. In the data set, it can be observed that design-and build projects show an average cost variance of 3.65%, while *traditional contracts* are 7.55% over budget on average. Traditional contract is therefore an increasing factor in the statistical model. Moreover, it is also found that the input *''own company - production people involved in choosing subcontractors and suppliers''* can be correlated to a higher cost variance. ''*Limited time to prepare production''* was also found to have a noticeable correlation with cost variance. The amount of subcontracted office workers hours in relation to in-house office workers hours also positively affects cost variance. Considering the project attributes, ''*technically challenging and advanced''* characteristic is identified as a factor behind cost overrun. A correlation is also found between ''*gross area''* and cost variance; bigger projects have a tendency to show higher deviation in percentage from planned cost.

Coefficient	*Factors*
w_1: -5.8128	x_1: *Client - planning the project well*
w_2: -5.7519	x_2: *Client - good decision-making*
w_3: 4.6225	x_3: *Contract type - traditional contract*
w_4: 5.3978	x_4: *Own company - production people involved in choosing subcontractors and suppliers*
w_5: 3.0754	x_5: *Subcontractors - office workers duration*
w_6: 1.0245	x_6: *Gross area*
w_7: 1.441	x_7: *Project technically challenging and advanced*
w_8: 1.7702	x_8: *Limited time to prepare production*
w_0: 4.012	

Table 4.3 List of the factors in pre-construction phase that have the highest correlation with cost variance

4.1.2.2 Time variance factors

In the result of the statistical analysis, four factors have negative impact on time variance whereas two factors have positive impact on time variance. Among the different client attributes, there are two that have influence on time variance. *''Client - planning the project''* well is an attribute that, if fulfilled, statistically decreases time variance. In other words, if the client plans the project well, it contributes to the likeliness that the project will be finished on time. It is also shown that encouraging innovation by the client also leads to lower time variance. Moreover, two increasing factors are found in the analysis of the contractor representatives survey. If time is limited for the contractor to prepare for production, it is likely that the percentage of time variance increases and that the project construction duration extends more than expected. *Technically challenging and advanced* project attributes also contributes to higher time variance. A correlation is also established between *percentage of prefabrication* and time variance. In the model, the coefficient for this factor is negative and projects with a high percentage of prefabrication (66-100%) on average overruns the time plan with 1.94% compared to 5,93% for projects with low (0-33%) percentage of prefabrication. The proportion between the duration of own craftsmen, subcontracted craftsmen, own office workers and subcontracted office workers is also found to have influence on time variance. The duration of own craftsmen stands for 30,68% of the total working hours on average. A small correlation is found between time variance and *''own company production people involved in choosing subcontractor and supplier''*. The factor was not included in the model.

Coefficient	*Factors*
w_1: -5.7689	x_1: *Client - planning the project well*
w_2: -3.9758	x_2: *Client - encouraging innovation*
w_3: 2.0444	x_3: *Limited time to prepare production*
w_4: 12.4581	x_4: *Project technically challenging and advanced*
w_5: -4.4745	x_5: *Percentage of prefabrication*
w_6: -10.296	x_6: *Own crew - craftsmen duration*
w_0: 7.0614	

Table 4.4 List of the factors in pre-construction phase that have the highest correlation with time variance

4.1.2.3 Contractor satisfaction factors

Among the 19 input variables, seven factors appear to have noticeable correlation with contractor. According to the result of the statistical analysis, it is important that the client show *good decision making*, presents *clear goals of the project* and *plan the project well*. *Encouragement of innovation* from the client, is also one factor that statistically makes the contractor more satisfied. No correlation was found between *type of contract* and the satisfaction of contractor representatives. Two factors associated to the main contractor's organisation are found in the model. It is shown that involving the *production people participate in choosing subcontractors and suppliers* will contribute to a higher satisfaction for the main contractor representatives. Respondents that have answered that production people were involved to a high degree in choosing subcontractors, are on average more satisfied; 11.77 out of 15 compared to 10.76 out of 15 in projects where production people were poorly involved. *The amount of subcontracted office workers months* is having effect on the degree of satisfaction. The result from the statistical analysis also shows that *support of the administrative work* from main contractor's organisation is a factor that can be correlated to the satisfaction of contractor representatives. Respondents that perceive that the own company highly supported the administrative work, generally answer higher satisfaction score, 11.38 compared to 10.35 out of 15. One more variable that is correlated to satisfaction, is to what degree the *project is prioritized by the own organisation*. This variable is not selected to be included in the final model, since the correlation with satisfaction is not significant enough for predictions to be based upon it. Likewise, the factor about the duration of own craftsmen in relation to own office workers and subcontracted craftsmen and office workers, is also found to be a correlating factor. It was also not included in the model, for the same reason.

Coefficient	*Factor*
w_1: *1.0347*	x_1: *Client - planning the project well*
w_2: *1.3854*	x_2: *Client - encouraging innovation*
w_3: *0.553*	x_3: *Client - good decision making*
w_4: *2.2765*	x_4: *Client - clearly presenting goals of the project*
w_5: *1.104*	x_5: *Own company - production people involved in choosing subcontractors and suppliers*
w_6: *1.1117*	x_6: *Own company - support of administrative work*
w_7: *2.253*	x_7: *Subcontractors - office workers duration*
w_0: *6.2312*	

Table 4.5 List of the factors in pre-construction phase that have the highest correlation with contractor satisfaction

4.2 Construction phase

The client survey considering the construction phase consists of questions regarding production conditions and the main contractor performance. The contractor survey considering the construction phase consists of questions about the client attributes, quality of consultants documents, piping and ventilation subcontractors performance, time plan, collaboration, external factors such as weather conditions and also site conditions such as groundwork and the extent of blasting work.

4.2.1 Result from client survey

The part of the client survey concerned with construction phase consists of 11 input variables. Here it is presented which of these 11 input variables are identified in the statistical analysis, as significant factors for the two different outputs; cost variance and client satisfaction.

4.2.1.1 Cost variance factors

The statistical analysis of the client survey shows that six factors in the construction phase have little to high correlation with the cost variance. Four factors with high correlation were selected for the final model: *''construction phase - good collaboration within the project team at site''*, *''construction phase - absence of disturbances and problems''*, *''main contractor - good cooperation''* and *''construction phase - good result''*. All four being weighted by negative coefficients, showing that they have a decreasing effect on cost variance. For instance, projects in which the client had answered that the collaboration at site is score 1 or 2 out of 5, show 12.11% cost variance on average. On the other hand, projects with score 4 or 5 show an average cost variance of -1.71%. For projects caused by a lot of disturbances and problems, score 1 or 2, the average cost variance is 6.99%. Projects with a low disturbance rate, score 4 or 5, cost variance tend to be -2.03% on average. Two other factors, *''main contractor - innovative production method''* and *''main contractor - tidy construction site''*, have a small correlation with the cost variance and were therefore not selected for the final model.

Coefficient	*Factor*
w_1: -5.9454	x_1: *Construction phase - good collaboration within the project team at site*
w_2: -5.4604	x_2:*Construction phase - absence of disturbances and problems*
w_3: -3.799	x_3: *Construction phase - good result*
w_4: -4.2777	x_4:*Main contractor - good cooperation*
w_0: 15.2651	

Table 4.6 List of the factors in construction phase that have the highest correlation with cost variance

4.2.1.2 Client satisfaction factors

In the statistical analysis of the client survey, six out of 11 variables in the construction phase are identified to be correlated to client satisfaction. Two factors, *''Main contractor - good cooperation''* and *''Main contractor - managed the design phase well''*, are not selected for the final model, since the building of the model requires to have the highest correlating input variables in order to obtain the best accuracy. The four selected factors have clear correlation to client satisfaction. The result from the analysis of the clients' survey show that in the production phase, *''construction phase - good collaboration within project team at site''*, *''construction phase - absence of disturbances''* and *''construction phase - good result''* are factors that are especially connected to client satisfaction. Furthermore, the production method is also found to be influential. *''Main contractor - innovative production method''* is also one variable that can be identified as a factor for client satisfaction. Results show that projects, in which the main contractor has established an innovative production method, clients are more satisfied with the project, 13,44 compared to 10,57 out of 15 for projects were main contractor did not use, or poorly established, an innovative production method.

Coefficient	*Factor*
w_1: *2.2539*	x_1:*Construction phase - good collaboration within project team at site*
w_2: *1.6259*	x_2:*Construction phase - absence of disturbances and problems*
w_3:*2.8202*	x_3:*Construction phase - good result*
w_4: *1.1718*	x_4: *Main contractor - innovative production method*
w_0: *6.5392*	

Table 4.7 List of the factors in construction phase that have the highest correlation with client satisfaction

4.2.2 Result from contractor survey

In the survey answered by the contractor representatives, 25 input variables are associated with the pre-construction phase. Here it is presented which of these input variables are identified as significant factors for the three different outputs; cost variance, time variance and contractor satisfaction.

4.2.2.1 Cost variance factors

Analysis of the 25 input variables of the construction phase in the contractor representatives survey, lead to the identification of five especially important factors behind cost variance. All factors in this formula are weighted by negative coefficients. Fulfilling these factors lead to a lower the cost of construction cost. Factors belonging to the performance of the client, consultants and project team are represented. First, the factor about client communication *''client - giving clear message on time''* is recognized to be correlating to the cost variance. In the result of the statistical analysis, the performance of the piping consultant is found to be of more importance than the other consultants. Statistically, the factor *''quality of documents - piping consultant''* appear to decrease cost variance if fulfilled. Performance of the piping work is also described by the two factors *''piping subcontractor - result as expected''* and *''piping subcontractor - result according to specifications in contract''*. The factors appear to be correlated to cost variance, however too feeble to be included in the formula. Likewise, the factor *''quality of documents - ventilation consultant''*, representing the performance of the ventilation consultant, is not included in the model, even though a small correlation is found.

It can be observed that in the construction phase, *''construction phase - good collaboration within project team at site''* and *''construction phase - follow time plan weekly''* also have the impact on decreasing the cost of construction and lowering the percentage between planned and actual cost. The situation in the construction phase is also described by the factor *''construction phase - absence of disturbances and problems''*. Projects with smooth construction processes without interruptions and complications, have better cost performance. On average, projects that were widely caused by disturbances and problems (score 1 or 2) have 11.98% cost variance on average compared to 2.59% for projects with less complications (score 4 or 5).

Coefficient	*Factor*
w_1: -4.4299	x_1: *Client - giving clear message on time*
w_2: -0.4374	x_2: *Construction phase - good collaboration within project team at site*
w_3: -7.1309	x_3: *Construction phase - absence of disturbances and problems*
w_4: -2.8115	x_4: *Construction phase - follow time plan weekly*
w_5: -6.9273	x_5: *Quality of documents - piping consultant*
w_0: 18.9767	

Table 4.8 List of the factors in construction phase that have the highest correlation with cost variance

4.2.2.2 Time variance factors

Out of the 25 construction phase related input variables from the contractor survey, five variables are identified to have significant influence on time variance. Four factors have decreasing effect on time variance while one factor, ''construction phase - extensive blasting work'', has increasing impact (w_3:7.2356). Projects that did not, or only performed a small amount of blasting work, have an average time overrun of 2.38%. This can be compared to 9,16% time overrun on average for projects that performed blasting work to a high extent. Among the four decreasing factors, two of them are considering the performance of the architect and the ventilation subcontractor. It is found in the statistical analysis that good quality of the architect's drawings decreases time variance (w_1: -7.6376). Projects with good or very good drawings made by the architect have an average time variance of 0.81%. In comparison, the average time variance is 8.24% for projects where the contractor respondents have answered that the architectural drawings were bad or very bad. A correlation is also found between time variance and the quality of documents provided by the electrical engineer, structural engineer and piping consultant. The factors were however not included in the formula, since the correlations were smaller than for other factors. It is found in the statistical analysis that ventilation subcontractor using innovative production methods can contribute to a lower time variance (w_2: -11.0663). Projects, in which the ventilation subcontractor have been considered to be innovative, show 0.86% in average time variance. The average time variance for all projects in the contractor survey is 3.53%. Also, a small correlation is found for ''ventilation subcontractor - kept agreed time frame''. Moreover, the two factors ''construction phase - absence of disturbances and problems'' and ''construction phase - follow time plan weekly'' are found to be influencing the output.

Coefficient	*Factor*
w_1: -7.6376	x_1: *Quality of documents - architect*
w_2: -11.0663	x_2: *Ventilation subcontractor - innovative production methods*
w_3: 7.2356	x_3: *Construction phase - extensive blasting work*
w_4: -9.5326	x_4: *Construction phase - absence of disturbances and problems*
w_5: -7.1495	x_5: *Construction phase - follow time plan weekly*
w_0: 25.0187	

Table 4.9 List of the factors in construction phase that have the highest correlation with time variance

4.2.2.3 Contractor satisfaction factors

Six of the 25 variables in the construction phase are identified to be influencing the contractor satisfaction. Following six performance influencers all have positive coefficients and hence increasing effect on contractor satisfaction in the model. In the result of the statistical analysis, it is seen that cooperation and collaboration between participants is influencing contractor satisfaction. *''Construction phase - good collaboration within project team at site*'' and *''client - good cooperation in project''* are such factors. *''Client - good cooperation in project''* is the most apparent performance influencer (w_1: 2.8794), followed by *''construction phase - absence of disturbances and problems''* (w_5: 2.8596). Absence of disturbances and problems is also complemented by *''construction phase - followed time plan weekly''* and *''ventilation subcontractor - work to minimize interruptions and problems''*. The performance of the ventilation contractor is also represented by *''ventilation contractor - result according to expectations''*. Some factors are discarded since the correlation with the output is minimal. These are *''quality of documents - piping consultant''*, *''piping subcontractor - innovative production''*, *''client - giving clear message on time''* and *''own company - enough provision crew''*.

Coefficient	*Factor*
w_1: 2.8794	x_1: *Client - good cooperation in project*
w_2: 1.3673	x_2: *Ventilation subcontractor - work to minimize disturbances and problems*
w_3: 1.0437	x_3: *Ventilation subcontractor - result according to expectations*
w_4: 0.3171	x_4: *Construction phase - good collaboration within project team at site*
w_5: 2.8596	x_5: *Construction phase - absence of disturbances and problems*
w_6: 1.2676	x_6: *Construction phase - follow time plan weekly*
w_0: 4.6173	

Table 4.9 List of the factors in construction phase that have the highest correlation with contractor satisfaction

4.3 Testing the adequacy of the models

Table 4.10 summarizes the root mean square errors of each model. Each model has a root mean square error, also called standard deviation. Comparing the results, it can be seen that the lowest standard deviation for construction cost predictions is 9.7913%. Secondly, the RMSE for predictions for construction cost variance that is derived from pre-construction information is 10.0592%. Time variance comes in the second place, the highest error for this performance measure is 16.0156%. Satisfaction comes in the third place when it comes to predictability. Predicting the satisfaction of contractor from data gathered from the pre-construction phase has an error rate of 16.646%.

Model	*Standard deviation of residuals (RMSE)*
Pre-construction, contractor - cost variance	10.0592%
Construction phase, contractor - cost variance	9.7913%
Pre-construction phase, contractor - time variance	16.0156%
Construction phase, contractor - time variance	15.3446%
Pre-construction phase, contractor - satisfaction	16.646%
Construction phase, contractor - satisfaction	15.0706%
Pre-construction phase, client - cost variance	12.7671%
Construction phase, client -cost variance	13.0603%
Construction phase, client - satisfaction	13.528%
Construction phase, client - satisfaction	12.458%

Table 4.10 Testing the adequacy. A comparison is made of the error of each model.

Statistical analysis of client survey – error distribution

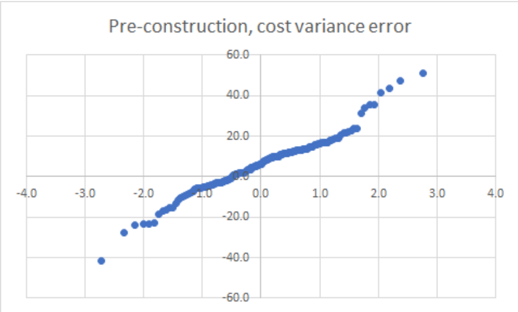

Figure 4.1 Quantile-Quantile plot chart

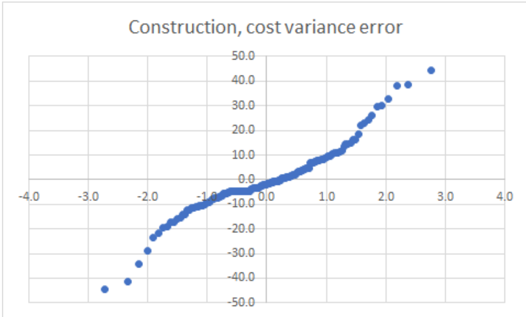

Figure 4.2 Quantile-Quantile plot chart

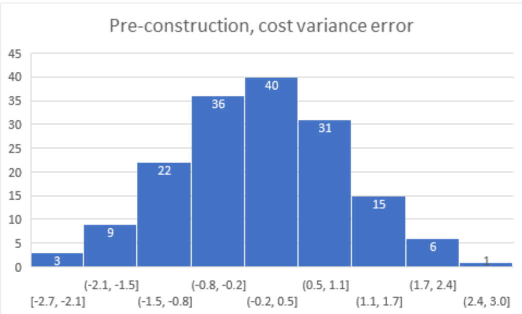

Figure 2.3 Histogram

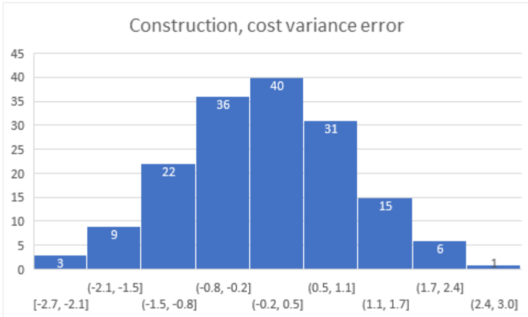

Figure 4.4 Histogram

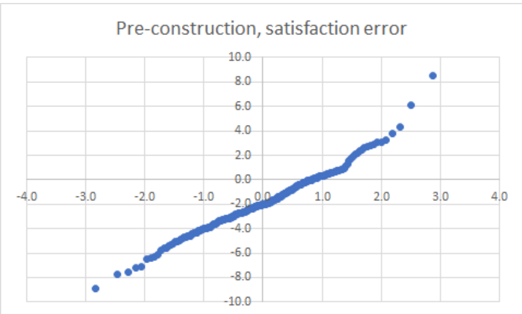

Figure 4.5 Quantile-Quantile plot chart

Figur 4.6 Quantile-Quantile plot chart

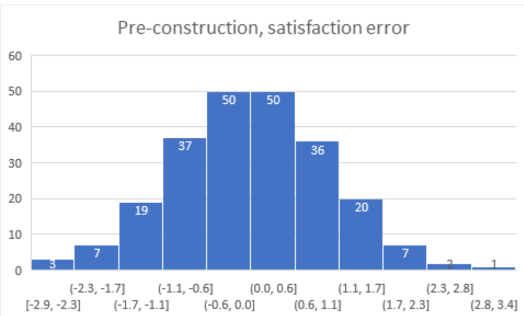

Figure 4.7 Histogram

Figure 4.8 Histogram

Statistical analysis of contractor survey – error distribution

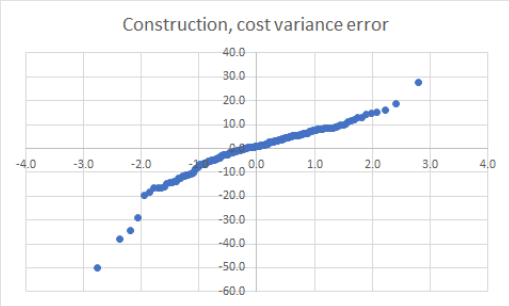

Figure 4.9 Quantile-Quantile plot chart

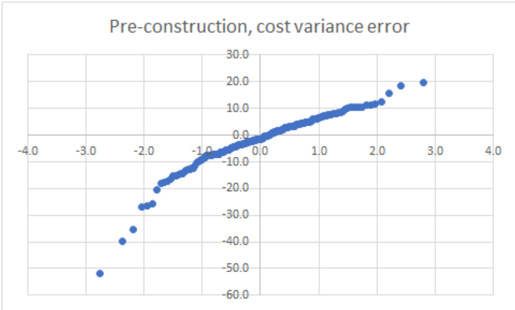

Figure 4.10 Quantile-Quantile plot chart

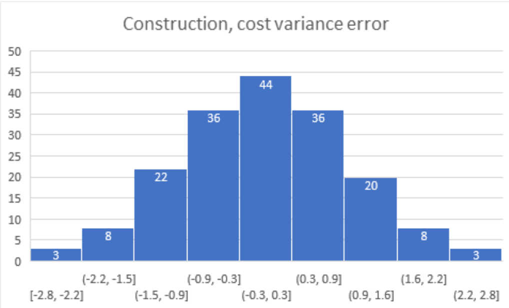

Figure 4.11 Histogram

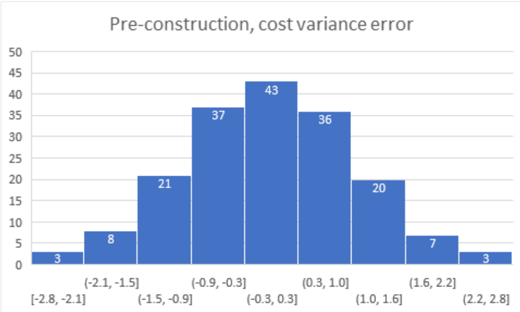

Figure 4.12 Histogram

Figure 4.13 Quantile-Quantile plot chart

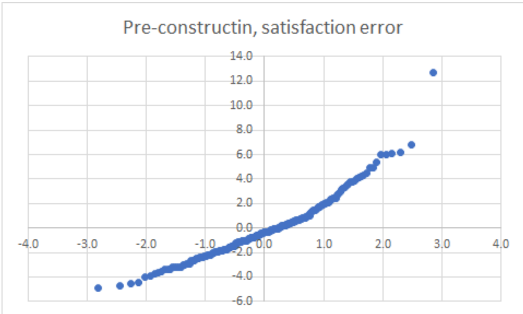

Figure 4.14 Quantile-Quantile plot chart

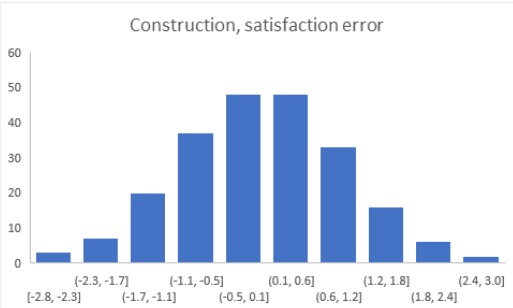

Figure 4.15 Histogram

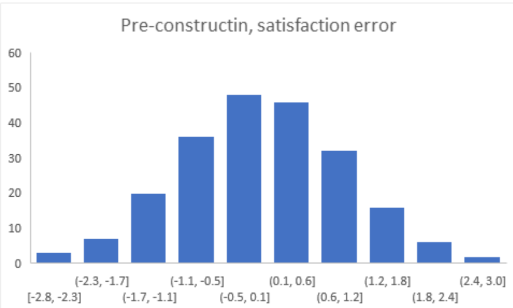

Figure 4.16 Histogram

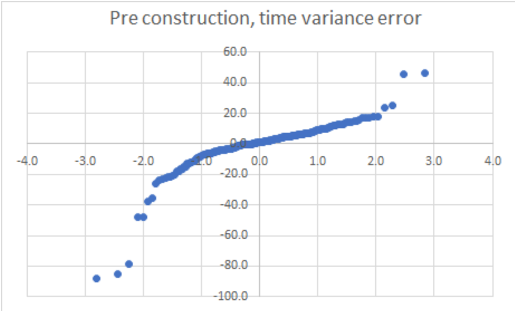

Figure 4.17 Quantile-Quantile plot chart

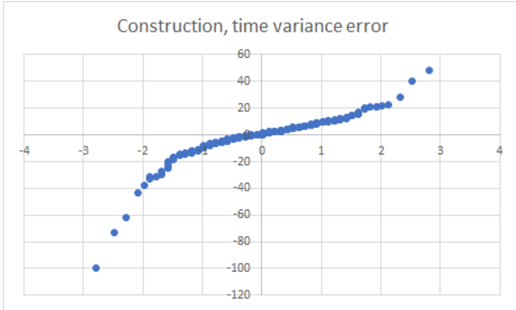

Figure 4.18 Quantile-Quantile plot chart

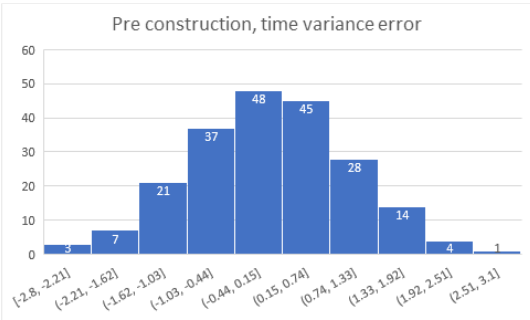

Figure 4.19 Histogram

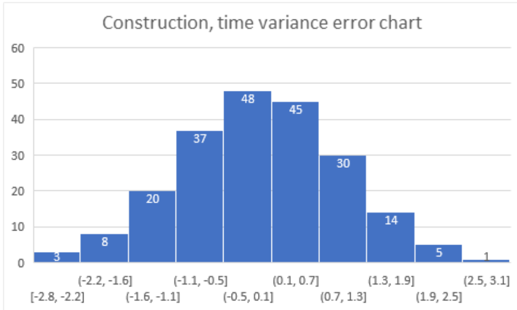

Figure 4.20 Histogram

5. Discussion

In this chapter, it is discussed how the findings of the statistical analysis are consistent with the findings made in the literature review. First, general project conditions considering project attributes and external factors are discussed. Thereafter, the influence of the different actors in the project is discussed. Finally, the discussion aims to evaluate whether performance actually can be predicted and how the prediction model presented in this report can be used. Additional references are added to strengthen the discussion and to find explanations for the identified phenomenons.

5.1 General project conditions

The result of the regression analysis shows that project attributes are influential on time variance and cost variance, while no pattern or correlation between project attributes and client- or contractor satisfaction is found that signify the impact project attributes have on satisfaction. It is found in the statistical analysis of the client survey that cost variance is influenced by project attributes such as “main procurement criteria - reference project”, “Main procurement criteria - environmental aspects” and ‘’Type of contract = trades contract”. Cost variance in the contractor statistical analysis is influenced by “Contract type = traditional contract”.

It is evident from the statistical analysis of both the client and contractor survey that traditional contracts are correlated with cost overrun. As argued by Naoum and Egbu (2016), it is known that the design process in traditional contracts is more influenced by the designer, hence the contractor’s task is to construct what the designer has developed without the involvement of the contractor. Traditional contracts create fragmentation between the design and the building process which makes rework and changes to be done more frequent during construction process and causing higher costs for the projects.

In the statistical result, ‘’environmental aspects’’ or ‘’reference project’’ as main procurement criteria have positive impact on cost underrun, while ‘’earlier collaboration’’ or ‘’specific competence of the contractor’’ do not show any significant impact. This does not mean that ‘’environmental aspects’’ or ‘’reference project’’ are more preferable than the other procurement criterias. There may be underlying reasons to why procurement criteria such as specific competence is not identified as factors behind cost variance. A contractor with specific competence, that can provide better quality on the final product and save cost and time, will not be identified in the correlation-based feature selection if such a contractor is more expensive. According to Alzahrani and Emsley (2013), contractors with relevant experience from previous projects can have positive impact on the success of a project.

Regarding ‘’environmental aspects’’ as a procurement criteria, it has been seen from the literature that this type of requirement is associated with the procurement method.

The literature review about sustainability made by Naoum and Egbu (2016), shows that it is agreed in the literature that environmental regulations require iterations at the design stage and highly suggests for more collaboration between the professionals involved in projects. Therefore, one can argue that the emphasis on procuring a contractor according to environmental criterias can increase project performance because carrying out a project in accordance with environmental requirements, calls for more collaboration which lead to an overall performance gain. In the literature review made by Naoum and Egbu (2016), it is also argued that traditional contracts have the characteristic of limiting the opportunity for the contractor to influence the design and planning and therefore, it is more difficult for the contactor and his partner suppliers to influence or achieve the overall sustainability targets. Therefore, traditional contracts are not the best contract when having environmental aspects as procurement criteria and it can be argued that design- and build contracts are better organizational environments to achieve sustainability requirements. The projects analysed in this report, that have embraced environmental procurement criterias have had design- and build contract.

Moreover, cost variance in the contractor statistical analysis is influenced by “Gross area” and” project technically challenging and advanced”. Time variance is influenced by the “Percentage of prefabrication”, ‘’extensive blasting works’’ and “project technically challenging and advanced”. The amount of prefabrication has a decreasing effect on time variance while the other two attributes have increasing effect on time variance. Josephson (2013) found that projects with prefabricated structures are generally more expensive than in situ-built structures. In the model presented in this report, the amount of prefabrication is a performance influencer for time but not for cost. It is therefore reasonable to believe that prefabricated constructions save time but does not lead to lower costs. It can be concluded that a technically challenging and advanced project have a tendency to have a high time- and cost variance. As stated by Lam and Wong (2009), construction buildability is highly affecting project performance (time, cost, quality and safety). Also, the performance influencer ‘’extensive blasting works’’ is found to be affecting time variance. A conclusion can be made that the complexity of the project, as presented by Chan, Scott and Chan (2004), is a factor behind project performance. “Gross area” is also one project attribute that is shown to be important for cost variance. The result of the statistical analysis indicates that the cost variance adheres to the gross area. In other words, larger projects do on average have higher deviation from planned cost than smaller projects. This indicates that it is more difficult to estimate the cost of big projects.

The statistical analysis of the client survey shows that ‘’Design phase - absence of disturbances’’ is important to the cost variance, it has a decreasing effect. ‘’Absence of disturbances’’ is shown to be important in the construction phase for cost variance and the satisfaction of the client and contractor representatives. The factor ‘’absence of disturbances’’ does not aim at any specific disturbance. An analysis of the dataset, made by Koch and Lundholm (2018) shows that, among swedish construction projects, weather is the biggest source of interruptions, followed by ground conditions and design flaws. It is reasonable to believe that ‘’absence of disturbances’’ follows this distribution. Weather conditions and ground conditions appearing to be some of the most important factors behind project performance according to Swedish studies as well as international studies. However, these factors are not distinguished in the statistical analysis of this report. The reason is that the output (cost-, time variance or satisfaction) is not changing accordingly with the complication of groundwork or bad weather. For instance, projects where the contractor representatives have answered that the projects were disturbed by bad weather conditions, the average cost variance is 5.69%. Projects with no disturbance by bad weather, the average cost variance is 4.89%. Evidently, there is no distinct difference between the two states of condition. Koch and Lundholm (2018) indicates that weather conditions is a broad problem throughout the swedish construction industry. Bad weather conditions is an element in projects that perform bad but also in projects that perform well. This is the reason to why this factor is not dominating in the statistical analysis, even though it is a broad and common problem.

5.2 Project organisation

The project organisation consists of the main contractor, consultants, subcontractors, suppliers and the client itself. The performance and actions of the parties are found to be correlated with project performance. What actions and attributes of the different parties that are correlated with project performance are discussed in this section. Worth to notice is that there are in many cases underlying reasons behind the actions and performance of the parties. This is considered in the discussion, that aims to provide an understanding of the empirical results.

5.2.1 Client

It is shown in the result of the statistical analysis that the client’s role is influential for both pre-construction and construction phase, as well as the influence expands to the three key performance objectives; cost, time and satisfaction. It has been concluded that ‘’client planned the project well’’ has a decreasing influence on cost variance and time variance as well as increasing influence on contractor’s satisfaction. According to Walker (1995) and Chan, Scott and Chan (2004), the client’s ability to collaborate with the design- and construction and contribute to design and construction are factors behind project success. It is reasonable to believe that the performance influencer ‘’client planned the project well’’ can represent these success factors in the statistical analysis. Moreover, ‘’client showing good decision making’’, ‘’encouraging innovation’’ and ‘’clearly presenting the goals of the project’’ are all affecting project success in the pre-production phase in one or more key performance indicator. On the other hand, construction phase is affected most by ‘’client achieved good cooperation in project’’ and ‘’giving clear message on time’’. Following quote out of the book by Boyd and Chinyio (2008) provides a general understanding for the problems facing the construction client:

‘’The problem of engaging the industry as a whole, the problem of telling the industry what is wanted, the problem of ensuring the design addresses what is wanted, the problem of paying for this design, the problem of employing a contractor in order to make the design real, the problem of facing the changes in conception and the problem of sorting it all out at the end.’’ (p.16)

The findings in the statistical analysis is consistent with the findings in the literature. The performance influencer ‘’client achieved good cooperation in project’’ can be related to what in the literature review is mentioned as the client’s ability to collaborate and manage the relationships in the project, see Jha and Iyer (2006), Walker (1995). Odeh and Battaineh (2002) found that slow decision-making by the client is one of the most common causes behind time overruns and Le-Hoai, Dai Lee and Lee (2008) found that changes made by the client is one of the most common causes behind time and cost-overrun. Cherns and Bryant (1984) problemify these attributes of the client. The authors argue that the client is usually a complex organisation with several sub-stakeholders and that problems and unresolved conflicts within this organisation lead to dilemmas in the construction phase such as delays and changes in design. The extensive role of the client in both pre-construction and construction phase can be explained with the perspective about the client discussed by Boyd and Chinyio (2008). In the book, it is explained how complex it can be to understand the requirements of the client. First, buildings are not just expected to be technical structures, they are also social, financial and symbolic purposes, which create a list of different needs. Moreover, the client needs can also be explicit and implicit, in other words there are requirements that client express clearly and other that he conceals either consciously or unconsciously. Usually, the briefing stage is when these requirements are to be discovered and negotiated (Boyd and Chinyio, 2008). Ann et al. (2007) argue that the project briefing is important for the success of a construction project. However, as explained by Boyd and Chinyio (2008), the complex set of multiple important stakeholders, such as an owner of the building and a user, makes the negotiation in the briefing stage complicated because the various and, in some cases, conflicting needs. The authors call for an implementation of the model presented by Barrette and Stanley (1999), which suggests a modern way in briefing which consists of five solutions in an interactive engagement process. It includes the need for client empowerment, appropriate user involvement, wider involvement of the project teams in the process and visualization for improved communication.

The result of the statistical analysis of the client survey show ‘’Project program - compliance with time plan’’, ‘’project program - good result’’ have correlation with the client satisfaction and will increase the client satisfaction if fulfilled. It is also shown that ‘’pre-design phase - well thought through description of goals’’, ‘’pre-design phase - analyse of suitable type of contract’’ have positive influence on the client’s own satisfaction in a project. The variable ‘’pre-design phase - investigation of plot’’ has decreasing effect on cost variance if well performed. The client’s ability to contribute to and manage the design stage is highlighted in the theory. It is reasonable to believe that this is affected by the client experience in managing construction projects as stated by Chan, Scott and Chan, (2004). The performance in the design phase is represented several times by different variables. ‘’Design phase - compliance with time plan’’, ‘’Design phase - good collaboration within project team’’ and ‘’Design phase - good result’’. This is in compliance with the findings made by Koch and Lundholm (2018), which show that problems with the design, either design flaws or problems in the design work, are sources behind disturbances in construction projects. The performance of the architect is discussed in section 5.2.3.

5.2.2 Main contractor

The statistical analysis shows that the cost of construction is affected by two variables that are related to the contractor, “own company production people involved in choosing subcontractors and suppliers” and “limited time to prepare for production”. Satisfaction of a contractor is affected by “own company production people involved in choosing subcontractors and suppliers” and “own company support administrative work”. Time variance is affected by “limited time to prepare for production” and “own craftsmen duration”. These are aspects according to the distinction of pre-construction phase. In the statistical analysis, it is shown that client satisfaction is affected by ‘’contractor innovative production method’’, “production at site-good result” and “good collaboration on site”. It is observed in the statistical analysis that cost- and time variance as well as contractor satisfaction are all affected by “compliance with timeplan” and “absence of disturbances”. For cost variance, “good collaboration within project team” is influential.

In the literature review, it is shown that site management and contractor having adequate experience are important factors behind project success (Sambasivan and Soon, 2007; Odeh and Battaineh, 2002). Poor site management is considered to be one of the most common factors behind unsatisfactory performance in construction projects (Jha and Iyer, 2008). This is compatible with the results found in the statistical analysis in terms of good production on site and good collaboration on site aspects. In the literature, problems with schedule were found to be influencing project success (Atalla and Hegazy, 2003). This is also presented in the result of the regression analysis. The factor “compliance with timeplan” is found to be affecting the three outputs in the analysis. The factor “compliance with timeplan” increases satisfaction as well as it lowers cost and time measures. Moreover, it is generally known that top management support is a driver for the site manager to increase the quality of the project and to strive for success. It is also generally known that site managers have a large load of administrative work that also consumes time. When the company provides support for administrative work, site managers seem to work more effectively. One important factor that has been shown in the empirical results of the regression analysis is that when the duration of contractors’ own craftsmen is higher, the time variance is lower. Previous study by Alzahrani and Emsley (2013) show that the adequacy of labour resources has great impact on project success. As discussed in their literature review, researchers generally agree to that the human resource aspect is crucial because of the roles people play in projects.

The last aspect to be discussed in the contractor section is “own company production people involved in choosing subcontractors and suppliers”. It is important to note that this aspect has a positive influence on the site managers’ satisfaction, as well as it can increase cost variance. One would say that it is more logical to lower the cost by choosing subcontractors and suppliers that people in the production prefer to work with. This might be explained with interplay of price and trust when choosing subcontractors. The study of Hartmann, Ling and Tan (2009) finds that contractors are more likely to value price quality when they choose subcontractors or suppliers and even if they have a previous trusted relationship with previous suppliers, they would choose the lower cost and more quality component one. This finding is inconsistent with the regression results, it is assumed that the involvement of choosing subcontractors and suppliers might affected to price of the procurement because of previous collaboration preferences.

5.2.3 Consultants and subcontractors

As mentioned before, design flaws are one of the most common disturbances in the Swedish construction industry, hence the result of the project brief (project program) and design phase is vital for project performance. Several performance influencers are identified in the project brief (project program) and the design phase. For instance, the absence of disturbances, compliance with time plan, good collaboration within the project team and good result of the phases. It can be argued that the architect is one important actor in the project brief and the design phase and that the performance of the architect can be a source for disturbances in these phases. The statistical analysis of the client survey, shows that a well performing architect in the design phase, contribute to lower cost variance as well to higher client satisfaction. Also, the statistical result of the contractor survey shows that the quality of the architect’s documents is important for the time variance. In the construction phase, the architect’s drawings are put to test and the contractor is the one to judge whether the architect’s performance is adequate or not. In the literature study, it is found that the skills and performance of the architect is important for project success. This lead to following question: Is it about skills and performance? According to Boyd and Chinyio (2008), often architects consciously skip to complete the design and leave out some things to be solved at the construction site. Which lead to time overrun and higher time variance, as can be observed in the statistical analysis in this report. However, the architect is not to be blamed for this. As stated by Boyd and Chinyio (2008), these design flaws are the result of hard competition caused by clients’ focus on lower cost and therefore the architect provide incomplete drawings to keep the cost down. The conclusion that can be made from this is that it is the architect’s responsibility to perform what the client asks for. However, the client is responsible for giving the architect clear directives and be deliberate to pay for performed work. Construction buildability and complexity is also found to be important in the literature review as well as in the statistical analysis. It could be argued that the architect is one of the parties that are responsible for this. However, Boyd and Chinyio (2008) argues that it is a myth that ‘’architects go for creative designs at the expense of the clients’ needs.’’ (p.16). The conclusion that can be made from this reasoning is that it is not only about the skills of the architect. It is important that the client and architect agree on the design, the payment and what is included in the design work, so that the architect performs as desired. It is reasonable to believe that this communication and decision making, is affected by the complex nature of clients as discussed by Cherns and Bryant (1984).

The discussion about the role of the architect can also had about the ventilation and piping consultant. How they perform is important for the collaboration, the level of disturbances and the result of the design phase. For instance, it is found in the statistical analysis that the performance of the ventilation consultant in the design phase is important for the client satisfaction. Thereafter, the consultant’s performance is reflected in the quality of drawings, which can also be a source of disturbances at the construction site. According to the statistical analysis of the contractor survey, the quality of the piping consultant documents is one of the influential factors behind cost variance. In general, well performing projects, considering cost variance, have high score on the quality of the piping consultant’s drawings. Thereafter, it can be assumed that the quality of the drawings is reflected in the performance of the ventilation subcontractor’s performance, since its role is to execute what is made in the drawings. In the statistical analysis, the result show that the performance of the ventilation subcontractor is influential to contractor satisfaction. Ventilation subcontractor worked to minimize interruptions and result was as expected are such factors that are highly correlated with contractor satisfaction. Ventilation subcontractor using innovative production methods is one factor that is correlated to construction time variance, according to the result of the statistical analysis. The duration of “subcontracted office workers” in relation to own office workers, is also found to be correlated to both construction cost variance and contractor satisfaction.

5.3 Predicting performance

The purpose of the prediction models developed using linear regression is to provide a reference for practitioners to improve the performance if it is expected to have unsatisfactory outcome, such as cost overrun, or time overrun. However, it is difficult to know exactly how the models work in the industry, especially if it is used as a consultancy tool for project teams. If undesired project outcome is predicted and corrective actions are taken, the project outcome has the possibility to improve and therefore it will be hard to assess the predictability of the models. One way to test the predictability of the model is to collect another data set with approximately same size and thereafter test the models. Having a second data set for testing would create more confidence about the utility and benefit of the model. Cross validation as has been shown in the method as a valid way to develop the most generalized model (Olive, 2017; Witten et al., 2011). It is a reliable method to test the models since it enables to build and test models in order to give the most generalized one with the lowest root mean square error, which means the model with the best accuracy. The error rates presented in table 4.10, show that the best prediction error is the one associated with predicting contractor cost variance of construction (9.7913%). Secondly, the prediction for construction cost based on the pre-construction predictors variables is also performing better than other models, with a root mean square error of 10.0592%. It has been shown that predicting satisfaction is the hardest. This is quite expected because it is harder to know what makes people happy since every person might have different criteria when deciding how happy they are with a project.

It is also worth mentioning that the outputs of the analysis, cost variance and time variance, are spread over a large span of numbers. The highest cost variance in the client perspective data set is +43.68% cost variance and the lowest is -49%. Cost variance in the contractor's survey also ranges from +60% to -20%, time variance ranges from +100% to -55.55%. This makes it more challenging to obtain a precise prediction because the instances are deviating considerably from the average and thus standard deviation will be higher. As a result, the prediction is approximate and not absolute and most likely it gives a very approximate indicator of how well the project would go. However, since the Quantile-Quantile plot graphs show that the error is mostly normally distributed, one can say that 68% of the predictions should have an error rate that is within one standard deviation. The standard deviation of a prediction works as an uncertainty range for the predictions.

Considering the usefulness of the model, it is worth mentioning that the variance of cost or time does not point out the whole picture. Just because time or cost variance is smaller for some projects, it does not automatically mean that they perform better or have higher productivity than other projects. A low variance only tells a project's ability to meet the planned cost and time. A construction project with a low-cost variance, does not mean that it was performing well. A low-cost variance can be the result of inaccurate estimations of the project cost in the early phases. It can be argued that predictability of time and cost could also be considered as a performance output, as mentioned in the literature review in this report. In addition to this, some projects in this report, have low cost variance by nature. For instance, projects with design and build contracts. It can be interpreted that the cost in these projects, by nature, are more easily predicted. However, it does not mean that these projects perform better. However, it is generally believed that construction contractors do not necessarily want to give unnecessary high price offers, since it is often competition during tendering. Therefore, low cost variance does not equal good project performance unless the original tender was low.

6. Conclusion

Revisiting the assumptions made in the problem formulation; it was assumed that projects success or failure can be evaluated using key performance indicators. It was also assumed that performance influencers can be identified and thereafter linked to each performance objective; cost, time and satisfaction. The assumptions have been qualified through the theoretical study and the empirical results. The theory presents multiple key performance indicators and it is found that cost, time and quality are the most common ones. Empirical results show that factors with high correlation with project performance can identified and used to build prediction models.

The report started with describing the vision of Tobias Nordlund, who is the head of the office of Prolog in Gothenburg. He expressed that his vision is to be able to tell his clients that ''if you pay attention to these factors, the outcome of the project will be like this''. This report has identified the most influential factors behind project performance in Swedish construction projects. A read thread has been discovered that is also answering the research question number one.

RQ1: *What are the most important factors affecting the performance in Swedish construction projects*

The empirical part, see chapter 4, is considered to be a sufficient answer for the first research question. The highest correlated factors are identified and reflected upon in the discussion. Looking at the recurrent factors, it is concluded that absence of disturbances is a common factor in the pre-construction as well as in the construction phase. However, the sources of disturbances are not identified. External factors, such as weather conditions, are broadly mentioned in the theory as causes behind bad performance and project failures. In this report, this type of factor is not found to be highly correlated with project performance. The reason is that this is a common causes that is occurring in some way in almost all projects. In the statistical analysis, no significant difference in satisfaction, time- or cost overrun, could be found between projects highly affected by weather and those projects free from disturbances caused by bad weather. The conclusion is that weather is an element in projects that perform bad but also in projects that perform well. In other words, it is a broad and common problem.

The technical aspects of a building are found to be of importance and highly suitable for basing predictions upon. It can be concluded that prefabrication saves time, but not necessarily cost. Technically challenging and advanced projects in terms of buildability, have a tendency to go over time and budget. Extensive blasting work is one factor that is shown to be correlated to time overrun. If a geological examination is made, the need for extensive blasting work would not come as a surprise for the contractor. Projects, in which contractor feel that the client has not planned the project well and that the time for preparing production is limited, are most likely to show unsatisfactory results. Planning the project is key to keeping budget and time schedule as well as client and contractor satisfaction, which can be seen in the statistical result as well as it is highlighted in the theory. Although external factors and technical aspects of a building are important for project success, the most recurring factors behind project performance can be derived to human related factors.

It can be concluded that construction projects tend to be more expensive when traditional contracts are used and there are presumably many reasons behind that observation. The nature of traditional contracts, the gap between design phase and the production of the building is a reasonable explanation. The report also observes that the procurement criterias ‘’environmental aspects’’ and ‘’reference project’’ have positive impact on keeping down the cost variance. Reference project as a procurement criterion can make the client select a suitable contractor which lead to project success. Meanwhile, it is not found that procuring a contractor with specific experience lead to any performance gain. The explanation can be that, a specific contractor that provide a better result is not favoured in the statistical analysis, if the cost for procuring these contractors are higher. It is discussed that environmental aspects as a procurement criteria, demands better collaboration between the parties in the project and hence it leads to overall better performance. The conclusion is that it is not about the contract itself, it is about the environment it creates. A contract that facilitates collaboration and communication between parties will per se contributes to fulfilling or solving several of the other factors behind project performance. Added to this, a common goal the parties can strive for, such as environmental requirements, makes performance likely to be satisfactory for the involved parties.

Factors that are recurrent in the statistical result of the contractor’s survey are client’s ability to plan the project, make decisions and communicate as well as to encourage innovation. These attributes are repeated time after time in the literature. Nevertheless, the identification and acknowledgement of the importance of these attributes also need to be followed up by an investigation about why these are found to be important. It can be concluded that the client is a very important actor for project success, not least in the early stages of a project where the client should be clear about the project scope and to select a suitable type of contract for carrying out the project. Disturbances in the construction phase are often caused by design flaws. Architects and consultants are often blamed for this. However, it is the client’s role to address what is wanted and to pay for the design. It is discussed that architects often leave out things to be solved at site, because of clients’ focus on keeping the cost down. The discussion in this report open up for an understanding about the underlying causes behind client behaviour, so that the actual factors can be identified. It is very common to point out that it is important that the client provide clear goals and show good decision-making, however these are very subjective attributes of a client and presumably more objective attributes can be identified if investigating the client’s organisation more in depth.

The second research question can also be answered by explanations of phenomenons provided in the discussion, but also about the knowledge gained when developing the equations of the regression analysis model.

RQ2: *How can the identification of these factors be used in practice to help practitioners improve their performance?*

To summarise, the report provides statistical evidence for what factors that are important to consider. The identification of the most correlating factors can be used to build prediction models, as done in this report. It is proved that, the obtained models in this report are capable of providing predictions when testing the formulas on the originate dataset. The models can work as a consultancy tool for practitioners in the industry. They can provide an approximate prediction for projects cost and time variance as well as satisfaction. When unsatisfactory outcome is predicted, the factors contributing to that can be identified and potentially improved to achieve better final results in a project. Further studies should undertake to investigate how able the model is to predict the outcome of projects in a new dataset.

7. Suggestions for future research

This research has shown potentials for using data science in the construction industry sector. Swedish construction industry can use similar types of research to drive a new development scheme that is based on data driven models. The mixed approach of quantitative and qualitative studies has the significance of quantifying soft aspects and making them more measurable.

The aim of this report was to find a general ground that is covering all construction projects. However, it was found that it is not possible to establish a full understanding of the single practice or interaction attribute unless it is fully defined or put into detailed lens of objective measures. Especially ‘’collaboration’’, ‘’performance of consultants and subcontractors’’ and ‘’client decision making’’. Therefore, it is advised to conduct research that focus on one particular aspect beside the more general studies. The same approach of this research is applicable in more particular areas. For example, focusing on one process such as design, project brief, site management or digitalization. It is also possible to focus on a certain type of contract to study its features and quantify new measurable performance indicators. Another suggestion is to conduct a study for developing predictive models, as the ones developed in this report, but with larger and more representative dataset that can result in better quality. It is also advised to test different machine learning algorithms such as neural networks or decision trees, to find the most suitable for solving problems associated with the construction management context.

8. References

Alias, Z., Zawawi, E. M. A., Yusof, K., & Aris, N. M. (2014). Determining critical success factors of project management practice: A conceptual framework. *Procedia-Social and Behavioral Sciences*, *153*, 61-69.

Al-Momani, A. H. (2000). Construction delay: a quantitative analysis. *International journal of project management*, *18*(1), 51-59.

Alzahrani, J. I., & Emsley, M. W. (2013). The impact of contractors' attributes on construction project success: A post construction evaluation. *International Journal of Project Management*, *31*(2), 313-322.

Ann, T. W., Shen, Q., Kelly, J., & Hunter, K. (2007). An empirical study of the variables affecting construction project briefing/architectural programming. *International Journal of Project Management*, *25*(2), 198-212.

Attalla, M., & Hegazy, T. (2003). Predicting cost deviation in reconstruction projects: Artificial neural networks versus regression. *Journal of construction engineering and management*, *129*(4), 405-411.

Beatham, S., Anumba, C., Thorpe, T., & Hedges, I. (2004). KPIs: a critical appraisal of their use in construction. *Benchmarking: an international journal*, *11*(1), 93-117.

Blaikie, N. (2009). *Designing social research*. Polity.

Boyd, D., & Chinyio, E. (2008). Understanding the construction client. John Wiley & Sons.

Bryman, A., & Bell, E. (2011). *Business research methods* (3.th ed.). Oxford: Oxford University Press.

Chan, A. P., & Chan, A. P. (2004). Key performance indicators for measuring construction success. *Benchmarking: an international journal*, *11*(2), 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(2), 188-198.

Chan, A. P., Scott, D., & Chan, A. P. (2004). Factors affecting the success of a construction project. *Journal of construction engineering and management*, *130*(1), 153-155.

Cherns, A. B., & Bryant, D. T. (1984). Studying the client's role in construction management. Construction management and economics, 2(2), 177-184.

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*(3), 142-150.

Cox, R. F., Issa, R. R., & Ahrens, D. (2003). Management's perception of key performance indicators for construction. *Journal of construction engineering and management*, *129*(2), 142-151.

De Wit, A. (1988). Measurement of project success. *International journal of project management*, *6*(3), 164-170.

El-Gohary, K. M., Aziz, R. F., & Abdel-Khalek, H. A. (2017). Engineering approach using ANN to improve and predict construction labor productivity under different influences. *Journal of Construction Engineering and Management*, *143*(8), 04017045.

Frank, E., Hall, M.A., & Witten, I.H. (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.

Freeman, M. & Beale, P. (1992). Measuring project success. Project Management Journal, 23(1), 8–17.

Frimpong, Y., Oluwoye, J., & Crawford, L. (2003). Causes of delay and cost overruns in construction of groundwater projects in a developing countries; Ghana as a case study. International Journal of project management, 21(5), 321-326.

Frödell, M., Josephson, P. E., & Lindahl, G. (2008). Swedish construction clients' views on project success and measuring performance. *Journal of Engineering, Design and Technology*, *6*(1), 21-32.

Grau, D., Back, W. E., & Mejia-Aguilar, G. (2017). Organizational-Behavior Influence on Cost and Schedule Predictability. *Journal of Management in Engineering*, *33*(5), 04017027.

Hall, M. A. (1999). Correlation-based feature selection for machine learning.

Han, J., Pei, J., & Kamber, M. (2011). Data mining: concepts and techniques. Elsevier.

Hartmann, A., Ling, F. Y. Y., & Tan, J. S. (2009). Relative importance of subcontractor selection criteria: evidence from Singapore. Journal of construction engineering and management, 135(9), 826-832

Jha, K. N., & Iyer, K. C. (2006). Critical factors affecting quality performance in construction projects. *Total Quality Management and Business Excellence*, *17*(9), 1155-1170.

Josephson, P. E. (2013). Produktivitetsläget i svenskt byggande 2013, nybyggnad flerbostadshus och kontor. Sveriges Byggindustrier.

Kim, G. H., An, S. H., & Kang, K. I. (2004). Comparison of construction cost estimating models based on regression analysis, neural networksand case-based reasoning. *Building and environment*, *39*(10), 1235-1242.

Khosravi, S., & Afshari, H. (2011). A success measurement model for construction projects. In *International Conference on Financial Management and Economics IPEDR* (Vol. 11, pp. 186-190). IACSIT Press Singapore.

Koch, C., Lundholm, M. (2018). PRODUKTIVITETSLÄGET I SVENSK BYGGANDE 2014. Lokaler, Grupphus och Anläggning

Lam, P. T., & Wong, F. W. (2009). Improving building project performance: how buildability benchmarking can help. *Construction Management and Economics*, *27*(1), 41-52.

Le-Hoai, L., Dai Lee, Y., & Lee, J. Y. (2008). Delay and cost overruns in Vietnam large construction projects: A comparison with other selected countries. *KSCE journal of civil engineering*, *12*(6), 367-377.

Lowe, D. J., Emsley, M. W., & Harding, A. (2006). Predicting construction cost using multiple regression techniques. *Journal of construction engineering and management*, *132*(7), 750-758.

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

Mbugua, L. M., Harris, P., Holt, G. D., & Olomolaiye, P. O. (1999). A framework for determining critical success factors influencing construction business performance. In *Proceedings of the Association of Researchers in Construction Management 15th Annual Conference* (Vol. 1, pp. 255-64).

Mir, F. A., & Pinnington, A. H. (2014). Exploring the value of project management: linking project management performance and project success. International journal of project management, 32(2), 202-217.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis* (Vol. 821). John Wiley & Sons.

Naoum, S. G., & Egbu, C. (2016). Modern selection criteria for procurement methods in construction: A state-of-the-art literature review and a survey. *International Journal of Managing Projects in Business*, *9*(2), 309-336.

Nordstrand, U. (2008). Byggprocessen (Vol. 248). Liber.

Odeh, A. M., & Battaineh, H. T. (2002). Causes of construction delay: traditional contracts. *International journal of project management*, *20*(1), 67-73.

Ogunlana, S. O. (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*(3), 228-236.

Olive, D. J. (2017). Linear regression. Springer.

Sachdeva, J. K. (2009). *Business research methodology*. Himalaya Publishing House.

Sambasivan, M., & Soon, Y. W. (2007). Causes and effects of delays in Malaysian construction industry. *International Journal of project management*, *25*(5), 517-526

Sanvido, V., Grobler, F., Parfitt, K., Guvenis, M., & Coyle, M. (1992). Critical success factors for construction projects. Journal of construction engineering and management, 118(1), 94-111.

Tangen, S. (2005). Demystifying productivity and performance. International Journal of Productivity and performance management, 54(1), 34-46.

Trost, S. M., & Oberlender, G. D. (2003). Predicting accuracy of early cost estimates using factor analysis and multivariate regression. Journal of Construction Engineering and Management, 129(2), 198-204.

Walker, D. H. (1995). An investigation into construction time performance. *Construction Management and Economics*, *13*(3), 263-274.

Witten, I. H., Frank, E., Hall, M. A., & Holmes, G. (2011). Data mining : Practical machine learning tools and techniques with java implementations. San Francisco: Elsevier Science & Technology

Xiong, B., Skitmore, M., Xia, B., Masrom, M. A., Ye, K., & Bridge, A. (2014). Examining the influence of participant performance factors on contractor satisfaction: A structural equation model. International Journal of Project Management, 32(3), 482-491.

Yan, X., & Su, X. (2009). *Linear regression analysis: theory and computing*. World Scientific.