MASTER

To bid, or not to bid

how contractors should support the decision-to-bid during tendering for projects

van Blerk, L.

Award date:
2017

23th of August, 2017

To bid, or not to bid:

How contractors should support the decision-to-bid during tendering for projects

Master Thesis (1ZM96)

by:
L. (Luc) van Blerk

Student identity number: 0751489
l.v.blerk@student.tue.nl

In partial fulfilment of the requirements for the degree of

Master of Science
in Innovation Management

First university supervisor
Dr. Ir. Michel van der Borgh
Assistant Professor of Marketing
Fac. IE&IS - TU/e
W.v.d.borgh@tue.nl

Second university supervisor
Dr. Sarah Gelper
Assistant Professor of Marketing
Fac. IE&IS - TU/e
S.gelper@tue.nl

First company supervisor
Ir. Mart Corbijn van Willenswaard
Executive Manager Engineering
Business Unit Airports
Mart.corbijn.van.willenswaard@vanderlande.com

Second company supervisor
Gilbert Schapendonk
Senior Process Improvement Manager
Planning & Performance
Gilbert.schapendonk@vanderlande.com

TUE. School of Industrial Engineering.

Series Master Theses Innovation Management

Subject headings: tendering, CRM, data-analysis, multiple linear regression, decision-to-bid

LIST OF DEFINITIONS

Term	***Definition***
Contractor	The company which responds to an Invitation To Bid/ Invitation to Tender or Request For Proposal to obtain projects (Philbin, 2008; Vincler & Vincler, 1996).
Project owner	The company which sends out an Invitation to Bid/Invitation to Tender or Request For Proposal to several contractors (Philbin, 2008; Vincler & Vincler, 1996).
Tendering	A transparent procurement method, in which project owners invite contractors and suppliers by openly announcing the scope, specifications, terms & conditions and the evaluation criteria on which the bid will be chosen (Ballesteros-Pérez et al., 2012).
Bid	A bid is a formally submitted document by the contractor in response to the invitation to tender (Jaques, 2011).
Decision-to-bid	Prior to developing a bid, contractors decide whether to bid or not. From that moment, major investments in terms of sales and engineering hours are allocated to the development of the bid (Skitmore et al., 2006)
Baggage Handling System	A Baggage Handling System (BHS) transports, stores and sorts baggage on Airports using a conveyor system.
CRM	Customer Relationship Management (CRM) is a strategic approach that is concerned with creating improved shareholder value through the development of appropriate relationships with key customers and customer segments. CRM unites the potential of relationship marketing strategies and IT to create profitable, long-term relationships with customers and other key stakeholders. CRM provides enhanced opportunities to use data and information to both understand customers and co-create value with them. This requires a cross-functional integration of processes, people, operations, and marketing capabilities that are enabled by information, technology, and applications (Payne & Frow, 2005).

ABSTRACT

Contractors who try to obtain projects via tendering have to decide whether to bid or not. This decision-to-bid is an important point during the tender process, since from that moment on, major investments in terms of sales and engineering hours are allocated to the development of a bid (Skitmore et al., 2006).

However, this decision is influenced by numerous factors, and literature suggests that, in practice, the decision-to-bid is mainly based on a subjective assessment (Odusote & Fellows, 1992; Shash, 1993; Smith, 1995; Wanous et al., 1998). To support the decision-to-bid, quantitative models were rarely reported being valuable.

In addition, scholars state that little empirical work about tendering is reported in literature (Philbin, 2008). Moreover, it is often unknown how initial project characteristics relate to (1) the quotation costs of capturing a project via tendering and to the execution of the project in terms of (2) man-hours and (3) contribution margin.

This thesis uses a quantitative approach, to address how contractors should support the decision-to-bid, by gathering insights via multiple linear regression analyses, based on project data collected from contractors' IT systems. The data analysis provided valid models for all three aforementioned subjects. This report describes the approach and presents the results.

ACKNOWLEDGEMENTS

With this master thesis project, I completed my master Innovation Management, at the Eindhoven University of Technology (TU/e). The project was conducted at Vanderlande Industries in Veghel, from February 2017 until August 2017. The execution of the project would not be possible without the input and help of several people, therefore, I would like to show my gratitude to those who supported me throughout the process.

Firstly, I would like to thank Vanderlande Industries for giving me the opportunity to conduct my master thesis at Vanderlande. In particular, my first company supervisor Mart Corbijn van Willenswaard, and second company supervisor Gilbert Schapendonk. Mart, although your agenda was occupied throughout, you always made time to provide me with feedback or ideas. Moreover, you gave me the freedom to work autonomous within Vanderlande and connected me with several people - no matter their hierarchical importance within Vanderlande-, which I really enjoyed. Someday, I will also be able to summarize a long conversation, ad hoc, into a single sentence.

Gilbert, thanks for welcoming me and providing me a workplace at the World Class Operations (WCO) department. When I wasn't able to meet Mart or had some quick questions, I was always able to ask you. Additionally, I would like to thank Mario Sigmans and Jelte van Gils for providing me with useful information and an occasional walk after lunch. Furthermore, I would like to thank Ietje Penninga, Chrissy Donsbach, and Hendy de Geus for providing me the data.

Secondly, I would like to thank my TU/e supervisors, Michel van der Borgh and Sarah Gelper. Michel, although your ideas overwhelmed me sometimes, your expertise, both theoretically *and* practically, helped me to successively complete this master thesis project. Sarah, thanks for looking with me at the datasets in more detail, your tips regarding the data analysis were of great use.

Prior to conducting this master thesis project at Vanderlande, I couldn't imagine which problems I would encounter. In my opinion, finding, gathering and assessing the data on validity was by far the most difficult, though I believe it was worth the effort. That's why I would like to end this section with a catchy quote from Charles Babbage:

"Errors using inadequate data are much less than those using no data at all"

Luc van Blerk

Eindhoven, August 2017

MANAGEMENT SUMMARY

Introduction

Tendering is associated with a high level of complexity and uncertainty (Philbin, 2008). Especially at the beginning of a tender process, details regarding the project characteristics are lacking, making it difficult for contractors to assess whether to bid or not, i.e., the decision-to-bid. However, this decision-to-bid is an important point in the tender process, as from that moment on, major investments can be assigned to the development of a bid (Skitmore et al., 2006).

To better assess the decision-to-bid, Vanderlande was interested in obtaining insights regarding the (1) quotation costs (i.e., costs belonging to the development of a bid), (2) man-hours (i.e., the engineering hours necessary to finalize a project) and (3) the contribution margin of BHS projects at a very early stage. Hence, the following research question was used:

How should Vanderlande support the decision-to-bid, to enhance the allocation of sales and engineering resources, and to have insights regarding the contribution margin of BHS projects?

Literature study

A literature study was conducted to gain a better understanding of tendering, with special attention to the decision-to-bid. We adopted the definition of Ballesteros-Pérez and colleagues (2012) where tendering was defined as a transparent procurement method, in which project owners invite contractors by openly announcing the scope, specifications, terms & conditions and the evaluation criteria on which the bid will be chosen. Three tender methods have been identified, and pricing was found to be one of the key activities during the development of a bid (Jennings & Holt, 1998; Kadefors, 2005; Mochtar & Arditi, 2001). Overpricing a project may result in losing the tender, whereas underpricing may result in winning the tender, yet, the project might not be completed, or completed in a non-profitable way (Akintoye, 2000; Kwak & Watson, 2005).

Interestingly, tender literature was rather inconclusive regarding the decision-to-bid. It was found that the various factors influence the decision-to-bid (Odusote & Fellows, 1992; Shash, 1993; Wanous et al., 1998). While these studies have identified similar factors, little agreement can be found in their relative importance. As a result, Smith (1995) labeled the search for a definitive list of factors as "*a search for the Holy Grail*", as it is more likely that different contractors consider different factors for each project, and that intuitive and subjective judgments change over time.

Contrary to the factors influencing the decision-to-bid, scholars in the field of tendering are more conclusive regarding the assessment of the decision-to-bid. They argue that the decision-to-bid is mainly

based on a subjective assessment by the contractor (Odusote & Fellows, 1992; Shash, 1993; Smith, 1995; Wanous et al., 1998). Quantitative models were rarely reported being valuable to support the decision-to-bid and the succeeding tender process. Unfortunately, little empirical work, regarding how contractors really assess the decision-to-bid in practice, is reported in literature (Philbin, 2008). As a result, scholars even argue that the (subjective) way of working might not have a justifiable basis (Laryea, 2013; Laryea & Hughes, 2008; Philbin, 2008). Note that a possible reason for the lack of empirical studies might be assigned to the contractors themselves, since they are not eager to participate in such studies, due to the sensitive nature of the topic (Skitmore & Wilcock, 1994).

Besides the aforementioned findings, the literature study also provided a research opportunity regarding how contractors should support the decision-to-bid. That is, the use of contractors' knowledge-based IT systems as a basis for the decision-to-bid (Stader, 1997). In 1997, Stader argued that tendering, especially the decision-to-bid, is about gathering information and the subsequent analysis. A decade later, Metallo and colleagues (2007) stated that particularly CRM data is "*a huge information asset for the articulation of a commercial offer*".

Although promising, no study was found that actually incorporated insights, obtained from CRM and related IT systems, to support the decision-to-bid or the succeeding tender process. A possible explanation for the absence of such studies might be related to the fact that IT technologies were not widely adopted at the time of the surveys, i.e., most were conducted over two decades ago. Fortunately, nowadays, contractors store their bidding and project data in CRM systems (Ngai, 2005; Payne & Frow, 2005). Hence, analyzing project data stored in the CRM and related IT systems was found to be a promising opportunity to obtain valuable insights, to be used to support contractors' decision-to-bid.

Research method

The study is considered exploratory in nature, and given the longitudinal time horizon of the projects, archival research was found to be a suitable method (Saunders et al., 2009). Nonetheless, contrary to self-collected data, Vanderlande's archival data has been collected for a specific purpose, which was different from the subject under investigation (Denscombe, 2007). As a result, the ability to answer the research question was inherently constrained by the availability of the archival data. Consequently, recommendations for the firm cultivated from the absence of certain variables.

Project data of baggage handling systems was collected over a time span of seven years. This time span was chosen in agreement with Vanderlande, to make sure the included projects were relevant for the analysis, while still being comparable for future projects (i.e., not too outdated). After the data gathering, the data was evaluated on validity, and a single dataset was constructed by connecting all columns with each other using project identification numbers. This resulted in a dataset with 1504 project numbers. At first glance, this seemed like a huge dataset, however, many of the project numbers in the master data were not considered relevant for the analysis. As some project numbers only represented 'replace cranes' or 'update software', instead of 'BHS Airport X'. Therefore, to be included in the data analysis, a filter was applied to only include BHS projects which fitted the scope of this study. The filter required that the observations should at least contain a 'motor', a 'length of the belts', an 'electrical connection', and should be present in the CRM system.

Then, multiple linear regression (MLR) was used to make predictive models, where a hybrid stepwise approach was chosen to select the best subset of predictor variables. Importantly, early in the tender process, information regarding the BHS project was found to be scarce. However, over time, gradually more information is identified. Therefore, for the quotation costs and man-hours, we decided to make four MLR models per response variable, such that the resulting models can be applied over time.

Results & Conclusions

By applying MLR analysis we obtained valid models for all the three themes (i.e., quotation costs, man-hours, and contribution margin). The models allow Vanderlande to obtain various insights at an early stage in the tender process. In addition, the models only use a few predictor variables, increasing the practical use.

Quotation costs

Valid MLR models were obtained for the quotation costs, which explained up to 72% of the total sample variability. Hence, Vanderlande can now estimate the quotation costs at a very early stage, prior to the decision-to-bid. In addition, by having an idea of the expected quotation costs, Vanderlande can monitor the sales process more closely, i.e., whether the bid under or over performs in terms of quotation costs.

Further improvements to the quotation cost models are likely to occur once more predictor variables are added to the model. Importantly, various predictor variables have been mentioned and identified during the interviews. Though, we were unable to include all the possible variables. Either because the variables were not stored, not stored properly, or not stored for a decent period. More information regarding the identified predictor variables and recommendations for storing the data can be found in Appendix VII.

Man-hours

Valid MLR models were obtained for the man-hours of Mechanical Engineering, Project Management, Site Management, Project Leader Engineering, and Low Level Controls. As a result, Vanderlande can now estimate the necessary man-hours to successfully conduct a BHS project at an early stage, using only a few variables, and act accordingly. For instance, to check whether the estimated amount of man-hours are available during the project period or whether additional employees need to be hired/trained to conduct the project.

Producing valid MLR man-hours models for other functional roles proved to be more problematic, as we were unable to obtain sufficient observations for Integration Management (7 obs.), High Level Controls (14 obs.) and the Project Directors (0 obs.). This deficiency in observations can be explained due to a recent change in the registration of the man-hours, in addition, some of the roles were only used in a relatively small number of BHS projects.

As for the quotation costs models, further improvements to the man-hour models are likely to take place once more predictor variables are added to the model, and more observations are available to apply advanced statistics.

Contribution margin

Insights regarding the contribution margin of BHS projects were obtained using the 'latest estimate' of the contribution margin. By selecting the 'latest estimate' of the contribution margin, we included BHS projects which were not entirely completed yet. By doing so, we were able to enlarge the dataset considerably. This choice was supported by Vanderlande, as the 'latest estimate' generally gives a good impression regarding the final contribution margin. Subsequently, three subjects, related to the contribution margin, were investigated using MLR analysis.

First, based on the 'sold sales value' of the BHS projects, valid MLR models were found for the contribution margin, for both the 'as sold' and the 'latest estimate' situation. Due to confidentiality, we did not state the exact relation between the 'sold sales value' and the contribution margin (or whether the relation was positive or negative), nonetheless, we can state that the MLR models provided interesting insights. Interestingly, no additional effects for the categorical variables 'project type' and 'customer center' were found after conducting a statistical analysis of covariance. This might be due to the fact that the dataset was unbalanced, i.e., after incorporating the categorical variables the size of some subsets was quite small, resulting in rather big confidence intervals.

Secondly, we investigated whether we could obtain valid MLR models for the contribution margin, based on the 'sales value per quotation hour'. No significant MLR models were obtained, neither when we included the 'project type' nor the 'customer center' in an analysis of covariance.

And thirdly, we investigated whether we could obtain valid MLR models for the 'delta CM' (i.e., the difference between the 'as sold' contribution margin and the 'latest estimate' contribution margin) based on the 'sales value per quotation hour'. No significant MLR models were obtained, neither when we included the 'project type' and 'customer center' in an analysis of covariance. A reason for these insignificant outcomes might be due to the fact that many BHS projects have a relatively long project period (i.e., some over three years), in which various factors might have an influence on the 'latest estimate' CM value.

Recommendations

The study showed that valid MLR models can be obtained from contractors' IT systems to support the decision-to-bid. Nevertheless, improvements are necessary (and possible) to increase the predictive capabilities in the near future. Various predictor variables have been mentioned by employees of Vanderlande, however, we were unable to incorporate all the mentioned variables in the analysis. Either because the data was not stored, not stored for a decent period, or stored inappropriately. Hence, we proposed various recommendations as can be found in Appendix VII and VIII.

CONTENTS

1. INTRODUCTION

1.1. Company background

Vanderlande is a material handling and logistics automation firm, with its headquarters located in Veghel, the Netherlands. The company was founded in 1949, and by that time, Vanderlande produced machines for the textile industry (Vanderlande, 2017b). Over the years, the product portfolio changed considerably, and by the year 2017, Vanderlande has reached the status of an international player in the field of automated baggage handling systems for Airports, warehouse automation solutions, and sorting systems for parcel and postal services, as can be found in Figure 1 (Vanderlande, 2017b).

Figure 1. *General facts and figures of Vanderlande (Vanderlande, 2017a).*

Vanderlande is currently carrying out over 250 projects in 105 countries, as a result, most of their revenue is realized in foreign countries. In the calendar year 2016, Vanderlande realized a revenue of 1.1 billion euros, with earnings before interest and tax of 64 million euros (Vanderlande, 2017a).

Vanderlande states that there is a high level of interest in baggage handling systems (BHS), and reported a positive outlook for the coming years (Vanderlande, 2017a). Moreover, the underlying market conditions seem promising: the air traffic market is expected to be twice as big within 15 to 20 years (Vanderlande, 2017a). Consequently, the global need for baggage handling systems is expected to increase significantly over the coming years.

1.2. Problem statements

Vanderlande obtains most of its baggage handling projects through tendering and is increasingly invited by project owners to participate in tenders all over the world. However, as Vanderlande, such as any other contractor, has limited resources, they are unable to develop well-defined bids for *all available* tenders. Consequently, they are forced to select which projects to pursue, prior to the formal tender process. This is problematic, as it is unknown during the sales phase how the tender process will turn out in terms of quotation costs. Therefore, Vanderlande states it is difficult to assess whether to bid for a project or not, i.e. the decision-to-bid. Based on this, the following problem statement has been stated:

1. *Prior to developing the bid, knowledge regarding the expected quotation costs to obtain a BHS project via tendering is lower than desired (e.g., more knowledge regarding the expected quotation costs is needed to better assess the decision-to-bid)*

Once a bid is won, Vanderlande uses a detailed layout of the project to calculate the man-hours using a bottom-up approach. Unfortunately, this approach cannot be used prior to the decision-to-bid, as a detailed layout is absent. However, for Vanderlande to be able to act in advance – prior to developing the bid-, it is important to have an idea of the future workload (e.g. to hire or train new employees). Based on this information, the following problem statement has been stated:

2. *Prior to developing the bid, knowledge regarding the man-hours to execute a BHS project is lower than desired (e.g., more knowledge regarding the man-hours is needed to better assess the decision-to-bid).*

Prior to developing a bid, the actual contribution margin of a BHS project is unknown. Insights regarding the contribution margin of a certain BHS project would be of great advantage for Vanderlande to better assess the decision-to-bid, for instance, to bid on tenders with a high expected contribution margin. Based on this, the following problem statement has been stated:

3. *Prior to developing the bid, knowledge regarding the contribution margin is lower than desired (e.g., more knowledge regarding the contribution margin is needed to better assess the decision-to-bid.)*

1.3. Main research question

In tendering, contractors compete on the right for a business opportunity (Green, 1989; Laryea, 2013). Within the tender process a formal proposal – a bid - is made by the contracting firm, which incorporates the requirements of the project owner to a solution and accompanying price (Hackett, Robinson, & Statham, 2007). The tender process is often associated with a high level of complexity and uncertainty (Philbin, 2008). However, the (in)ability to manage the tender process can have a direct and major impact on the success of the firm (Philbin, 2008).

Importantly, contractors who want to capture projects via tendering come across two decisions.

First, contractors have to decide whether to bid or not – the decision-to-bid. This decision is influenced by numerous factors, and literature suggests that, in practice, the decision-to-bid is mainly based on a subjective assessment (Odusote & Fellows, 1992; Shash, 1993; Smith, 1995; Wanous et al., 1998). Quantitative models were rarely reported being valuable during the decision-to-bid and the succeeding tender process.

And second, once a contractor decides to bid, a final bid price has to be submitted to the project owner. Formation of the final bid price is divided into two separate stages: first, the variable costs and contribution margin are estimated, and second, *adjudication* is applied by the management to the estimated project price to create the final bid price (Brook, 2008).

Interestingly, both the decision-to-bid, as well as the pricing decision seem to be highly related to each other. As overpricing a project may result in losing the tender, whereas underpricing may result in winning the tender, yet in an unprofitable way (Akintoye, 2000; Kwak & Watson, 2005).

Unfortunately, little empirical work about tendering is reported in literature (Philbin, 2008). It is often unknown how initial project characteristics relate to the costs of capturing a project via tendering and to the execution of the project in terms of man-hours or contribution margin. As a result, scholars even argue that the current subjective way of working might not have a justifiable basis (Laryea, 2013; Laryea & Hughes, 2008; Philbin, 2008). Therefore, based on the problem statements of Vanderlande and the status quo in the tender literature, the following research question was formed:

How should Vanderlande support the decision-to-bid, to enhance the allocation of sales and engineering resources, and to have insights regarding the contribution margin of BHS projects?

1.4. Research scope

On an abstract level, Vanderlande has an input of BHS leads, which are processed, and the outcome is a finished BHS project, as can be found in Figure 2. This research will mainly focus on the input side, by investigating how to support the decision-to-bid, i.e., to have more knowledge to assess the decision-to-bid to only pursue the 'right' projects.

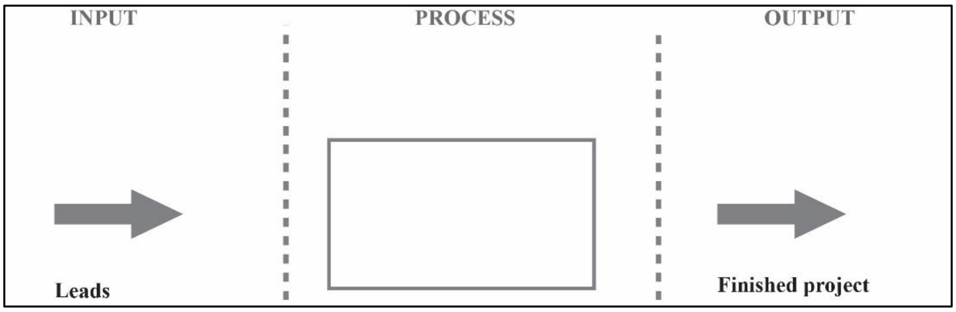


Figure 2. *Abstract input-process-output model of Vanderlande's BHS projects.*

Contrary to focusing on the input side, one could also consider to improve the process part, to achieve better project outcomes. However, we argue that Vanderlande should first make sure to only pursue the 'right' projects, before further optimizing the process part. By focusing on the input part, Vanderlande can omit the 'garbage in, garbage out' phenomenon. Consequently, Vanderlande does not have to struggle of processing 'bad' projects into successful outcomes.

Importantly, in agreement with Vanderlande, only BHS projects will be included, starting from the lead phase until the completion of the project and taken over by services. Service projects will be left out of scope, as these are usually separated from the actual construction of the project, an overview of the research scope within Vanderlande can be found in Figure 3.

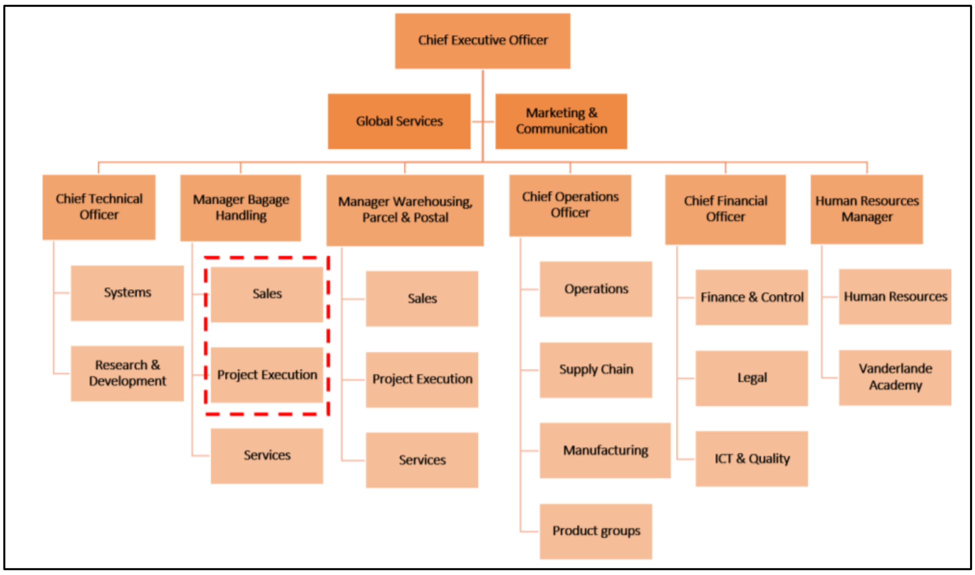


Figure 3. *Overview of the research scope (the red block) within Vanderlande.*

1.5. Research method

The aim of this research was to analyze BHS projects that were captured via tendering and completed at Vanderlande, to gather insights to be used to support the decision-to-bid during tendering for projects.

Little empirical research has been reported in literature (Philbin, 2008), therefore, applying a causal research using pre-developed theories was found to be unsuitable. As a result, given the nature of the research question and the extent of the existing knowledge on tendering, exploratory research seemed the most appropriate.

1.5.1. Process mining

Importantly, as the decision-to-bid is part of a tender process, we investigated whether it would be possible to apply process mining. Process mining uses system data to discover, check conformance, and improve real processes(Van der Aalst, 2016; Van der Aalst & Weijters., 2012). And as process mining has been adopted in a wide variety of organizations (e.g., municipalities, governmental agencies, insurance companies and banks), process mining seemed a promising method to analyze the BHS project data. Unfortunately, an initial investigation regarding the necessary event data revealed that the IT systems did not store all the relevant process events, and timestamps where often questionable. As a result, we were unable to make use of process mining. Nonetheless, we still argue that Vanderlande could gain a great amount of knowledge when they would apply process mining. For that reason, in Appendix VIII, we paid special attention regarding the storage of data for process mining.

1.5.2. Multiple Linear Regression analysis

Although process mining was found to be unfeasible for this research, the IT systems did store a great variety of information which could be explored using data mining tools. We selected Multiple Linear Regression (MLR) to be used for the data analysis, as this is one of the most popular data analysis methods, and has widely proven its use in decision making (Hair, Black, Babin, & Anderson, 2014). Moreover, contrary to sophisticated data analysis methods (e.g., Artificial Neural Networks), MLR was found to be a suitable method to discuss findings with employees from Vanderlande.

1.6. Contributions

Literature argues that contractors tend to rely on subjective assessments based on past experience and intuition in their decision-to-bid, rather than on quantitative data (Hwang & Kim, 2016; Smith, 1995). Though, Stader (1997) proposed the use of knowledge-based IT systems as a basis for the decision-to-bid. Additionally, Metallo and colleagues (2007) stated that particularly CRM data is "*a huge information asset for the articulation of a commercial offer*". Although promising, no study was found that actually incorporated insights, obtained from CRM and related IT systems, to support the decision-to-bid or the succeeding tender process. Hence, this thesis is the first research that empirically shows that contractors' IT systems can provide predictive models and insights, to be used during tendering for projects.

In addition, little empirical work about tendering is reported. As a result, scholars such as Laryea (2013) and Philbin (2008) argue that especially empirical work is necessary to bridge the gap between theory and practice. Given the real-world project data which is used in the analyses, this study fulfills in that desire. Furthermore, this thesis provides contractors important recommendations with respect to the storage of project data.

1.7. Outline of the report

In the first chapter of this thesis an introduction of the company is presented, and the problem statements are stated. The next chapter contains the theoretical background, in which special attention has been allocated to the decision-to-bid. Additionally, definitions will be provided and relevant topics will be explained in more detail. The theoretical background ends with a conclusion including the research opportunity regarding how contractors should support the decision-to-bid. The third chapter provides a detailed description regarding the research method, data collection, and data analysis. Chapter four presents the data analysis for the quotation costs, man-hours and contribution margin. Finally, the fifth chapter provides conclusions, both managerial and theoretical. Additionally, limitations of the study are presented and ideas for future research are proposed.

2. THEORETICAL BACKGROUND

This chapter summarizes the current tender literature. The chapter starts by providing a definition of tendering and its associated terms. In the second part, findings related to pricing tender projects are described. Lastly, the status quo of using CRM systems during tendering will be discussed.

2.1. Tendering

Tendering can be seen as the business practice in which contractors compete on the right for a business opportunity, in a changing and dynamic environment (Green, 1989; Laryea, 2013). Ballesteros-Pérez and colleagues (2012) defined tendering as a transparent procurement method, in which project owners invite contractors by openly announcing the scope, specifications, terms & conditions and the evaluation criteria on which the bid will be chosen. Briefly stated, the purpose of tendering, for a project owner, is twofold: first, to select a suitable contractor within a certain time, and secondly, to obtain a price for the project. Together this forms the basis for a contract between a project owner and a contractor (Murdoch & Hughes, 2008). Three tender methods have been identified in the literature, an overview of the three methods can be found in Appendix I.

Within tendering, a formal proposal – a bid – is made by the contracting firm, which incorporates the requirements of the project owner to a solution and accompanying price (Hackett et al., 2007). Tendering is often associated with construction projects, however, literature suggests that the process is not different in other industries (Laryea, 2013).

The tender process can be highly complex with many uncertainties, nevertheless, the ability to properly manage the tender process can have a direct and major impact on firms' success (Philbin, 2008). Unfortunately, contractor' bidding success rates are low, often only 15-30% of the submitted bids are awarded (Gann & Salter, 2000; Hicks et al., 2000; Konijnendijk, 1994; Laryea & Hughes, 2008). Importantly, if a project is not secured, spend resources during the tender process – quotation costs – turn into sunk costs, as the investment cannot be repaid by the lost project (Philbin, 2008; Smartt & Ferreira, 2015). For that reason, authors like Nickson (2003), Whitley (2006) and Philbin (2008) have proposed the discipline of bid management.

Bid management is concerned with aligning different functional aspects of the tender process, to efficiently turn sales resources into successful bids (Nickson, 2003; Philbin, 2008; Whitley, 2006). Bid management can be compared to the management of a project, yet the outcome is not a physical project but a bid for business (Whitley, 2006).

Interestingly, project owners are increasingly applying tendering in practice, further increasing the need for proper bid management (Eriksson & Westerberg, 2011; Hardeman, 2012; Paul & Gutierrez, 2005; Runeson & Skitmore, 1999). A first reason for the increasing usage of tendering might be found in the way tendering can be used to strategically seize bargains from contractors (Eriksson & Westerberg, 2011; Paul & Gutierrez, 2005; Runeson & Skitmore, 1999). Though, scholars have also found contradictory evidence in which contractors offer unrealistic low bids to win the tender, which might turn out not to be bargains after all (Hensher & Stanley, 2008; Jha & Iyer, 2006; Rooke et al., 2004; Smith & Bohn, 1999). A second reason for the increase in tendering is the case of governmental projects, as contractor selection by means of tendering is more often required by the nation's law (Eriksson & Westerberg, 2011; Paul & Gutierrez, 2005).

While the effective capturing of projects via tendering and the related discipline of bid management is a promising area of management, it has received little attention in literature (Philbin, 2008). According to Laryea (2013), especially empirical work is necessary to bridge the gap between theory and practice. Ironically, a reason for the lack of tender literature can be found at the contractor companies themselves, since contractors are not willing to participate due to the sensitive nature of the data (Skitmore & Wilcock, 1994).

2.2. Tender procedures

Tendering starts once a project owner sends out a Request For Information (RFI) document, in which potential contractors respond to questions from the project owner, either to further specify the project requirements or refine the award/evaluation mechanism (Seshadri, 2005). Consequently, the project owner pre-qualifies contractors to participate in the tender, mainly selected upon the projects the contractors accomplished in the past (Smith, 1995, pp. 123-126). An extensive explanation of all the tender phases can be found in Appendix II.

The actual formal tender process starts once the project owner sends information to several pre-qualified contractors (Eriksson & Westerberg, 2011; Philbin, 2008; Seshadri, 2005; Vincler & Vincler, 1996). Typically, the invitation document includes site plans, details of existing buildings and information of any unusual features or conditions. Although this information is not directly price-able, it gives the contractor context regarding the scope and associated risk of the project (Smith, 1995). In addition, often the project owner, together with consultants, perform detailed design work in order to create a more solid basis for tendering (Eriksson & Westerberg, 2011). However, in practice the amount of information can vary considerably, some project owners only provide key features, whereas others provide extensive information about design requirements, constraints, contractual conditions and performance (Seshadri, 2005; Smith, 1995). Although documentation might not include all details, it does give contractors an equal base, on which they can base their decision-to-bid (Jaques, 2011)

2.3. Decision-to-bid

The decision-to-bid can be seen as a turning point, since from that moment on, major investments in terms of sales and engineering hours are allocated to the development of the bid (Skitmore et al., 2006). Factors influencing contractors' decision-to-bid have been identified and prioritized on their relative importance by various scholars (Odusote & Fellows, 1992; Shash, 1993; Wanous et al., 1998). Interestingly, although several studies have identified similar factors, little agreement can be found in their relative importance. For instance, Shash (1993) identified and ranked factors affecting the decision-to-bid, with 'the need for work', 'number of competitors' and 'experience in similar projects' as the most important factors, whereas Wanous and colleagues (1998) reported 'fulfilling the tender conditions', 'financial capability of the client', and 'relation with/reputation of the client' as the most important factors affecting the decision to tender. A more comprehensive overview can be found in Appendix III.

In line with these inconclusive findings regarding factors influencing the decision-to-bid, Smith (1995) labeled the search for a definitive list as "*a search for the Holy Grail*", as it is more likely that different contractors consider different factors for each project, and that intuitive and subjective judgments change over time. Although finding a definitive list of factors affecting the decision-to-bid might be impossible, literature does argue that contractors tend to rely on subjective assessments based on past experience and intuition, rather than on quantitative data (Hwang & Kim, 2016).

2.4. Development of the bid

Once a contractor decides to pursue a tender (i.e., a 'go' has been given in the decision-to-bid), they enter the development phase, where the contractor will start creating the bid. During the formulation of a bid, contractors need to make many assumptions, as tender documents are often incomplete (Doloi, 2011; Seshadri, 2005). This is worrisome, as incorrect assumptions can turn to serious liabilities after winning the tender (Laryea & Hughes, 2008). To overcome some of this uncertainty, a site visit is often arranged to gain a more accurate impression of the project site before further bid development (Smith, 1995, pp. 22). Reasons for visiting the site may include general site details such as: site access, security, the condition and present use of buildings, or the presence of local labor and subcontractors.

One of the key activities during the bid development concerns the price setting, which is realized via two separate stages. Firstly, contractors estimate the costs for conducting the project, apply a percentage for overhead and include a mark-up: together this forms the sales estimate (Akintoye, 2000; Carr, 1989; Cassaigne, Kromker, Singh, & Wurst, 1997). Second, the sales estimate is converted into a sales price through *adjudication*, in which management takes the prevailing market conditions into account, and modifies the price accordingly (Smith, 1995).

2.5. Evaluation

Once the bid is finalized, the submitted bids are evaluated by the project owner. The criteria on which the bids are evaluated, and their relative weights, are known upfront, and usually compromises of: price, technical requirements, management capabilities, earlier experience, environmental and quality management systems, financial stability, collaborative skills, schedule and delivery aspects (Eriksson & Laan, 2007; Lam, Hu, Ng, Skitmore, & Cheung, 2001; Philbin, 2008). Of these criteria, the price seems to be a key parameter receiving the most weight, as project owners typically praise the lowest price above other parameters (Jennings & Holt, 1998; Kadefors, 2005; Mochtar & Arditi, 2001).

Once a bid is chosen, the tender is 'closed' and it reaches the status of a project. In practice, changes to the project can still occur, however, the contractor has been selected and will not change anymore.

Recapitulating, tendering is a procurement method for project owners to select a capable contractor for a suitable price (Ballesteros-Pérez et al., 2012). Within the tender process, the decision-to-bid can be seen as a turning point since from that moment on, major investments in terms of sales and engineering hours are allocated to the development of the bid (Skitmore et al., 2006). Interestingly, literature argues that contractors tend to rely on subjective assessments based on past experience and intuition during the decision-to-bid, rather than on quantitative data (Hwang & Kim, 2016).

2.6. Pricing tender projects

Within tendering, pricing the project is one of the key activities during the development of the bid. Not surprisingly, as the tender price often receives most of the weight during the evaluation of the bid, and consequently, is related to winning the tender (Jennings & Holt, 1998; Kadefors, 2005; Mochtar & Arditi, 2001).

Appropriate pricing of tender projects is of strategic importance (Akintoye, 2000). Overpricing a project may result is losing the tender, whereas underpricing may result in winning the tender, however, the project might not be completed, or completed in a non-profitable (Akintoye, 2000; Kwak & Watson, 2005). Extreme underpricing might even give the impression the contractor does not understand the requirements of a project owner, and as a result, loses the bid (Kwak & Watson, 2005). Hence, setting the right sales price is difficult, as this involved a trade-off between contractors making a reasonable profit, yet at the same time, making a chance of winning the tender (Cassaigne et al., 1997).

Traditionally, contractors apply a cost-based pricing strategy, in which they create a project price by calculating the total variable cost of the project (i.e. the labor cost, material cost and equipment cost), apply the cost of overhead and include a mark-up (i.e. risk/contingency and profit) (Akintoye, 2000; Carr, 1989; Cassaigne et al., 1997; Mochtar & Arditi, 2001). An overview of the various pricing elements which lead to the estimated project price can be found in Figure 4.

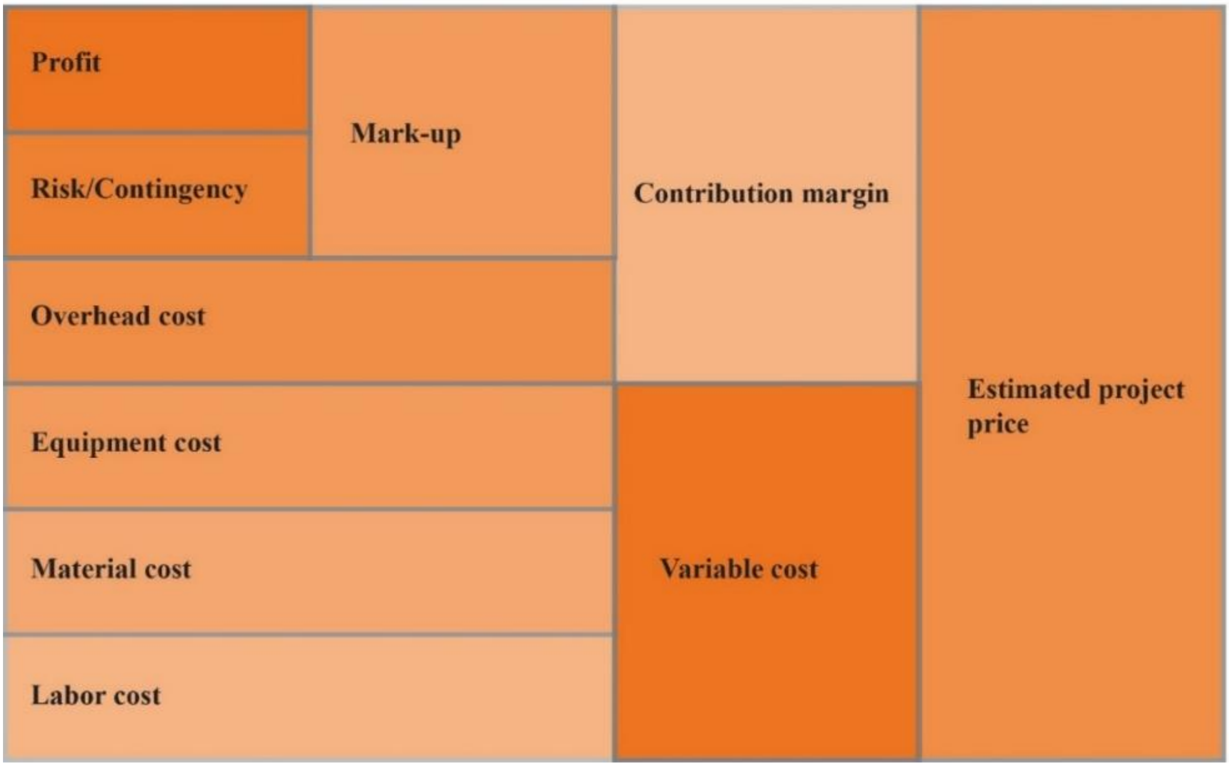


Figure 4. *Cost-based pricing overview, with on the left side the most basic pricing elements, while on the right side more generic terms are shown. Mark-up, for instance, is composed out of profit and risk. While the variable cost is the sum of the equipment, material and labor costs.*

Cost-based pricing is an inward oriented pricing strategy, ignoring consumer and competitive information (Noble & Gruca, 1999). However, in tender situations, contractors do take external information into consideration, by means of an *adjudication* (Mochtar & Arditi, 2001; Smith, 1995). Adjudication concerns the delicate and subjective activity of transforming the estimated project price into a final bid price, which corresponds to the prevailing market conditions and the firms' particular conditions (Laryea & Hughes, 2008; Smith, 1995). Consequently, the estimated project price can either be decreased, increased or kept unchanged. Both the estimation of the project price and the adjudication process are worked out in more detail in Appendix IV and Appendix V accordingly.

Briefly stated, the formation of the final bid price is divided into two separate stages: first, the variable costs and contribution margin are estimated, and second, management applies adjudication to the estimated project price to create the final bid price (Brook, 2008).

2.7. CRM and tendering

Interestingly, in the tender literature, quantitative models were rarely reported being valuable during the tender process, the decision-to-bid, or the price setting. A possible explanation for this low adoption might be found within the knowledge systems of the contractors at the time of the surveys; most studies were conducted over two decades ago, in a time where IT technology was not widely used. Nowadays, however, contractors store their leads, bidding and project data in their CRM systems (Ngai, 2005; Payne & Frow, 2005). Subsequently, analyzing contractor' CRM data might be useful to obtain valuable insights, to be used throughout the tender process. Therefore, this chapter summarizes the current status quo of using CRM insights during the tender process.

In the 1990s, companies began to adopt customer relationship management (CRM) technologies, in order to survive in the saturated competitive landscape (Heinrich, 2005; Keramati et al., 2010; Soltani & Navimipour, 2016). Within a decade, company interest in CRM technologies grew tremendously, making CRM technology a multi-billion dollar industry (Ngai, 2005; Payne & Frow, 2005).

Providing a definition for CRM is not straightforward: Various scholars argue that providing a definition of the CRM concept is problematic (Boulding, Staelin, Ehret, & Johnston, 2005; Garrido-Moreno & Padilla-Melendez, 2011; Richards & Jones, 2008; Winer, 2001). This is illustrated in a study by Payne and Frow (2005), who found a great variety in CRM definitions. The definitions ranged from a narrow view of implementing a technology to a broad approach of managing customer relationships. Even though there might still be a lack of consensus in defining CRM, for the purpose of this Thesis, we will adopt the extensive definition of Payne and Frow (2005):

> "*CRM is a strategic approach that is concerned with creating improved shareholder value through the development of appropriate relationships with key customers and customer segments. CRM unites the potential of relationship marketing strategies and IT to create profitable, long-term relationships with customers and other key stakeholders. CRM provides enhanced opportunities to use data and information to both understand customers and co-create value with them. This requires a cross-functional integration of processes, people, operations, and marketing capabilities that are enabled by information, technology, and applications*."

Scholars have linked the emergence of CRM technologies to the philosophy of relationship marketing, in which the old sales model of "design-build-sell" (the product-oriented view) is replaced by the "sell-build-redesign" (the customer-oriented view) (Keramati et al., 2010; Rygielski, Wang, & Yen, 2002; Xu, Yen, Lin, & Chou, 2002). Furthermore, the terms relationship marketing and CRM are often used interchangeably in literature (Parvatiyar & Sheth, 2001; Payne & Frow, 2005).

Numerous studies have investigated the relation of CRM implementation on firm performance and reported positive effects (Cao & Gruca, 2005; Hogan, Lemon, & Rust, 2002; Mithas, Krishnan, & Fornell, 2005; Palmatier et al., 2006; Payne & Frow, 2005). Not surprisingly, over the years, companies have adopted CRM technologies on a large scale (Krasnikov et al., 2009).

2.8. Insights from CRM data

One of the key functions of CRM systems is to provide a better understanding of customer data and provide guidelines for strategy (Khodakarami & Chan, 2014; Soltani & Navimipour, 2016; Tanner et al., 2005). Consequently, data mining tools can be used to gather knowledge about customers and allocate (human) resources accordingly (Ngai et al., 2009). Furthermore, data mining tools can extract hidden patterns or information from CRM databases, which can lead to new business opportunities (Ngai et al., 2009)

Importantly, CRM systems are able to capture a large part of the available information within a company, however, not the whole spectrum. Therefore, companies who want to tailor their offerings should integrate, or at least link, their CRM system with other systems such as Enterprise Resource Planning (ERP) or pricing tools (Kopczak & Johnson, 2003). By doing so, they can increase their revenue, effectively allocate their resources and maximize the associated profits (Kopczak & Johnson, 2003).

2.9. CRM supporting the tender process

One of the first scholars to have linked tendering with knowledge-based IT systems – not CRM systems in particular – is Stader (1997). Stader (1997) argued that tendering, especially the decision-to-bid, is mainly about gathering information and the subsequent analysis. Hence, he proposed to use of knowledge-based IT systems as a basis for decision making during the tender process.

A decade later, Metallo and colleagues (2007) related the development of a bid to the complete enterprise knowledge system, with CRM systems in particular. They stated that the stored information is *"a huge information asset for the articulation of an intelligent commercial offer"* and can, for instance, be used in the price setting.

According to Cao and Gruca (2005), CRM begins with acquiring the "right" customer, and a company should really make an effort to discover who the best customers are, and in specific who are not. In tendering this is not different, as contractors conduct the decision-to-bid in an attempt to only pursue the 'right' projects. Where the 'right' projects can be, for instance, the projects who possess a reasonable profit. However, at the moment, this decision-to-bid is mainly based on incomplete documents and subjectively assessed (Smith, 1995).

Fortunately, as mentioned previously, CRM systems have been adopted all over the world by now. These CRM systems may record a great variety of information, such as: project leads, the conducted projects, the activities, the time of entering, and by whom the data was inserted (Jans et al., 2011). Consequently, the stored data can be explored using data mining tools (Ngai, 2009), and as CRM is related to the *'intelligent use of data and technology'* (Boulding et al., 2005), insights might be valuable during the decision-to-bid.

Interestingly, no study has been found that proposed to apply data mining tools to CRM data, with the goal to support decision making during tendering or the succeeding project execution. A possible reason for the lack of empirical studies might be assigned to the contractors, since they are not eager to participate in such studies, due to the sensitive nature of the data (Skitmore & Wilcock, 1994).

Summarizing, in the ever demanding competitive landscape of tendering, it is of strategic importance to fully utilize available information. By using data mining tools on CRM data, insights might be obtained that can support both the decision-to-bid and the succeeding tender process. Although this idea seems promising, not a single study was found in which data mining tools were used on CRM data, and related IT systems, to support the decision-to-bid and the proceeding project process.

2.10. Conclusions and research gap

Prior to the development of a bid, contractors face the decision whether to bid, or not. This decision-to-bid can be seen as a turning point, since from that moment on, major investments in terms of sales and engineering hours are allocated to the development of the bid (Skitmore et al., 2006). Literature argues that contractors tend to rely on subjective assessments based on past experience and intuition in their decision-to-bid, rather than on quantitative data (Hwang & Kim, 2016). Unfortunately, little empirical work, regarding how contractors really assess the decision-to-bid in practice, is reported in literature (Philbin, 2008). As a result, scholars even argue that the way of working might not have a justifiable basis (Laryea, 2013; Laryea & Hughes, 2008; Philbin, 2008).

Stader (1997) proposed the use of knowledge-based IT systems as a basis for the decision-to-bid. A decade later, Metallo and colleagues (2007) stated that the CRM data is "*a huge information asset for the articulation of a commercial offer*". Although promising, no study was found that actually incorporated insights, obtained from CRM and related IT systems, to support the decision-to-bid or the succeeding tender process. A possible explanation for the absence of such studies might be related to the fact that IT technologies were not widely adopted at the time of the surveys; most were conducted over two decades ago. Fortunately, nowadays, contractors store their bidding and project data in CRM systems (Ngai, 2005; Payne & Frow, 2005). Analyzing project data stored in the CRM and related IT systems might be a promising opportunity to obtain valuable insights, to be used during the decision-to-bid and the succeeding tender process. An illustration of this idea can be found in Figure 5.

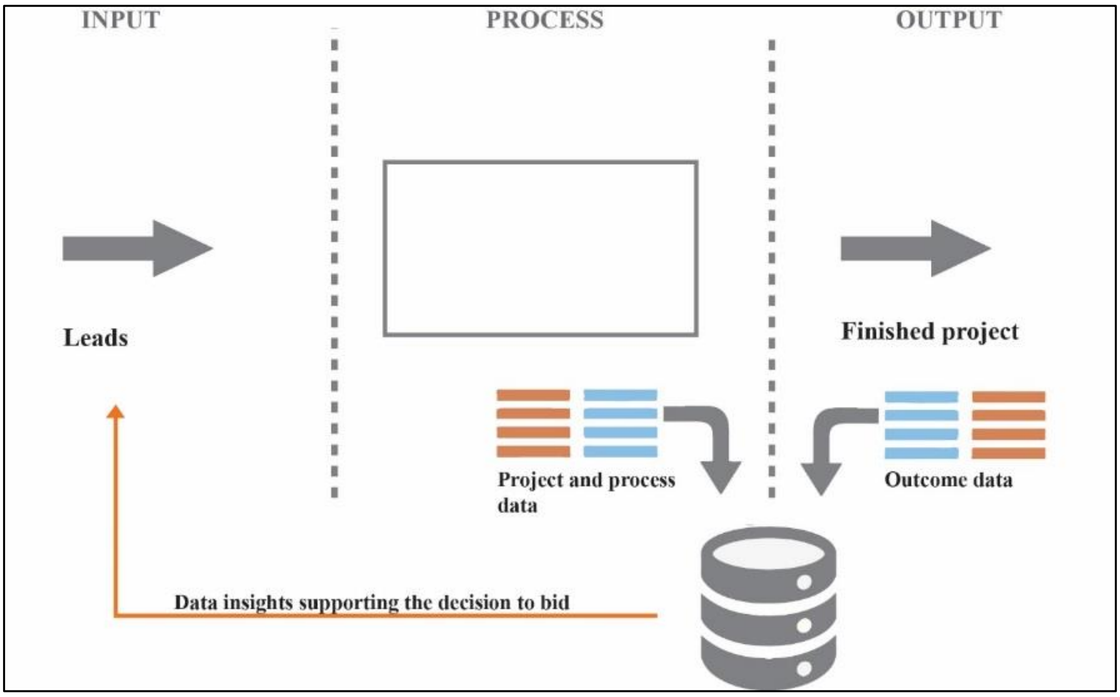


Figure 5. *Applying data mining tools to obtain insights from project data stored in CRM related IT systems, with the intention to support the decision-to-bid.*

3. METHODOLOGY

This study is considered exploratory, making use of archival research within Vanderlande. Moreover, it uses multiple linear regression (MLR) analysis, with the goal of making predictive models and to gather insights, to be used to support the decision-to-bid.

Little empirical work about tendering is reported in literature (Laryea, 2013; Philbin, 2008). Not surprisingly, Laryea (2013) argues that especially empirical work is necessary to bridge the gap between theory and practice. As a result, it is nearly impossible to apply a causal research, due to the absence of pre-developed theory. Therefore, given the nature of the research questions and the extent of existing knowledge on tendering, exploratory research seems the most appropriate. By doing so, one is able to investigate whether and how company's quantitative CRM and related system data might be useful to support the decision-to-bid.

Given the nature of the research context, mainly quantitative methods will be applied rather than qualitative methods. Though, initial interviews will be conducted to identify which insights (response variables) might be valuable to support the decision-to-bid. Subsequently, predictor variables will be identified which might influence the response variables. Based on this information, quantitative methods will be applied to gather insights using the available project data. We expect that not all mentioned predictor variables are stored by the firm (or stored inappropriately), hence, many recommendations regarding the data storage are expected to cultivate from the various interviews.

Interestingly, prior to conducting the research, we investigated whether it would be possible to apply process mining to gather insights regarding the tender process. Unfortunately, a preliminary search regarding the necessary event data revealed that the IT systems did not store all the relevant process events, and timestamps were often open to discussion. As a result, we were unable to make use of process mining for the topic under investigation. Nonetheless, we still argue that Vanderlande could gain a great amount of knowledge when they would apply process mining to the tender process. For that reason, in Appendix VIII, we provided recommendations regarding the storage of data for process mining.

Although process mining was found to be unfeasible for this thesis, the IT systems did store a great amount of data which could be explored using data analysis tools. Therefore, the following sections will explain more about the data collection, the data analysis and the variable selection which was applied in this research.

3.1. Data collection methods

According to Saunders and colleagues (2009), there are roughly seven research strategies: experiment, survey, case study, action research, grounded theory, ethnography, and archival research. Each research strategy can be used for exploratory, causal, or descriptive research designs (Yin, 2003). Furthermore, the strategies are not mutually exclusive, e.g., a survey can be part of a case study (Saunders, Lewis, & Thornhill, 2009). Taken into account the research questions and company context of this research, the most appropriate strategy is the archival research. Archival research makes use of administrative records and documents, consequently, they are part of the reality being studied, rather than being collected for research purposes (Saunders et al., 2009).

The ability to answer the research question is inherently constrained by the availability of the archival data. Even when data is available, archival data may not contain the exact information needed to conduct the research (Saunders et al., 2009). Therefore, it is important to point out some of the advantages and disadvantages of using this research strategy.

3.1.1. Advantages of archival data

One of the main advantages to use archival data, over gathering data yourself, are the savings in terms of time and money (Ghauri & Grønhaug, 2005). By using archival data, one is able to obtain a far larger data set, contrary to self-collected data. Furthermore, given the longitudinal time horizon of the subject under study, archival data is usually the only possibility to undertake such studies (Saunders et al., 2009).

3.1.2. Disadvantages of archival data

Contrary to self-collected data, Vanderlande’s archival data has been collected for a specific purpose, which is different from the subject under investigation (Denscombe, 2007). Consequently, obtained archival data might be unsuitable for the research (Saunders et al., 2009). To overcome this issue, one can collect data from multiple sources, or address the research question only partially.

Importantly, when one intends to combine data sets, project identifiers may differ, or have been changed over time (Saunders et al., 2009). Alternatively, the data may represent the interpretations of those who produced them, rather than an objective image. Therefore, the Vanderlande data must be evaluated carefully before conclusions are drawn, which will be discussed next.

3.1.3. Evaluating archival data

According to Stewart and Kamins (1993), evaluating archival data is time well spent, as rejecting data early in the research can save much time later on. Hence, the *validity* and *reliability* of the archival data are important.

Validity is the extent to which the data measures what you intend to measure (Bordens & Abbott, 2011). For instance, in this study, the overall sales value of a project is an important variable, however, this information is stored in at least three systems within Vanderlande (CRM system, ERP system, and a pricing system). Unfortunately, not all systems are continuously updated with the latest changes, and as a result, some system data might not be valid. This example highlights the importance of evaluating the data sets, as conclusions might be drawn from the wrong information. To overcome the issues of validity, obtained data will be checked with employees from Vanderlande, and where possible, cross-checked with other data sets.

In addition to validity, the *reliability* of a measure concerns the ability to produce similar results when repeated measurements are made under identical conditions (Bordens & Abbott, 2011). The data sets under investigation are obtained in-house and stored by Vanderlande. Therefore, the reliability of the data is high, as one can obtain the same data, and consequently, produce similar results.

3.2. Data collection

The study will use project data from baggage handling systems (BHS). And, as large BHS projects can have a lead-time of over five years, a time span of seven years has been chosen to make sure a relatively large number of completed projects is included in the dataset (January 2010 until June 2017). This time span has been selected in agreement with Vanderlande, prior to the data gathering, to make sure the included projects are relevant for the analysis (i.e., not too outdated).

Based on insights from interviews with employees from Vanderlande, project data is gathered throughout the project's lifecycle, with the aim of gathering BHS specifications, quotation costs, pricing, man-hours, and financial outcome data per project. Obviously, baggage handling projects within Vanderlande go through multiple phases, starting in the lead phase, and ending with a takeover by the project owner. During these phases, project data is stored in various systems, therefore, the data needs to be collected from multiple systems. An overview of the relevant IT systems within Vanderlande can be found in Appendix VI.

Once the raw system data is collected, and evaluated on validity, the data sets will be linked with each other using project identification numbers, to produce a single data set. From that moment on, the data will be analyzed. The data analysis method will be explained next.

3.3. Multiple Linear Regression Analysis

The predictive models will be made using multiple linear regression (MLR). MLR is one of the most popular data analysis methods and widely used in business decision making, e.g., business forecasting models (Hair et al., 2014). The basic formula of MLR is:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon$$

Where Y represents the *response* variable, while $x_1, x_2 \dots x_n$ are labelled the *predictor* variables. The β_0 determines the intercept, while $\beta_1, \beta_2 \dots \beta_n$ are called regression coefficients and describe the slope (Sheather, 2009). The final term ε is termed the residual, or the prediction error. MLR applies the least squares procedure, in which the values of β_0, $\beta_1 \dots \beta_n$ are estimated, such that the sum of squared residuals of prediction is minimized (Hair et al., 2014; Sheather, 2009). MLR is appropriate if there is a linear relationship between the predictors and the response variables, however, non-linear relationships can be included by transforming the data such that it behaves more or less like a linear relationship (Teetor, 2011)

3.4. Data analysis

To answer the main research question, an attempt will be made to produce valid predictive models to support the decision-to-bid, during the sales phase of a BHS project.

3.4.1. Response variables

Based on initial interviews with sales and operational departments of Vanderlande, three response variables have been selected, whose values are held to be important during the decision-to-bid. The three response variables are:

1. **Quotation costs in Euro**, i.e., costs belonging to the development of a bid. Quotation costs may include all costs between the lead-phase and the signed contract, e.g., costs of the sales hours, operational or legal departments.

   Explanation: Major investments in terms of quotation costs are allocated to the development of the bid (Skitmore et al., 2006). Knowledge regarding the expected quotation costs for the development of a bid is believed to be valuable, for instance, to properly allocate the limited sales resources over the available tenders (Vanderlande, 2017). In addition, once the estimated quotation costs are known, the sales department can monitor the sales process.

2. **Man-hours**, i.e., the number of hours necessary to finalize a BHS project. Within Vanderlande, the man-hours are categorized per role:
   - *Mechanical Engineering hours* *(ME hours)*
   - *Project Management hours* *(PM hours)*
   - *Project Leader Engineering hours* *(PLE hours)*
   - *Integration Management hours* *(IM hours)*
   - *Project Director hours* *(PD hours)*
   - *Site Management hours* *(SM hours)*
   - *Low Level Control hours* *(LLC hours)*
   - *High Level Control hours* *(HLC hours)*

   Explanation: For Vanderlande to be able to act in advance – prior to developing the bid –, it is important to have an idea of the future workload. Knowledge regarding the expected man-hours can, for instance, be used to assess the decision-to-bid whether a project fits with the available man-hours (Vanderlande, 2017).

3. **Contribution Margin (CM),** i.e., the percentage of contribution margin is calculated by dividing the absolute contribution margin by the project price, as can be found in Figure 4 (page 22). The contribution margin includes the profit, risk, and overhead of a project. Where overhead is by far the largest component, and is typically a fixed percentage. Hence, the contribution margin can be seen as a measure of profitability of a project, i.e., a higher CM relates to more profit.

   Explanation: Information regarding the expected contribution margin (c.q. profit) can be used to assess the decision-to-bid for various tender projects, for instance, to pursue the BHS project(s) with the highest CM percentage.

In addition to the aforementioned response variables, insights regarding the 'win ability' of a bid, based on historical bids, would be of significant value. In which 'win ability' is defined as the chance of winning a certain bid. Unfortunately, this analysis was impossible to address in this study, as project data of lost bids was rarely stored in the systems.

3.5. Predictor variables

Based on interviews with sales and operational departments, various predictor variables have been identified which might influence the quotation costs, man-hours or contribution margin of a project. As expected, only a portion of the mentioned variables was logged in Vanderlande's CRM and related IT systems. The predictor variables, which were available for the analysis were:

- **CRM sales value (Euro),** i.e., the initial sales value of the BHS in Euro, which is logged in the CRM system early in the sales phase.
- **Number of motors**, i.e. the total number of motors which are used in the BHS. This value is stored in a pricing system.

- **The length of belts (meter),** i.e. the total length of the belts which are used in the BHS project. This value is stored in a pricing system.
- **The number of electrical connections**, i.e., the total number of electrical connections in the BHS. This value is stored in a pricing system.
- **Delta sales**, i.e. the difference between the CRM sales value and the sales value of the signed contract. For instance, when CRM sales value is € 2.000.000 and the signed value is € 2.300.000 then the delta sales is € 300.000.
- **Project type**, i.e., the BHS project type which was sold (categorical). These types were labeled A, B or C.
- **Customer center**, i.e., the customer center from which the BHS was sold (categorical). This can be Europe-Middle-East-Africa (EMEA), North-America (NA), United-Kingdom (UK) and Asia-Pacific (APAC).

A comprehensive list of predictor variables identified in the various interviews can be found in Appendix VII. In addition, Appendix VII also provides general recommendations regarding the storage of project data.

3.6. Variables over time

Early in the tender process, information regarding the project is scarce. Over time, gradually more information is identified, as can be seen in Figure 6.

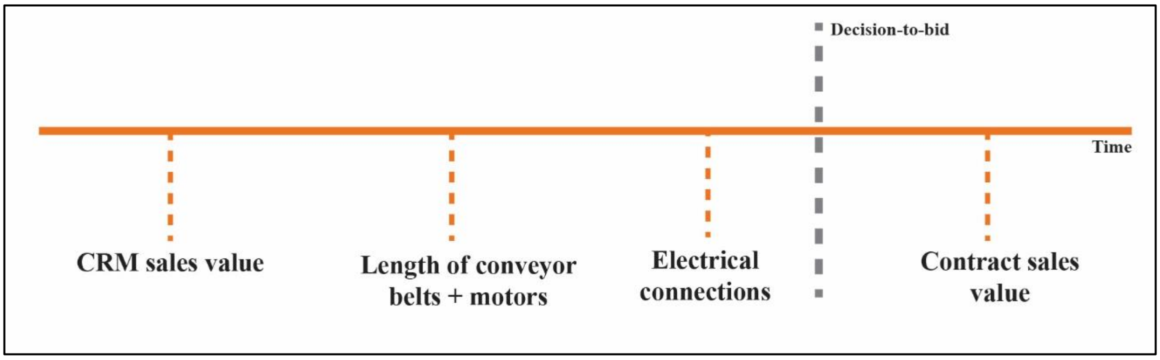


Figure 6. *Identification of information during tendering for projects.*

Based on this practical finding, four predictive models will be made per response variable:

- Model 1: Y ~ CRM sales value
- Model 2: Y ~ CRM sales value + motors + length
- Model 3: Y ~ CRM sales value + motors + length + connections
- Model 4: Y ~ CRM sales value + motors + length + connections + delta sales

3.7. Variable selection methods

As we intend to make four predictive models per response variable, we want to select the 'best' subset of the variables to be enclosed in the models. Adding and removing individual predictor variables manually can become rather difficult, as there are multiple subsets possible, e.g., with a choice of five predictor variables there are 32 possible subsets ($2^5 = 32$). Fortunately, there are two methods for choosing the 'best' subset of predictor variables, namely, the all possible subsets method and stepwise methods (backward elimination and forward selection) (Sheather, 2009). An explanation of both variable selection methods can be found in Appendix VIII.

This study will adopt a hybrid stepwise method, which is a combination of the backward stepwise elimination and the forward stepwise selection method (James, Witten, Hastie, & Tibshirani, 2013). The hybrid method adds variables, while it can also eliminate variables that no longer provide an improvement to the model. The hybrid stepwise method is closely related to the 'all possible subset method', while having the computational efficiency of forward and backward stepwise selection (James et al., 2013), as can be found in Appendix VIII.

3.8. Predictability

The accuracy of the predictions by the various MLR models will be assessed by the Mean Absolute Percentage Error (MAPE). This measure is suitable for this study, as the data are positive and much greater than zero. In addition, by applying percentage errors, the resulting measures become scale-independent, which is beneficial to compare performance across the different MLR models. For that reason, as the series in this study have different scales, the MAPE is preferred over, for instance, the Mean Absolute Error (MAE) or the Mean Square Error (MSE) (Goodwin & Lawton, 1999).

The MAPE is defined as the average of percentages errors and expressed as:

$$MAPE = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{A_t - P_t}{A_t}\right|$$

Where A_t is the actual value and P_t is the predicted value.

Note that the MAPE uses the absolute values prior to averaging. This gives a better understanding of the true error rate of the prediction. However, as the MAPE takes the absolute value, only the true error in terms of percentage is known, not whether the model under or over predicts. For the purpose of this study, we argue that this is not a large problem, as we are interested in providing valid predictive models with the smallest true error, and not whether the model under or over predicts.

4. DATA ANALYSIS RESULTS

The gathered data was brought together in a single master data set, which contained 1504 project numbers of Vanderlande over the past seven years. At first glance, this seemed like a huge dataset, however, many of the project numbers in the master data set were found to be not relevant for the analysis. For instance, some project numbers only represented ‘replace cranes’ or ‘update software’, instead of ‘BHS Airport X’. Therefore, a filter was applied to only include BHS projects which were found to be relevant. The filter required that the observations should at least contain a ‘motor’, a ‘length of the belts’ for the system, ‘electrical connections’ in the system, and should be present in the CRM system. In addition, depending on the response variable, requirements were added to the filter, e.g., whether the project was finished or not. This resulted in different datasets per response variable.

After applying the filter, the data was investigated in the software package ‘R’. A correlation matrix was made, as can be found in Appendix X, and the data was checked for normality by means of histograms, Quantile-Quantile plots (QQ-plots) and Shapiro-Wilk tests. Many of the response and the predictor variables were found to have high levels of skewness and kurtosis (long right tail), which was not beneficial for future analyses. Therefore, variables with a high level of skewness were log-transformed to behave more like a normal distribution, i.e., where log refers to ‘log to the base *e’* or natural logarithms. An example of the effects of a log-transformation on the response variable ‘quotations costs’ can be found in Figure 7.

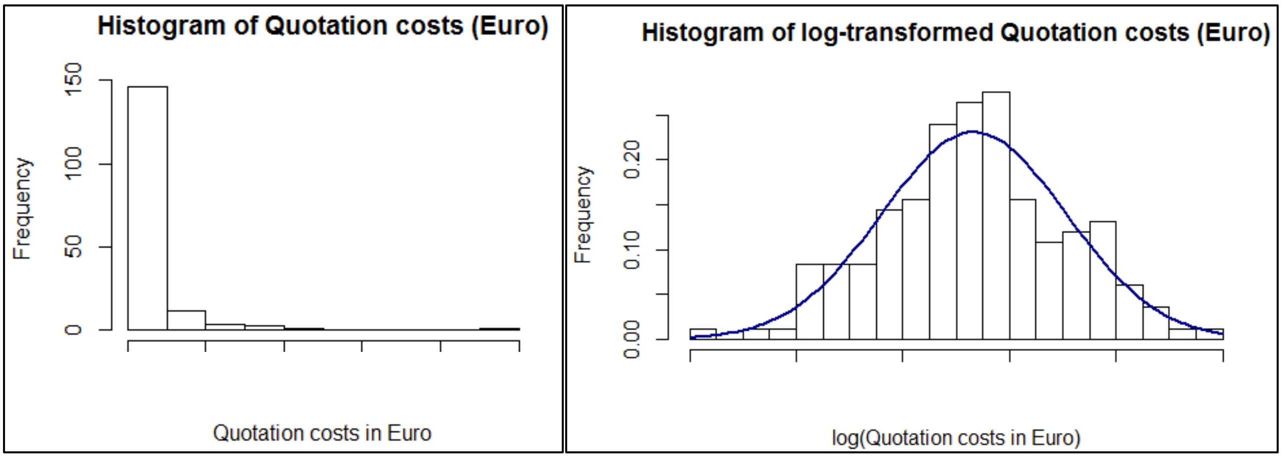


Figure 7. *Left histogram showing the long tail to the right (skewness & kurtosis), while the right histogram shows the log-transformed data, behaving more like a normal distribution.*

4.1. Predicting Quotation costs

Quotation costs are the costs belonging to the development of a bid. These costs may contain the costs of the sales hours, or the costs of the hours from any other department supporting the bidding process, e.g., operational or legal departments. Knowledge regarding the expected quotation costs is believed to be valuable, for instance, to properly allocate the limited sales resources over the available tenders (Vanderlande, 2017c). In addition, once the estimated quotation costs are known, the sales department can monitor the sales process more closely.

Prior to applying MLR for the response variable 'quotation costs', the Mahalanobis distance was calculated to detect outliers in the dataset. Based on visual assessments using boxplots, and a rule of thumb $\left(\frac{Mahalanobis\ Distance}{Nr\ of\ Variables} > 3\right)$, outliers where removed from the analysis. After the outlier removal, 81 of the 84 observations remained in the dataset.

Then, scatter plots were drawn to visually assess whether the relationship between the response and predictor variables was linear. An example of the response variable Quotation costs (log-transformed) versus the predictor variable CRM sales value (log-transformed) can be found in Figure 8. No visual evidence was found which violated the MLR assumption of linear relationships, e.g., no quadratic or cubic relationships were found.

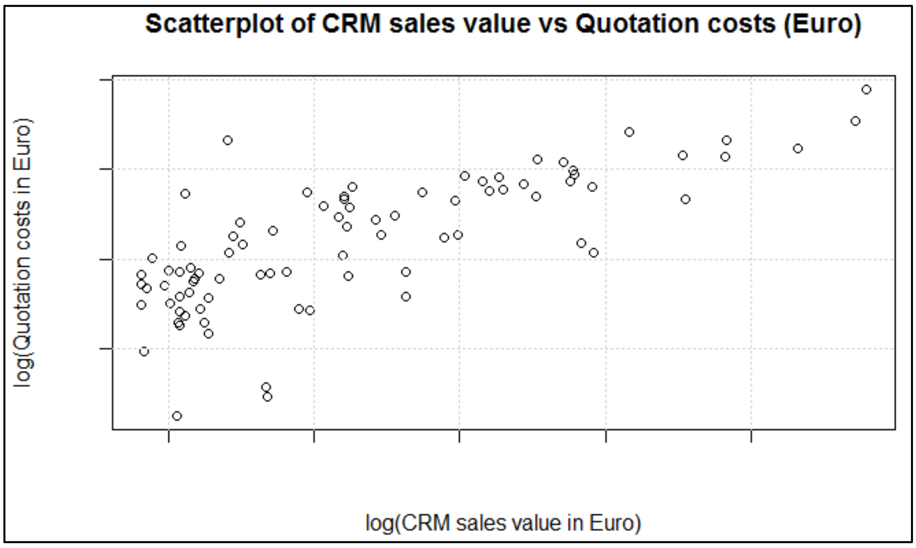


Figure 8. *Scatter plot of CRM sales value vs quotation costs, both log-transformed. The scatter plot shows evidence of a linear relationship.*

As stated in the method sector, four MLR models per response variable will be made, such that the models can be applied over time, i.e., early in the process when little information is known, until later in the process when more information is known. Table 1 (page 37) shows the four MLR models for the log 'quotation costs'; both the full models and the models resulting from the hybrid stepwise regression are shown. Prior to interpreting the MLR results, the validity of the models was checked, as is explained in the next section.

4.1.1. Validation of the quotation cost models

The validity of the models was checked in four ways. First, by assessing the residual plots. The residual plots showed no clear structure, therefore, we stated that the models passed the first test, an example of model 1 can be found in Figure 9.

The second assumption, whether the errors were normally distributed, was checked by means of a normal Q-Q plot, an example of model 1 can be found in Figure 9. As can be seen in the Figure, the majority of the points are on, or close to the line in the Q-Q plot, suggesting that the data is normally distributed, hence the assumption was met.

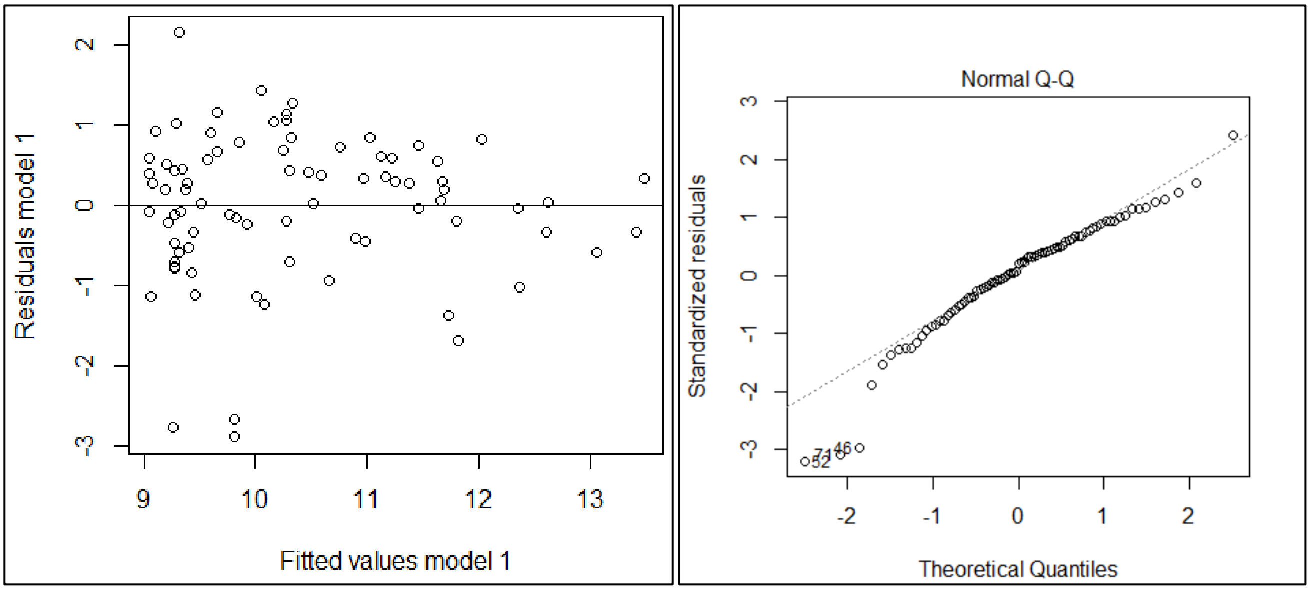


Figure 9. *Residual plot (on the left) and normal a Q-Q plot (on the right) for model 1.*

Thirdly, the assumption of constant variance of the errors was checked by a formal Breusch-Pagan test. The Breusch-Pagan test examines whether heteroscedasticity is present. The tests indicated homoscedasticity (p-value > 0.05), therefore, the third assumption was also met.

Lastly, linearity was checked by means of added variable plots. No evidence/trends were spotted to include quadratic terms. As a result, based on the checked assumptions, we conclude that the predictive MLR models for the log-transformed quotation costs are valid.

Table 1.
Summary of MLR Analysis for variables predicting log Quotation Costs in Euro (N=81). Both the full models, as the (hybrid) stepwise regression models are shown.

	Full MLR models			**MLR models after hybrid stepwise variable selection (per model)**		
Model 1	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	1.20		B_0	1.20	
log Sales Value CRM	B_1	0.08	<0.000	B_1	0.08	<0.000
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.621	0.616		0.621	0.616	60%
Model 2	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	1.59		B_0	1.59	
log Sales Value CRM	B_1	0.14	<0.000	B_1	0.14	<0.000
log Nr of Motors	B_2	0.17	<0.000	B_2	0.17	<0.000
log Length of Belts	B_3	0.16	0.002	B_3	0.16	0.002
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.711	0.699		0.711	0.699	49%
Model 3	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	1.8		B_0	1.40	
log Sales Value CRM	B_1	0.14	<0.000	B_1	0.13	<0.000
log Nr of Motors	B_2	0.45	0.695	-	-	-
log Length of Belts	B_3	0.16	0.010	B_3	0.14	0.005
log Nr of Elec. Connections	B_4	0.36	0.121	B_4	0.13	<0.000
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.720	0.705		0.719	0.708	46%
Model 4	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	1.96		B_0	1.40	
log Sales Value CRM	B_1	0.15	<0.000	B_1	0.13	<0.000
log Nr of Motors	B_2	0.47	0.736	-	-	-
log Length of Belts	B_3	0.16	0.010	B_3	0.14	0.005
log Nr of Elec. Connections	B_4	0.36	0.120	B_4	0.13	<0.000
log Delta Sales	B_5	0.01	0.835	-	-	-
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.720	0.701		0.719	0.708	46%

Notes:

- *Delta Sales was calculated by subtracting the 'Sales Value CRM' from the 'Signed Contract Value'.*
- *The MAPE has been calculated by splitting the dataset (81 obs.) at random into a trainset (65 obs.) and a testset (16 obs.). The data from the trainset was used to estimate the regression coefficients using the stepwise regression method. Then, data from the testset was used to predict the response variable, followed by calculating the MAPE. This process was repeated ten times and averaged.*

4.1.2. Interpretation of MLR results obtained via stepwise regression

In model 1 (Table 1), an MLR model was calculated to predict the log-transformed quotation costs based only on the log-transformed CRM sales value. A valid regression equation was found, with a R^2 of 0.62, i.e., 62% of the total sample variability in the response variable was explained by the model. Based on the regression equation of model 1, the log-transformed quotation costs can be predicted by the following formula:

$$log(Quotation\ costs\ in\ Euro) = \beta_0 + \beta_1 * log(Sales\ value\ CRM\ in\ Euro) + \varepsilon$$

Interpretation of the formula is based on elasticity, i.e., a BHS project with 1% more log(sales value CRM), relates to β_1% more log(quotation costs).

In model 2, the predictor variables ‘log Nr of motors’ and the ‘log Length of belts’ were added to the MLR model. This resulted in an increase from a R^2 of 0.62 towards a R^2 of 0.71. More importantly, when we compare model 2 with model 1, we find that the $R^2_{Adj.}$ increased from 0.62 (model 1) to 0.69 (model 2). Therefore, we can state that the added predictor variables improved the MLR model.

After adding predictor variable ‘log Nr of Elec. Connections’ in model 3, the significance of the added variable was found to be insignificant. However, after applying stepwise regression, ‘log Nr of Motors’ was removed from the model, and ‘log Nr of Elec. Connections’ did improve the model. The $R^2_{Adj.}$ increased slightly from 0.699 towards 0.708.

Finally, in model 4, ‘log Delta Sales’ was added, which was found to be insignificant. Also after applying stepwise regression the ‘log Delta Sales’ was removed. Hence, the remaining model was equal to model 3.

For predictive actions, we used the MLR models resulting from the stepwise regression method, as these included the variables which improved the MLR models significantly. Clearly, the predictive capability increased over time. Early in the sales phase, when only CRM sales values are known, the MAPE is 60%. While the last model (model 3) has a MAPE of 46%. Further improvements are likely to occur once more (categorical) variables are included, for instance, the BHS project type (3 types) and the customer center (4 locations). However, given the size of the data set (81 obs.) we were unable to incorporate these categorical variables, as this would cause issues for the statistical analysis, i.e., the resulting data subsets (81 obs. divided by 9) would be too small for the statistical tests, such as an analysis of covariance. Even adding one categorical variable proved to be problematic, as the dataset was highly unbalanced, i.e., 60 of the 81 observations (75%) were from a single customer center, making it hard to conduct the statistical tests between the two other groups (11 and 13 obs.)

Summarizing, the obtained MLR models for predicting quotation costs proved to be valid models. Therefore, by making use of the MLR models, Vanderlande can predict the quotation costs during tendering for projects at a very early stage, prior to the decision-to-bid.

In addition, with respect to the main research question, we found evidence that Vanderlande's CRM data can be used to support the decision-to-bid during tendering for projects.

4.2. Predicting man-hours for the Mechanical Engineers

Prior to developing of a bid, it is important to have an idea of the required man-hours to conduct a project. For instance, to assess whether a new project fits with the available man-hours in a certain period, or to train/hire/inform employees in advance (Vanderlande, 2017c).

This section will focus on the man-hours for the mechanical engineers, which is one of the functional roles within Vanderlande. Typically, once the contract is signed, there is not much time to adjust the workforce anymore as the mechanical engineers start right after the project is won. Hence, having a prediction of the expected man-hours at an early stage (i.e. the decision-to-bid) was found to be valuable to be able to act in advance (Vanderlande, 2017c)

Prior to applying MLR for the response variable log(Mechanical Engineering hours), the Mahalanobis distance was calculated to detect outliers in the dataset. Based on visual assessments using boxplots, and a rule of thumb $\left(\frac{Mahalanobis\ Distance}{Nr\ of\ Variables} > 3\right)$ outliers where removed from the analysis. After the outlier removal, 82 of the 85 observations remained in the dataset. Then, scatter plots were drawn, to visually assess whether the relationship between the response and predictor variables was linear. An example of the response variable 'log(Mechanical Engineering hours)' versus the predictor variable 'log(CRM sales value)' can be found in Figure 10. No visual evidence was found which violated the MLR assumption of linear relationships, e.g., no quadratic or cubic relationships where found.

As stated in the method sector, four MLR predicting models will be made per response variable, such that the models can be applied over time, i.e., early in the process when little information is known, until later in the process when more information is known. Table 2 (page 43) shows the four MLR models for 'log (Mechanical Engineering hours)'; both the full models and the models resulting from the hybrid stepwise regression are shown. Before interpreting the MLR results, the validity of the models was first checked, as is explained in the next section.

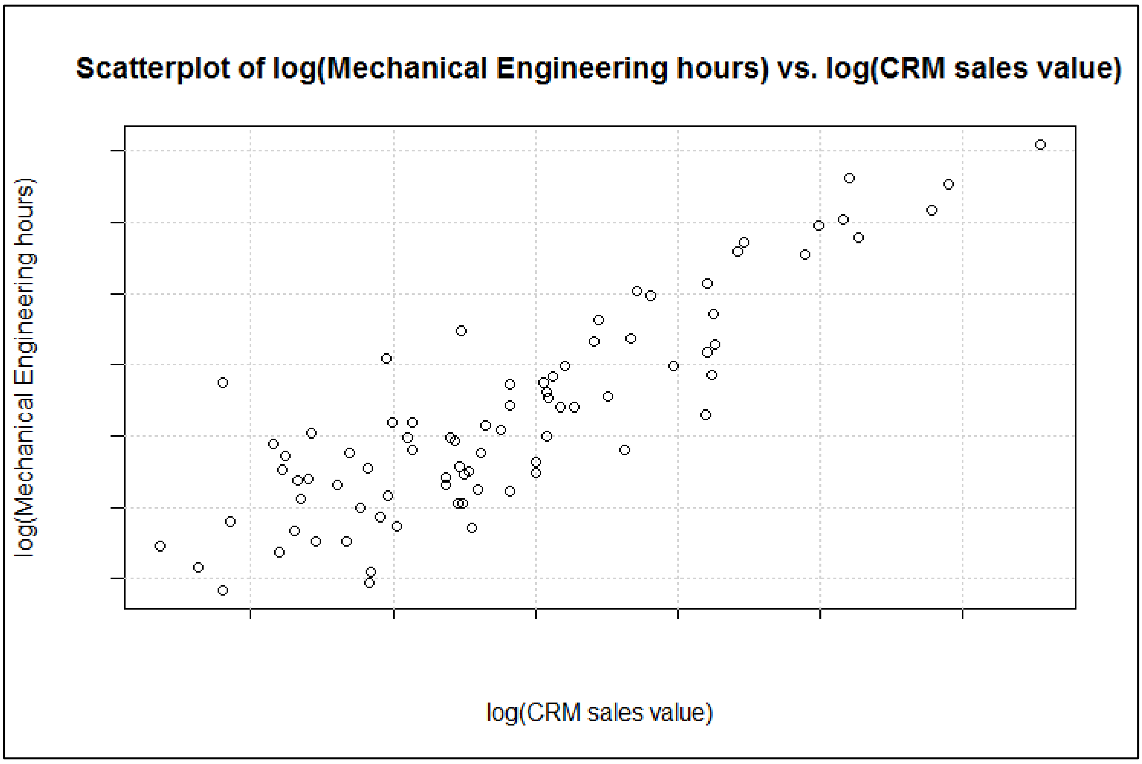


Figure 10. *Scatter plot of log(CRM sales value) vs log(Mechanical Engineering hours). The scatter plot indicates a linear relationship.*

4.2.1. Validation of the Mechanical Engineering hours models

The validity of the models was checked in four ways. First, by assessing the residual plots. The residual plots showed no clear structure, therefore, we stated that the models passed the first test. An example of model 1 can be found in Figure 11.

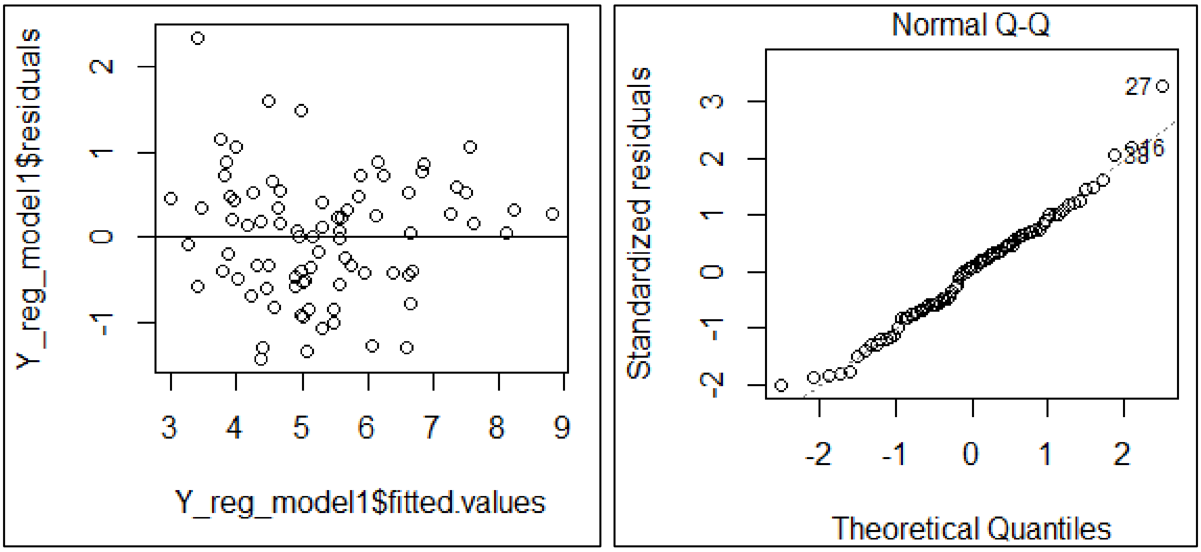


Figure 11. *Residual plot (on the left) and a normal Q-Q plot (on the right) for model 1 (log(Mechanical Engineering hours) based on log(CRM sales value)).*

The second assumption, whether the errors were normally distributed, was checked by means of a normal Q-Q plot, an example of model 1 can be found in Figure 11. The majority of the points were found to be on the line in the normal Q-Q plot, hence, the assumption was met.

Thirdly, the assumption of constant variance of the errors was checked by a formal Breusch-Pagan test. The Breusch-Pagan test examines whether heteroscedasticity is present. The tests indicated homoscedasticity (p-value > 0.05), therefore, the third assumption was also met.

Lastly, linearity was checked (again) by means of added variable plots. No evidence/trends were spotted to include quadratic terms. As a result, based on the checked assumptions, we conclude that the predictive MLR models for the 'log(Mechanical Engineering hours)' are valid.

4.2.2. Interpretation of MLR results obtained via stepwise regression

In model 1 (Table 2), an MLR model was calculated to predict the 'log(Mechanical Engineering hours)' based only on the 'log(CRM sales value)'. A valid regression equation was found, with a R^2 of 0.75, i.e., 75% of the total sample variability in the response variable was explained by the model. Based on the regression equation of model 1, the 'log(Mechanical Engineering hours)' can be predicted by the following formula:

$$lo\,g(Mechanical\ Engineering\ Hours) = \beta_0 + \beta_1 * log(Sales\ value\ CRM\ in\ Euro) + \varepsilon$$

Interpretation of the formula is based on elasticity, i.e., a BHS project with 1% more log(sales value CRM) relates to β_1% more 'log(Mechanical Engineering hours)'.

In model 2, the predictor variables 'log Nr of motors' and the 'log Length of belts' were added to the MLR model. This resulted in an increase from a R^2 of 0.75 towards a R^2 of 0.82. More importantly, when we compare model 2 with model 1, we find that the $R^2_{Adj.}$ increased from 0.75 (model 1) to 0.82 (model 2). Therefore, we can state that the added predictor variables improved the MLR model.

After adding predictor variable 'log Nr of Elec. Connections' in model 3, the significance of the added variable was found to be insignificant.

Table 2.

Summary of MLR Analysis for variables predicting ***log Mechanical Engineering Hours*** *(N=82). Both the full models, as the (hybrid) stepwise regression models are shown.*

	Full MLR models			MLR models after hybrid stepwise variable selection (per model)		
Model 1	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	0.84		B_0	0.84	
log Sales Value CRM	B_1	0.06	<0.000	B_1	0.06	<0.000
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.751	0.750		0.751	0.750	56%
Model 2	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	1.04		B_0	1.04	
log Sales Value CRM	B_1	0.10	<0.000	B_1	0.10	<0.000
log Nr of Motors	B_2	0.11	0.020	B_2	0.11	0.020
log Length of Belts	B_3	0.12	0.019	B_3	0.12	0.019
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.823	0.816		0.823	0.816	48%
Model 3	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	1.13		B_0	1.04	
log Sales Value CRM	B_1	0.10	<0.000	B_1	0.10	<0.000
log Nr of Motors	B_2	0.30	0.061	B_2	0.11	0.020
log Length of Belts	B_3	0.12	0.035	B_3	0.12	0.019
log Nr of Elec. Connections	B_4	0.26	0.264	-	-	-
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.826	0.817		0.823	0.816	48%
Model 4	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	1.04		B_0	0.96	
log Sales Value CRM	B_1	0.09	<0.000	B_1	0.09	<0.000
log Nr of Motors	B_2	0.27	0.139	B_2	0.10	0.069
log Length of Belts	B_3	0.11	0.016	B_3	0.11	0.009
log Nr of Elec. Connections	B_4	0.24	0.374	-	-	-
log Delta Sales	B_5	0.01	<0.000	B_5	0.01	<0.000
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.856	0.847		0.855	0.847	42%

Notes:

- *Delta Sales was calculated by subtracting the 'Sales Value CRM' from the 'Signed Contract Value'.*
- *The MAPE has been calculated by splitting the dataset (82 obs.) at random into a trainset (66 obs.) and a testset (16 obs.). The data from the trainset was used to estimate the regression coefficients using the stepwise regression method. Then data from the testset was used to predict the response variable, followed by calculating the MAPE. This process was repeated ten times and averaged.*

Finally, in model 4, 'log Delta Sales' was added, which was found to be significant. Also after applying stepwise regression, 'log Delta Sales' remained in the model. Moreover, the $R^2_{Adj.}$ increased from 0.82 (model 2 and 3) to 0.85 (model 4). Therefore, we can state that the added predictor variable 'log Delta Sales' improved the MLR model.

For predictive actions, we used the MLR models resulting from the stepwise regression method, as these included the variables which improved the MLR models significantly. Clearly, the predictive capability increased over time. Early in the sales phase, when only CRM sales values are known, the MAPE was found to be 56%. While the last model (model 4) has a MAPE of 42%. Further improvements are likely to occur once more (categorical) variables are included, for instance, the BHS project type (3 types) and the customer center (4 locations). However, given the size of the data set (82 obs.) we were unable to incorporate these categorical variables, as this would cause issues for the statistical analysis, i.e., the resulting data subsets (82 obs. divided by 12) would be too small for the statistical tests, such as an analysis of covariance. Even adding only one categorical variable proved to be problematic, as the dataset was highly unbalanced, i.e., 65 of the 81 observations (80%) were from a single customer center, making it hard to conduct the statistical tests between the three other groups (1, 5 and 14 obs.)

Summarizing, the obtained MLR models for predicting the Mechanical Engineering hours proved to be valid models. Therefore, by making using of the MLR models during tendering, Vanderlande can predict the necessary Mechanical Engineering hours to successfully conduct a BHS project. By having an idea about the future needed man-hours, Vanderlande can train or hire additional employees during a very early stage.

In addition, with respect to the main research question, once again we found evidence that using CRM and related system data can be used to support the decision-to-bid during tendering for projects.

MLR models for the man-hours of the other roles, i.e., the Project Management, Site Management, Project Leader Engineering and Low Level Controls can be found in Appendix XI.

Unfortunately, due to a lack of remaining observations after applying the filter, we were not able to obtain valid MLR models for Integration Management (7 obs.), High Level Control (14 obs.) and Project Director Hours (0 obs.). This deficiency in observations can be explained by to the following:

- *The registration of the Integration Management hours was changed a few years ago, hence, only the most recent data proved to be usable.*
- *Both 'High Level Control' and 'Project Director' are only used in a relatively small number of projects, hence, the number of observations was rather small.*

4.3. Insights regarding the Contribution Margin

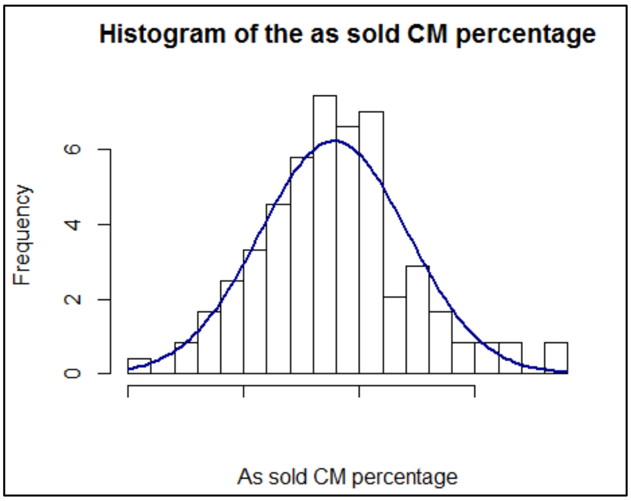
The contribution margin (CM) includes the profit, risk, and overhead of a project. In which overhead is by far the largest component, and is typically a fixed percentage. Consequently, the CM can be seen as a measure of profitability of a project, i.e., a higher CM relates to more profit. Insights regarding the CM were found to be valuable during tendering for projects, for instance, to only pursue tenders with a reasonable profit (Vanderlande, 2017c).

Interestingly, the CM is influenced to a large extent by the adjudication process, as can be found in Appendix V. In the adjudication process, the CM can be adjusted downward as well as upward. Obviously, the final bid price should be stated above the project costs to make a profit, but below perceived value to have a chance of winning (Hackett et al., 2007). Hence, the final project price is not merely based on the costs of a project, but rather market-oriented and established on intuition (Mochtar & Arditi, 2001).

Importantly, adjudication generally only affects the contribution margin, and not the variable costs (e.g., mechanical engineering hours) (Laryea & Hughes, 2008). Therefore, we found it more appropriate to label this section as 'gathering insights', rather than obtaining 'predictive models', as the CM values are typically based on intuition, and are not merely based on the project characteristics. Consequently, the layout of this section is slightly different from the previous two sections.

To gather insights regarding the contribution margin (CM) of a BHS project, a filter was applied to only include completed BHS projects. Unfortunately, the resulting dataset proved to be too small to incorporate categorical variables, such as the BHS project types (3 types) and the customer centers (4 customer centers), i.e., splitting the dataset would give problems for the statistical tests. Fortunately, the company systems did store the Latest Estimate (LE) of the CM percentage. Consequently, by selecting the LE CM, we were able to enlarge the dataset significantly, from 81 to 126 projects. This approach was supported by Vanderlande, as the LE CM usually gives a good indication whether the 'as sold' CM percentage would be reached, or not.

Based on visual assessments using boxplots, and a rule of thumb $\left(\frac{Mahalanobis\ Distance}{Nr\ of\ Variables} > 3\right)$ outliers where removed from the analysis. After the outlier removal, 121 of the 126 observations remained in the dataset. Then, summary tables were made to better understand the data, and the data was checked on normality by means of histograms and Shapiro-Wilk tests. Both the 'As sold CM percentage' and the 'LE CM percentage' were found to behave like a normal distribution, as can be found in Figure 12. Based on this, we decided not to transform these response variables.



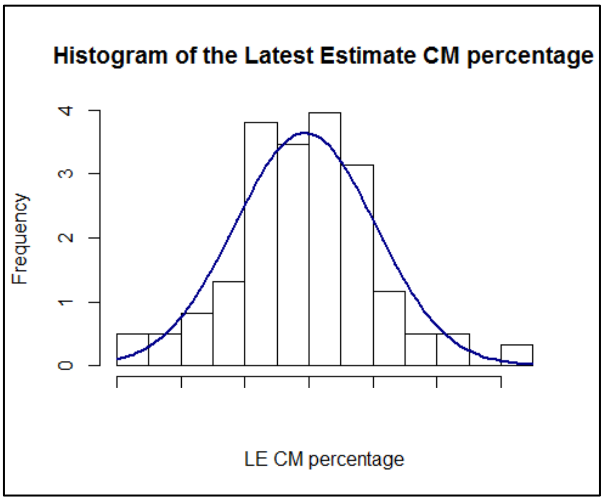


Figure 12. *Left histogram showing the As sold CM percentage, while the right histogram shows the Latest Estimate CM percentage, both behaving more or less like a normal distribution.*

To gather insights to support the decision-to-bid, three questions regarding the CM percentage have been chosen to investigate in more detail. These topics will be clarified and investigated in the following three sections:

1. *How do the project characteristics in terms of 'as sold sales value', 'customer center', and 'project type' relate to the 'As sold CM percentage' and the 'LE CM percentage'?*

To answer this question, scatterplots were drawn and MLR was applied on the dataset. The resulting scatterplots based on the 'As sold Sales value' revealed rather random patterns of points, however; the fitted MLR lines had a clear direction. (Due to confidentiality, the scatterplots are not shown, and the relation is not explained in more detail).

Two MLR models were calculated for the 'As sold CM percentage' and the 'LE CM percentage' based on the 'As sold Sales value'. A significant regression equation was found for the 'As sold CM percentage' (F(1, 119)=11.4, $p < 0.001$, with an R^2 of 0.09. The 'As sold CM percentage' can be calculated by the following formula:

$$As\ sold\ CM\ percentage\ = \beta_0 + \beta_1 * log(As\ sold\ sales\ in\ Euro) + \varepsilon$$

Similarly, a significant regression equation was found for the 'LE CM percentage' (F(1, 119)=11.4, $p < 0.001$, with an R^2 of 0.10. The 'LE CM percentage' can be calculated by the following formula:

$$LE\ CM\ percentage\ = \beta_0 + \beta_1 * log(As\ sold\ sales\ in\ Euro) + \varepsilon$$

Importantly, both models are (highly) significant, but note that the R^2 values are rather small (0.09-0.10). Therefore, we do not recommend the equations to be used for predictive actions, nor was this the purpose of this analysis. The analysis was only used to show there is a small, yet reliable relationship.

To investigate the effects of the categorical variables ‘Customer center’ and ‘Project type’ on both the ‘As sold’ and ‘Latest Estimate CM percentage’, scatterplots were made. The scatterplots revealed fitted MLR lines with rather large confidence intervals. Furthermore, the confidence intervals overlapped each other to a large extent. Scatterplots with the categorical variable ‘Customer Center’ gave similar results. Consequently, we argue that there are no additional effect of the categorical variables ‘Project Type’ and ‘Customer center’ on the ‘As sold CM percentage’. (Due to confidentiality we were not able to show the scatterplots)**.**

This statement was further supported by a statistical analysis of covariance. A one-way analysis of covariance was conducted, to determine whether a statistically significant difference was present for ‘log(As Sold sales value)’ on the ‘As Sold CM percentage’ controlling for the ‘project type’. No significant effect was found for the slopes of the regression lines of ‘log(As sold sales value) on ‘As sold CM percentage’ after controlling for ‘Project type’, $F(2,115)=0.52$, $p = 0.60$. In addition, there was no significant effect for the intercept of the regression lines $F(2,117)=0.27$, $p = 0.76$. Moreover, analysis of covariance for the response variable ‘LE CM percentage’ gave comparable results, i.e., the categorical variables ‘Project Type’ and ‘Customer center’ had no significant effect on the ‘LE CM percentage’.

Summarizing, we found a significant relationship between the ‘As sold sales value’ and the CM percentages, both in the ‘As sold’ phase and for the ‘Latest Estimate’ situation. However, we found no additional effects for the categorical variables ‘Project type’ and ‘Customer center’ on the CM values.

2. *How does the ‘Sales value per quotation hour’ relate to the ‘As sold CM percentage’ and the ‘Latest estimate CM percentage’?*

The aforementioned question was formulated based on meetings with sales and operational departments. We argued that the when Vanderlande would spend relatively less ‘quotation hours’ per project’s sales value (i.e., more sales value per quotation hour), Vanderlande would know less about the project prior to executing the project, consequently, we argued that the resulting ‘As sold CM percentage’ and the ‘Latest Estimate CM percentage’ would be lower. To test this statement, scatterplots were drawn and MLR was applied on the dataset. The scatterplots showed no relation at all: both the fitted lines of the ‘As sold CM percentage’ and the ‘Latest Estimate CM percentage’ where nearly perfectly horizontal, with rather big confidence intervals in both directions, an example for the scatterplot of the ‘As sold CM percentage’ can be found in Figure 13.

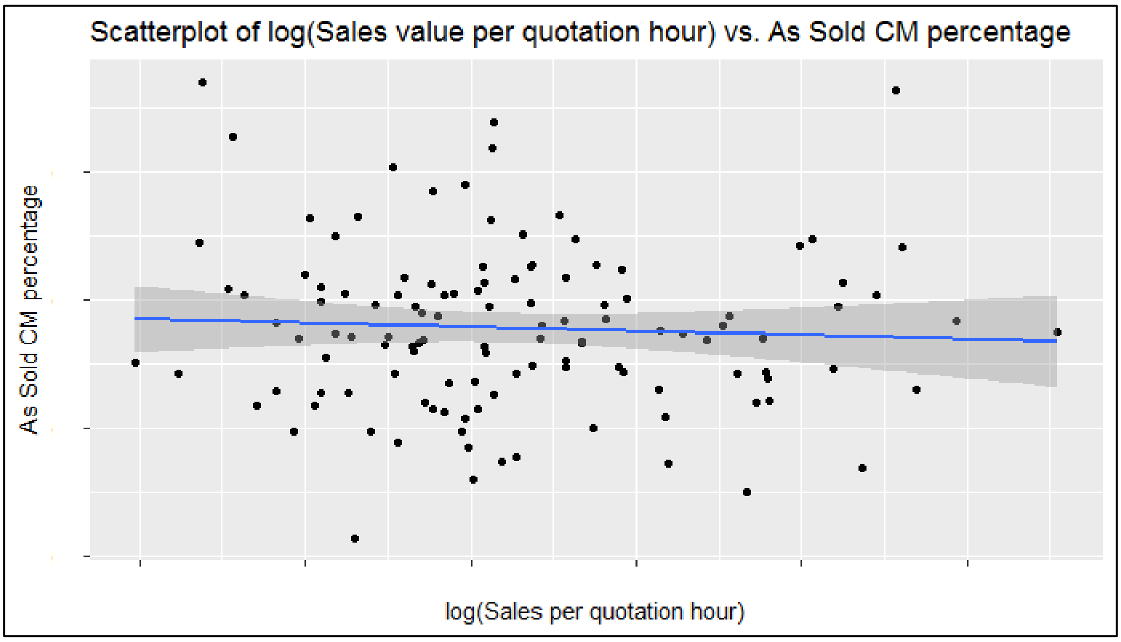


Figure 13. *Scatterplot based on 'log(Sales per Quotation hour)' vs the 'As sold CM percentage'.*

Two MLR models were calculated for the 'As sold CM percentage' and the 'LE CM percentage' based on the 'log(Sales per quotation hour)'. Two non-significant regression equations were found for the 'As Sold CM percentage' (F (1,119) =0.37, p 0.55), and 'LE CM percentage (F(1,119)=0.61, p 0.44). In addition, the effects of the categorical variables 'Customer center' and 'Project Type' were investigated using scatterplots and an analysis of covariance. The scatterplots showed large overlapping confidence intervals, and the analysis of covariance showed no significant effects of the 'Customer center' or the 'Project Type'.

To conclude this section, no significant MLR models where obtained for the 'As sold CM' and the 'Latest estimate CM', based on 'Sales per quotation hour'.

3. *What is the effect of investing (relatively) fewer quotation hours per project's sales value, during the development of a bid, on the difference between the 'As sold CM percentage' and the 'Latest estimate CM percentage'?*

The aforementioned question was formulated based on meetings with sales and operational departments. We argued that when Vanderlande would spend relatively fewer 'quotation hours' per project's sales value (i.e., capture more sales value per quotation hour), Vanderlande would know relatively less about the BHS project. We argued that this relatively lower level of knowledge could lead to increased differences between the 'As sold CM' and the 'Latest Estimate CM', both positively or negatively.

The ‘Delta CM’ was calculated by subtracting the ‘As Sold CM percentage’ from the ‘Latest Estimate CM percentage’. Where positive ‘Delta CM’ values correspond to BHS projects which turn out to have higher ‘Latest Estimate CM’ values than the ‘As sold CM percentage’. In addition, the absolute value of the ‘Delta CM’ was also calculated. Histograms of the Delta CM and the Absolute Delta CM value can be found in Figure 14.

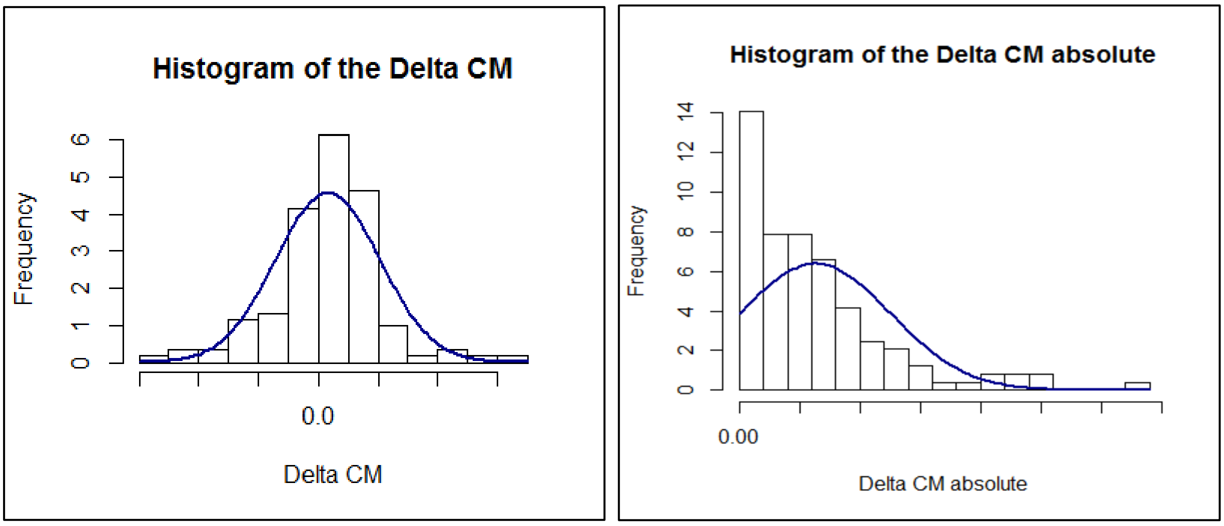


Figure 14. *Histograms of the Delta CM and the Absolute Delta CM.*

To investigate this further, scatterplots were drawn and MLR was applied on the dataset. The resulting scatterplots based on the ‘log(Sales per Quotation hour)’ vs the ‘Delta CM’ and the ‘Delta CM absolute’ revealed rather random patterns. In fact, the scatterplots showed no relation at all: both fitted lines of the ‘Delta CM’ and the ‘Delta CM absolute’ where nearly perfectly horizontal, with rather big confidence intervals in both directions, as can be found in Figure 15.

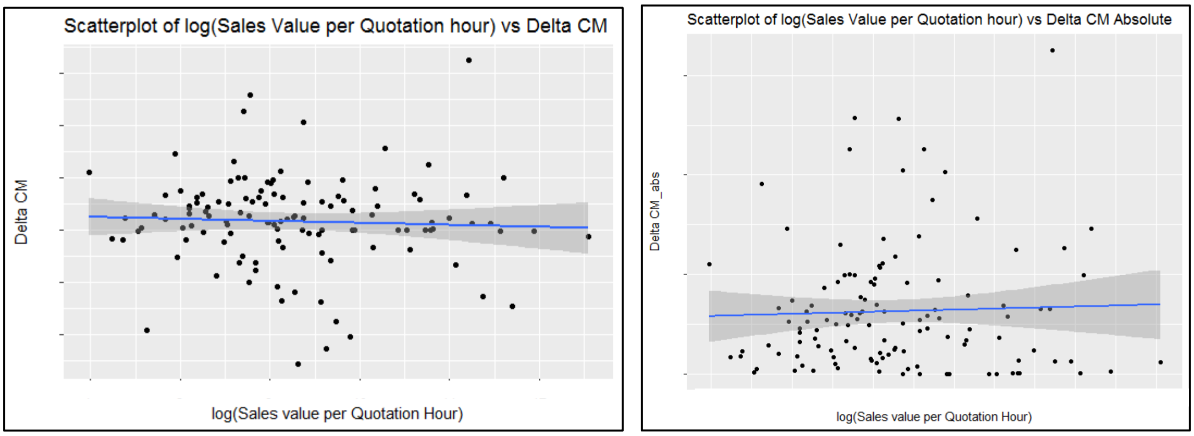


Figure 15. *Scatterplots based on the ‘log(Sales value per quotation hour)’ vs. the ‘Delta CM’ and the ‘Delta CM absolute’.*

Two MLR models were calculated for the ‘Delta CM’ and the ‘Delta CM Absolute’, based on the ‘log (Sales per quotation hour) ’. Two non-significant regression equations were found for the ‘Delta CM’ ($F(1,119) = 0.28$, p 0.60), and ‘Delta CM absolute’ ($F(1,119) = 0.17$, p 0.68). However, when we incorporated the categorical variable project type (3 types) into the analysis, the resulting scatterplots revealed some interesting fitted lines, as can be found in Figure 16.

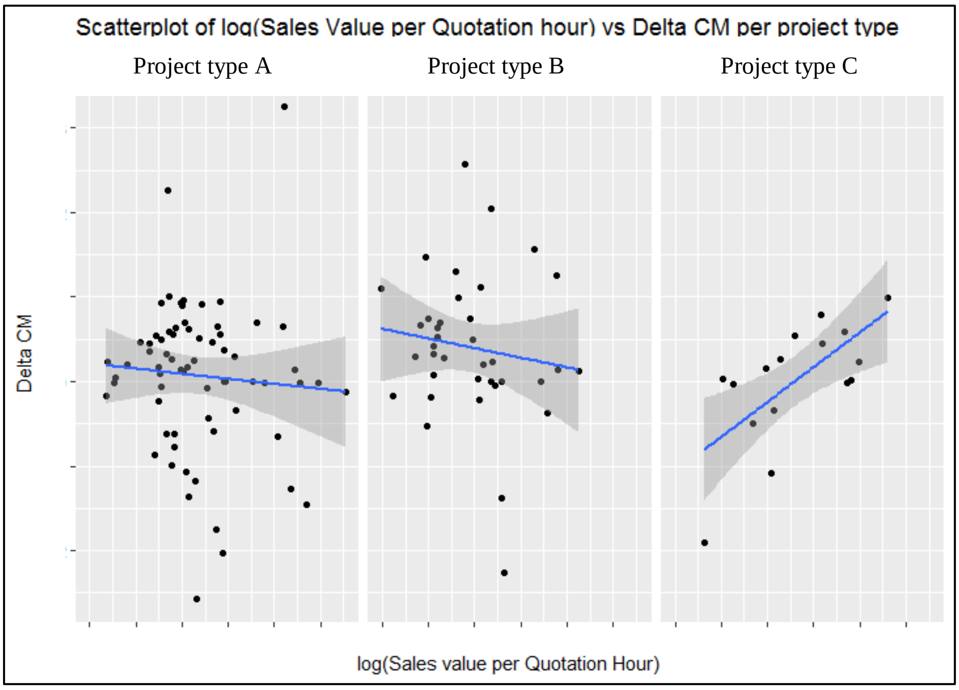


Figure 16. *Scatterplots of the 'log (Sales value per quotation hour)' vs. the 'Delta CM' per project type.*

Project type A and B showed declining fitted lines with confidence intervals overlapping to a large extent, while project type C showed an increasing fitted line. This cross-over interaction was further examined using an analysis of covariance. A one-way analysis of covariance was conducted to determine whether the regression lines 'log(Sales value per quotation hour)' on the 'Delta CM' controlling for 'project type' where statistically different from each other. There was no significant effect for the slopes of the regression lines of 'log(Sales Value per quotation hour)' on 'Delta CM' after controlling for 'Project type', $F(2,115)=2.89$, $p = 0.059$. In addition, we found no significant effect on the intercept of the regression lines $F(2,117)=1.92$, $p = 0.15$. Similar results were found for the effect of the categorical variable 'customer center', i.e., 'customer center' did not result in significantly different regression lines for 'Delta CM'.

Summarizing, no significant relationships were found between the 'Sales per quotation hour' and the difference between the 'As sold CM percentage' and 'Latest Estimate CM percentage'. Neither when we took into account the 'project type' or the 'customer center'.

To translate this to practice: based on this analysis, we did not find support for Vanderlande to invest relatively more 'Quotation hours' per project's sales value, during the development of the bid, to decrease the difference between the 'As sold CM percentage' and 'Latest Estimate CM percentage'.

5. CONCLUSIONS

Vanderlande was interested in obtaining insights regarding the quotation costs, man-hours, and contribution margins of BHS projects at a very early stage. Guided by findings from an initial literature study, we have chosen to focus on the decision-to-bid during tendering for projects, as this was both seen theoretically and practically relevant.

Prior to the research, a literature study was conducted to obtain theoretical insights regarding tendering for projects, which provided interesting insights. First, scholars argue that contractors tend to rely on subjective assessments during the decision-to-bid, rather than on quantitative data (Hwang & Kim, 2016; A. Smith, 1995). Moreover, literature suggests that it is likely that different contractors consider different factors for each project for the decision-to-bid, and that intuitive and subjective judgments change over time. Secondly, little empirical research regarding how contractors really assess the decision-to-bid is reported in literature (Philbin, 2008). As a result, scholars even argue that the subjective way of working might not have a justifiable basis (S. Laryea, 2013; S. Laryea & Hughes, 2008; Philbin, 2008). A possible reason for the lack of empirical studies might be assigned to the contractors themselves, since they are not eager to participate in such studies, due to the sensitive nature of the topic (Skitmore & Wilcock, 1994).

Taking into account the problem statements of the company and the status quo in the literature, the following research question was used in this thesis:

How should Vanderlande support the decision-to-bid, to enhance the allocation of sales and engineering resources, and to have insights regarding the contribution margin of BHS projects?

In addition to the aforementioned findings in the literature study, the literature also provided direction towards a possible solution on how contractors should support the decision-to-bid. That is, the use of contractors' knowledge-based IT systems as a basis for the decision-to-bid (Stader, 1997). In 1997, Stader argued that tendering, especially the decision-to-bid, is about gathering information and the subsequent analysis. A decade later, Metallo and colleagues (2007) stated that particularly CRM data is "*a huge information asset for the articulation of a commercial offer*".

Although promising, no study was found that actually incorporated insights, obtained from CRM and related IT systems, to support the decision-to-bid or the succeeding tender process. A possible explanation for the absence of such studies might be related to the fact that IT technologies were not widely adopted at the time of the surveys; most were conducted over two decades ago.

Fortunately, nowadays, contractors store their bidding and project data in CRM systems (Ngai, 2005; Payne & Frow, 2005). Therefore, this study focused on using contractors' IT systems, with CRM data in particular, to support the decision-to-bid during tendering for projects. This was considered both theoretically and practically relevant.

The study was exploratory and made use of archival research at Vanderlande. Moreover, it used quantitative data methods, with the goal of making predictive models and to gather insights, to be used during the decision-to-bid. Though, the availability of the project data proved to be a difficult endeavor when trying to apply more sophisticated data analysis tools (i.e. process mining). Consequently, we decided to collect project data from the various IT systems, evaluated the data on validity and combined the obtained data into a single database. Then, multiple linear regression (MLR) was applied to obtain predictive models and insights to be used to support the decision-to-bid.

5.1. Managerial implications

This study obtained valid MLR models for all three subjects, which allows Vanderlande to obtain a quick indication for various response variables at an early stage, using only a few predictor variables.

Quotation costs

Valid MLR models were obtained for the quotation costs, which explained up to 72% of the total sample variability. Hence, Vanderlande can now estimate the quotation costs at a very early stage, prior to the decision-to-bid. In addition, by having an idea of the expected quotation costs, Vanderlande can monitor the sales process more closely, i.e., whether the bid under or over performs in terms of quotation costs.

Further improvements to the quotation cost models are likely to occur once more predictor variables are added to the model. Importantly, various predictor variables have been mentioned and identified during the interviews. Though, we were unable to include all the possible variables. Either because the variables were not stored, not stored properly, or not stored for a decent period. More information regarding the identified predictor variables and recommendations for storing the data can be found in Appendix VII.

Man-hours

Valid MLR models were obtained for the man-hours of Mechanical Engineering, Project Management, Site Management, Project Leader Engineering, and Low Level Controls. As a result, Vanderlande can now estimate the necessary man-hours to successfully conduct a BHS project at an early stage, using only a few variables, and act accordingly. For instance, to check whether the estimated amount of man-hours are available during the project period or whether additional employees need to be hired/trained to conduct the project.

Producing valid MLR man-hours models for other functional roles proved to be more problematic, as we were unable to obtain sufficient observations for Integration Management (7 obs.), High Level Controls (14 obs.) and the Project Directors (0 obs.). This deficiency in observations can be explained due to a recent change in the registration of the man-hours, in addition, some of the roles were only used in a relatively small number of BHS projects.

As for the quotation costs models, further improvements to the man-hour models are likely to take place once more predictor variables are added to the model and more observations are available to apply more advanced statistics.

Contribution margin

Insights regarding the contribution margin of BHS projects were obtained using the 'latest estimate' of the contribution margin. By selecting the 'latest estimate' of the contribution margin, we included BHS projects which were not entirely completed yet. By doing so, we were able to enlarge the dataset considerably. This choice was supported by Vanderlande, as the 'latest estimate' generally gives a good impression regarding the final contribution margin. Subsequently, three subjects, related to the contribution margin, were investigated using MLR analysis.

First, based on the 'sold sales value' of the BHS projects, valid MLR models were found for the contribution margin, for both the 'as sold' and the 'latest estimate' situation. Due to confidentiality, we did not state the exact relation between the 'sold sales value' and the contribution margin (or whether the relation was positive or negative), nonetheless, we can state that the MLR models provided interesting insights. Interestingly, no additional effects for the categorical variables 'project type' and 'customer center' were found after conducting a statistical analysis of covariance. This might be due to the fact that the dataset was unbalanced, i.e., after incorporating the categorical variables the size of some subsets was quite small, resulting in rather big confidence intervals.

Secondly, we investigated whether we could obtain valid MLR models for the contribution margin, based on the 'sales value per quotation hour'. No significant MLR models where obtained, neither when we included the 'project type' nor the 'customer center' in an analysis of covariance.

And thirdly, we investigated whether we could obtain valid MLR models for the 'delta CM' (i.e., the difference between the 'as sold' contribution margin and the 'latest estimate' contribution margin) based on the 'sales value per quotation hour'. No significant MLR models where obtained, neither when we included the 'project type' and 'customer center' in an analysis of covariance. A reason for these insignificant outcomes might be due to the fact that many BHS projects have a relatively long project period (i.e., some over three years), in which various factors might have an influence on the 'latest estimate' CM value.

Although note that, by default, every project in the data set is different from each other, making it hard to invest the above phenomena in isolation.

5.2. Theoretical implications

Firstly, scholars in the field of tendering, such as Philbin (2008) and Laryea (2013), argue that empirical studies are needed to obtain new insights. As this thesis was conducted at a contractor, this study fulfills the desire of the scholars, by providing new findings based on practical research.

Secondly, this study provided evidence, by means of valid MLR models, to use contractors' quantitative system data (i.e., CRM, ERP, planning, and pricing software) to support the decision-to-bid during tendering for projects, as was proposed by Stader (1997) and Metallo and colleagues (2007). Particularly, to the best of our knowledge, this thesis is the first research that shows how initial project characteristics, such as initial sales value, can provide the contractor with useful insights regarding the quotation costs, man-hours and contribution margin at an early stage.

Thirdly, the study has identified several response variables which are believed to be important during the decision-to-bid. In addition, various (potential) predictor variables have been found and recommendations were provided to incorporate these variables in the near future.

5.3. Limitations

The study is considered exploratory, making use of archival data from the company. Importantly, as archival research makes use of administrative records, the data was not collected for the specific purpose of this research (Saunders et al., 2009). This proved to be challenging while obtaining project data for the analyses. For instance, we investigated whether we could apply process mining on the tender process, where we discovered that the IT systems did not store all the relevant process events, and timestamps were often questionable. As a result, we decided not to make use of process mining, although it seemed a promising method to explore the data.

Likewise, for the multiple linear regression analysis, obtaining the necessary data proved to be a difficult endeavor. Particularly, as the project data was stored in various IT systems, we had to build our own dataset by using project identification numbers. Consequently, we had to check whether the data was valid to make sure the dataset contained no errors. This process was found to be time-consuming and various assumptions had to be taken to produce a single dataset (e.g., the usage of a single currency rate to transform the currencies to Euros, while in practice these currencies can vary over time). In addition, for some response variables, the obtained data set was too small to obtain valid models, and many variables were logged inconsistently. Taken together, the used dataset was constructed with the best intentions, though, we had to make assumptions regarding some variables, and consequently, the dataset might still contain a few inaccuracies.

5.4. Future research

This thesis showed that, among other things, a contractor is able to obtain an idea of the needed man-hours for a project at an early stage. Though, future research should be conducted to improve the predictability of the models. Yet, this might still not be enough for practical usage. Since the models only provide information regarding the expected man-hours for the *whole* project, but not how these hours are allocated during the project period. In other words, unknown is how many hours are needed on an operational level, i.e. how many man-hours are expected on a weekly basis. Therefore, we argue that future research should not only improve the predictability of the models, but should also focus on obtaining insights for the operational level.

Additionally, we were unable to obtain insights regarding the win ability of a tender (i.e., the chance of winning a bid). We believe knowledge regarding the win ability would be valuable to support the decision-to-bid, hence, future research is needed with regard to the win ability of biddings.

Lastly, the study showed that insights can be obtained, at an early stage, using only a few predictor variables. Nonetheless, the study used project data from only one contractor. While there are only a few BHS contractors, it would be of great value to check whether the findings are more generic.

6. REFERENCES

Akintoye, A. (2000). Analysis of factors influencing project cost estimating practice. *Construction Management and Economics*, *18*(1), 77–89.

Ashworth, A., & Skitmore, R. M. (1983). *The Effectiveness of Estimating in the Construction Industry*. *Accuracy in estimating*. Englemere.

Atkinson, R. (1999). Project management: cost time and quality two best guesses and a phenomenon, it's time to accept other success criteria. *International Journal of Project Management*, *17*(6), 337–342. Retrieved from http://eprints.bournemouth.ac.uk/3187/1/licence.txt

Ballesteros-Pérez, P., González-Cruz, M. C., & Cañavate-Grimal, A. (2012). Mathematical relationships between scoring parameters in capped tendering. *International Journal of Project Management*, *30*(7), 850–862. Retrieved from http://dx.doi.org/10.1016/j.ijproman.2012.01.008

Basu, R. (2013). Managing quality in projects: An empirical study. *International Journal of Project Management*, *32*(1), 178–187. Retrieved from http://dx.doi.org/10.1016/j.ijproman.2013.02.003

Bordens, K. S., & Abbott, B. B. (2011). *Research Design and Methods*. (M. Ryan, Ed.) (8th Ed.). New York: McGraw-Hill.

Boulding, W., Staelin, R., Ehret, M., & Johnston, W. J. (2005). Management Roadmap : What Is Known , Potential Pitfalls. *Journal of Marketing*, *69*(Oct), 155–166. Retrieved from http://www.atypon-link.com/AMA/doi/abs/10.1509/jmkg.2005.69.4.155

Brook, M. (2008). *Estimating and tendering for construction work*. Elsevier/Butterworth-Heinemann.

Cao, Y., & Gruca, T. S. (2005). Reducing Adverse Selection Through Customer Relationship Management. *Journal of Marketing*, *69*(Oct), 219–229.

Carr, R. I. (1989). Cost estimating principles. *Journal of Construction Engineering and Management*, *115*(4), 545–551.

Carricano, M. (2014). Pricing myopia: do leading companies capture the full value of their pricing strategies? *Management Decision*, *52*(1), 159–178.

Cassaigne, N., Kromker, M., Singh, M. G., & Wurst, S. (1997). Decision support for effective bidding in a competitive business environment. In *IEEE International Conference* (pp. 3591–3596).

Chan, E. H. W., & Au, M. C. Y. (2007). Building contractors' behavioural pattern in pricing weather risks. *International Journal of Project Management*, *25*(6), 615–626.

CIOB. (1997). *Code of estimating practice* (6th ed.). Englemere: Longman.

Denscombe, M. (2007). *The Good Research Guide* (3rd Ed.). Buckingham: Open University Press.

Doloi, H. K. (2011). Understanding stakeholders' perspective of cost estimation in project management. *International Journal of Project Management*, *29*(5), 622–636.

Dysert, L. R. (2008). An Introduction to Parametric Estimating. In *AACE International transactions* (pp. 1–8).

Elhag, T. M. S., Boussabaine, A. H., & Ballal, T. M. A. (2005). Critical determinants of construction tendering costs: Quantity surveyors' standpoint. *International Journal of Project Management*, *23*(7), 538–545.

Eriksson, E., & Laan, A. (2007). Procurement effects on trust and control in client-contractor relationships. *Engineering, Construction and Architectural Management*, *14*(4), 387–399.

Eriksson, P. E., & Westerberg, M. (2011). Effects of cooperative procurement procedures on construction project performance: A conceptual framework. *International Journal of Project Management*, *29*(2), 197–208.

Gann, D. M., & Salter, A. J. (2000). Innovation in Project-Based, Service-Enhanced Firms: The Construction of Complex Products and Systems. *Research Policy*, *29*(7), 955–972.

https://doi.org/10.1016/S0048-7333(00)00114-1

Garrido-Moreno, A., & Padilla-Melendez, A. (2011). Analyzing the impact of knowledge management on CRM success: The mediating effects of organizational factors. *International Journal of Information Management*, *31*(5), 437–444. https://doi.org/10.1016/j.ijinfomgt.2011.01.002

Ghauri, P., & Grønhaug, K. (2005). *Research Methods in Business Studies: A practical Guide* (3rd Ed.). Harlow: Financial Times Prentice Hall.

Goodwin, P., & Lawton, R. (1999). On the asymmertry of the symmetric MAPE. *International Journal of Forecasting*, *15*, 405–408. https://doi.org/10.2466/PR0.85.7.1059-1069

Green, S. D. (1989). Tendering: Optimization and rationality. *Construction Management and Economics*, *7*(1), 53–63.

Hackett, M., Robinson, I., & Statham, G. (2007). *The aqua group guide to procurement, tendering and contract administration*. John Wiley and Sons Ltd.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate Data Analysis* (Ed. 7th). Edinburgh Gate: Pearson Education Limited.

Hardeman, D. . (2012). Aanbestedingsgedrag opdrachtgevers. Retrieved from http://www.eib.nl/pdf/aanbestedingsgedrag_opdrachtgevers_aanbesteding_en_transactiekosten_2009_2011.pdf

Heinrich, B. (2005). Transforming strategic goals of CRM into process goals and activities. *Business Process Management Journal*, *11*(6), 709–723. https://doi.org/10.1108/14637150510630873

Hensher, D. A., & Stanley, J. (2008). Transacting under a performance-based contract: The role of negotiation and competitive tendering. *Transportation Research Part A: Policy and Practice*, *42*(9), 1143–1151. https://doi.org/10.1016/j.tra.2008.05.004

Hicks, C., Mcgovern, T., & Earl, C. F. (2000). Supply Chain Management A Strategic Issue In Engineer To Order Manufacturing.pdf. *International Journal of Production Economics*, *65*, 179–190.

Hogan, J. E., Lemon, K. N., & Rust, R. T. (2002). Customer Equity management. *Journal of Service Research*, *5*(1), 4–12. https://doi.org/10.1177/1094670502005001002

Hurvich, C. M., & Tsai, C. H. (1989). Regression and time series model selection in small samples. *Biometrika*, *76*, 297–307.

Hwang, J. S., & Kim, Y. S. (2016). A bid decision-making model in the initial bidding phase for overseas construction projects. *Journal of Civil Engineering*, *20*(4), 1189–1200. https://doi.org/10.1007/s12205-015-0760-y

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning*. (G. Casella, S. Fienberg, & I. Olkin, Eds.). New York: Springer New York.

Jans, M., Van Der Werf, J. M., Lybaert, N., & Vanhoof, K. (2011). A business process mining application for internal transaction fraud mitigation. *Expert Systems with Applications*, *38*(10), 13351–13359. https://doi.org/10.1016/j.eswa.2011.04.159

Jaques, E. (2011). *Bid management*. London: Kogan Page Ltd.

Jennings, P., & Holt, G. D. (1998). Prequalification and multi-criteria selection: a measure of contractors' opinions. *Construction Management and Economics*, *16*(6), 651–660. https://doi.org/10.1080/014461998371944

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. https://doi.org/10.1080/14783360600750444

Kadefors, A. (2005). Fairness in interorganizational project relations: norms and strategies. *Construction Management and Economics*, *23*(8), 871–878. https://doi.org/10.1080/01446190500184238

Keramati, A., Mehrabi, H., & Mojir, N. (2010). A process-oriented perspective on customer relationship management and organizational performance: An empirical investigation. *Industrial Marketing Management*, *39*(7), 1170–1185. https://doi.org/10.1016/j.indmarman.2010.02.001

Khodakarami, F., & Chan, Y. E. (2014). Exploring the role of customer relationship management (CRM) systems in customer knowledge creation. *Information & Management*, *51*(1), 27–42. https://doi.org/10.1016/j.im.2013.09.001

Konijnendijk, P. a. (1994). Coordinating marketing and manufacturing in ET0 companies. *International Journal of Production Economics*, *37*(May), 19–26.

Kopczak, L. R., & Johnson, M. E. (2003). MITSloan Management Review The Supply-Chain. *MIT Sloan Management Review*, 27–34.

Krasnikov, A., Jayachandran, S., & Kumar, V. (2009). The Impact of Customer Relationship Management Implementation on Cost and Profit Efficiencies: Evidence from the U.S. Commercial Banking Industry. *Journal of Marketing*, *73*(6), 61–76. https://doi.org/10.1509/jmkg.73.6.61

Kumaraswamy, M. M., & Anvuur, A. M. (2008). Selecting sustainable teams for PPP projects. *Building and Environment*, *43*(6), 999–1009. https://doi.org/10.1016/j.buildenv.2007.02.001

Kwak, Y. H., & Watson, R. J. (2005). Conceptual estimating tool for technology-driven projects: Exploring parametric estimating technique. *Technovation*, *25*(12), 1430–1436. https://doi.org/10.1016/j.technovation.2004.10.007

Kwakye, A. A. (1994). *Understanding tendering and estimating*. Aldershot: Gower Publishing Company.

Lam, K. C., Hu, T., Ng, S. T., Skitmore, M., & Cheung, S. O. (2001). A fuzzy neural network approach for contractor prequalification. *Construction Management and Economics*, *19*(2), 175–188. https://doi.org/10.1080/01446190150505108

Laryea, S. (2013). Nature of Tender Review Meetings. *Journal of Construction Engineering and Management*, *139*(8), 927–940. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000662

Laryea, S., & Hughes, W. (2011). Risk and Price in the Bidding Process of Contractors. *Journal of Construction Management and Engineering*, *137*(4), 248–258. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000293

Laryea, S., & Hughes, W. P. (2008). How contractors price risk in bids: theory and practice. *Construction Management and Economics*, *26*(9), 911–924. https://doi.org/10.1080/01446190802317718

Lauras, M., Marques, G., & Gourc, D. (2010). Towards a multi-dimensional project Performance Measurement System. *Decision Support Systems*, *48*(2), 342–353. https://doi.org/10.1016/j.dss.2009.09.002

Li, H. (1994). Neural network models for intelligent support of mark-up estimation. *Engineering, Construction and Architectural Management*, *3*(1), 69–81. https://doi.org/10.1016/0167-9236(94)90018-3

Metallo, G., Cuomo, M. T., & Festa, G. (2007). Relationship Management in the Business of Quality and Communication. *Total Quality Management & Business Excellence*, *18*(2), 119–133. https://doi.org/10.1080/14783360601051859

Mithas, S., Krishnan, M. S., & Fornell, C. (2005). Customer satisfaction. *Journal of Marketing*, *69*(Oct), 201–209. https://doi.org/10.1016/B0-08-043076-7/04273-X

Mochtar, K., & Arditi, D. (2001). Pricing strategy in the US construction industry. *Construction Management and Economics*, *19*(January 2015), 405–415. https://doi.org/10.1080/01446190010020372

Murdoch, J., & Hughes, W. (2008). *Construction Contracts* (Fourth edi). New York: Taylor & Francis.

Ngai, E. W. T. (2005). Customer relationship management research (1992- 2002). *Marketing*

Intelligence & Planning, *23*(6), 582–605. https://doi.org/10.1108/02634500510624147

Ngai, E. W. T., Xiu, L., & Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications*, *36*(2), 2592–2602. https://doi.org/10.1016/j.eswa.2008.02.021

Nickson, D. (2003). *The bid manager's handbook*. Gower Publishing Company.

Noble, P. M., & Gruca, T. S. (1999). Industrial Pricing : Theory Practice and Managerial. *Marketing Science*, *18*(3), 435–454.

Odusote, O. O., & Fellows, R. F. (1992). An examination of the importance of resource considerations when contractors make project selection decisions. *Construction Management & Economics*, *10*(2), 137–151.

Palmatier, R. W., Gopalakrishna, S., & Houston, M. B. (2006). Returns on Business-to-Business Relationship Marketing Investments: Strategies for Leveraging Profits. *Marketing Science*, *25*(5), 477–493. https://doi.org/10.1287/mksc.1060.0209

Parvatiyar, A., & Sheth, J. (2001). Customer relationship management: Emerging practice, process, and discipline. *Journal of Economic and Social Research*, *3*(2), 1–34.

Paul, A., & Gutierrez, G. (2005). Simple probability models for project contracting. *European Journal of Operational Research*, *165*(2), 329–338. https://doi.org/10.1016/j.ejor.2004.04.005

Payne, A., & Frow, P. (2005). A Strategic Framework for Customer. *Journal of Marketing*, *69*(Oct), 167–176. https://doi.org/10.1509/jmkg.2005.69.4.167

Philbin, S. P. (2008). Bid management: A systems engineering approach. *Journal of High Technology Management Research*, *19*(2), 114–127. https://doi.org/10.1016/j.hitech.2008.10.004

Pickett, T. W. (2005). Preparing a Basis of Estimate. In *AACE International Transactions* (pp. 1–5).

Rasmusen, E. (2001). *Games and information: An introduction to game theory* (third). Oxford: Blackwell.

Richards, K. A., & Jones, E. (2008). Customer relationship management: Finding value drivers. *Industrial Marketing Management*, *37*(2), 120–130. https://doi.org/10.1016/j.indmarman.2006.08.005

Rooke, J., Seymour, D., & Fellows, R. (2004). Planning For Claims: An Ethnography of Industry Culture. *Construction Management and Economics*, *22*(6), 655–662. https://doi.org/10.1080/014461904200026324

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

Rygielski, C., Wang, J.-C., & Yen, D. C. (2002). Data mining techniques for customer relationship management. *Technology in Society*, *24*(4), 483–502. https://doi.org/10.1016/S0160-791X(02)00038-6

Saunders, M., Lewis, P., & Thornhill, A. (2009). *Research methods for business students* (Fifth Ed.). Essex: Pearson Education Limited.

Seshadri, S. (2005). *Sourcing strategy: Principles, policy and designs*. *Sourcing Strategy: Principles, Policy and Designs*. Springer Science+business Media Inc. https://doi.org/10.1007/b106908

Seshadri, S., & Mishra, R. (2004). Relationship marketing and contract theory. *Industrial Marketing Management*, *33*(6), 513–526. https://doi.org/10.1016/j.indmarman.2004.03.004

Shash, A. A. (1993). Factors considered in tendering decisions by top UK contractors Factors considered in tendering decisions by top UK contractors. *Construction Management and Economics*, *11*(2), 111–118.

Sheather, S. J. (2009). *A Modern Approach to Regression Analysis*. (G. Casella, S. Fienberg, & I. Olkin, Eds.). New York: Springer Science+business Media Inc.

Skitmore, M., Runeson, G., & Chang, X. (2006). Pricing or Neoclassical Microeconomic. *Construction Management and Economics*, *24*(7), 773–783.

Skitmore, M., & Wilcock, J. (1994). Estimating processes of smaller builders. *Construction Management and Economics*, *12*(2), 139–154. https://doi.org/10.1080/01446199400000020

Smartt, C., & Ferreira, S. (2015). Model Based Systems Engineering with Department of Defense Architectural Framework. *Systems Engineering*, *18*(1), 71–86. https://doi.org/10.1002/sys

Smith, A. (1995). *Estimating, Tendering and Bidding for construction*. Hampshire: MacMillan Press LTD.

Smith, G. R., & Bohn, C. M. (1999). Small to Medium contractor contingency and assumption of risk. *Journal of Construction Engineering and Management*, *125*(2), 101–108.

Soltani, Z., & Navimipour, N. J. (2016). Customer relationship management mechanisms: A systematic review of the state of the art literature and recommendations for future research. *Computers in Human Behavior, 61*(March), 667–688. https://doi.org/10.1016/j.chb.2016.03.008

Stader, J. (1997). An intelligent system for bid management. *The International Journal of Project & Business Risk Management*, *3*(August), 299–314. Retrieved from citeseer.nj.nec.com/stader97intelligent.html

Stewart, D. W., & Kamins, M. A. (1993). *Secondary Research: Information Sources and Methods* (2nd Ed.). Newbury Park: Sage.

Tanner, J. J., Ahearne, M., Leigh, T. W., Mason, C. H., & Moncrief, W. C. (2005). CRM in sales-intensive organizations: A review and future directions. *Journal of Personal Selling & Sales Management*, *25*(2), 169–180. https://doi.org/10.1080/08853134.2005.10749057

Teetor, P. (2011). *R Cookbook: Proven Recipes for Data Analysis, Statistics, and Graphics*. (M. Loukides, Ed.) (First Ed.). Sebastopol: O'Reilly Media.

van der Aalst, W. (2016). *Process Mining*. (2th Ed., Ed.) (Vol. 55). Springer. https://doi.org/10.1007/978-3-662-49851-4

Van Der Aalst, W. M. P., Reijers, H. A., & Song, M. (2005). Discovering social networks from event logs. *Computer Supported Cooperative Work*, *14*, 549–593. https://doi.org/10.1007/s10606-005-9005-9

van der Aalst, W. M. P., & Weijters., A. J. M. M. (2012). Process Mining. *ACM Transactions on Management Information Systems*, *3*(2), 1–17. https://doi.org/10.1145/2229156.2229157

Vanderlande. (2017a). *Contents list*. *Annual Report CY2016*. Retrieved from https://www.vanderlande.com/media/1797/annual-report-cy2016.pdf

Vanderlande. (2017b). History. Retrieved May 15, 2017, from https://www.vanderlande.com/about-vanderlande/history/

Vanderlande. (2017c). Master thesis interviews. Veghel.

Vincler, E. J., & Vincler, N. H. (1996). Producing Persuasive Proposals. *Journal of Management in Engineering*, *12*(October), 20–24.

Wanous, M., Boussabaine, A.H. and Lewis, J. (1998). Tendering factors considered by Syrian contractors. In *ARCOM, 14th Oxford. Annual Conference Proceedings* (Vol. 2, pp. 535–43).

Whitley, M. (2006). Bid, or not to bid ... *Engineering Management*, (November 2006), 40–43.

Winer, R. S. (2001). A framework for customer relationship management. *California Management Review*, *43*(4), 89–105.

Wong, C. H., Holt, G. D., & Harris, P. (2000). Multi-criteria selection or lowest price? Investigation of UK construction clients' tender evaluation preferences. *Construction Management & Economics*, *18*(7), 767–774. https://doi.org/10.1108/eb021187

Xu, Y., Yen, D. C., Lin, B., & Chou, D. C. (2002). Adopting customer relationship management technology. *Industrial Management and Data Systems*, *102*(8), 442–452. https://doi.org/10.1108/02635570210445871

Yin, R. K. (2003). *Case Study Research: Design and Method* (3rd Ed.). London: Sage.

Zeng, J., An, M., & Smith, N. J. (2007). Application of a fuzzy based decision making methodology to construction project risk assessment. *International Journal of Project Management*, *25*(6), 589–600. https://doi.org/10.1016/j.ijproman.2007.02.006

7. APPENDIX I – TENDER METHODS

Most of the projects within Vanderlande are captured via 'selective tendering', however, 'open tendering' and 'negotiation' methods for capturing projects have been found in the literature. Therefore, this Appendix explains and highlights the differences between the tender methods described in literature.

The conventional tender method – open tendering – begins with advertising the brief details of a project, consequently, interested contractors receive further information and participate in the tender process. Typically, the participating contractors might not all be suitable for a successful completion of the project, i.e. they might just be in the need for (any) work (Murdoch & Hughes, 2008, pp. 130). As a consequence, this tender method does not result in high-quality bids, and in practice, only about 5% of the contractors' submitted bids are successful (Murdoch & Hughes, 2008, pp. 130). An illustration of open tendering can be found in Figure A1.

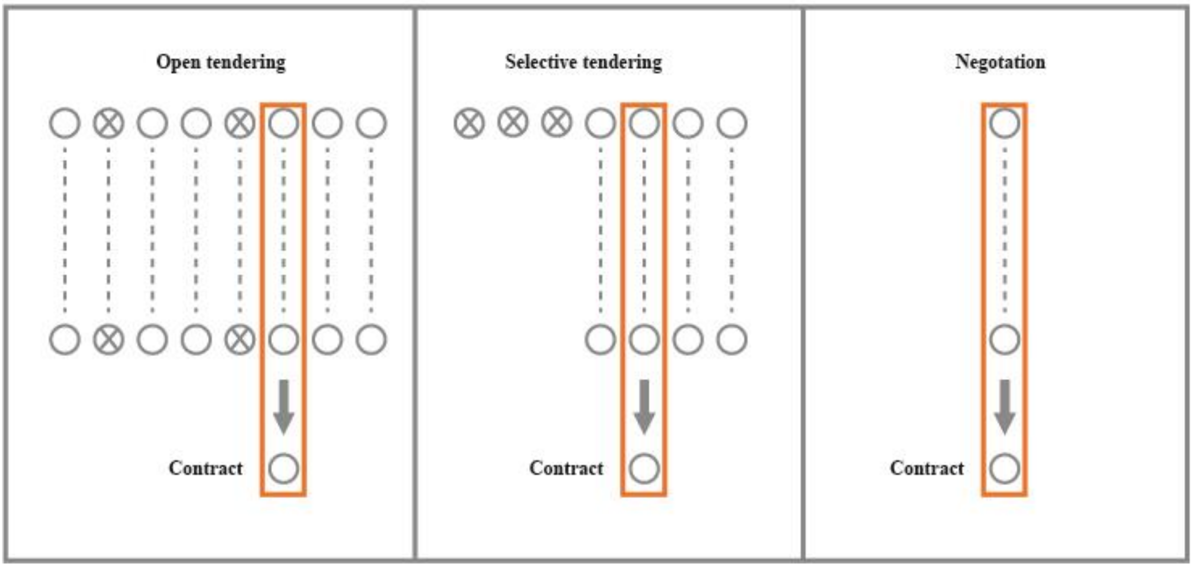


Figure A1. *General procurement methods to select an appropriate contractor for a project. Crossed circles indicate unsuitable contractors. In selective tendering, illustrated in the middle, unsuitable contractors are omitted to participate further by means of a pre-qualification. In negotiation, only one (capable) contractor is selected to realize an agreement.*

To overcome the issue of unsuitable contractors participating in the tender, selective tendering has emerged (Smith, 1995). In selective tendering, the project owner pre-qualifies a limited number of contractors who possess the necessary skills before they are formally invited to the tender, as can be found in Figure A1. (Eriksson & Westerberg, 2011). Interestingly, project owners want a large number of competing contractors in the tender, to reassure the 'best' deal at the time of the tender, whereas contractors want a small number of competing contractors, to have a reasonable chance of winning the project (Murdoch & Hughes, 2008). Consequently, although the sheer size of the project might affect the number of competing contractors, Smith (1995) argues that four to five competing contractors are proficient to execute the selective tender process successfully. For project owners, it is even advised to reveal the number of selected contractors in the tender process (Seshadri, 2005, pp. 162). An explanation for this advice is that when information is hidden from the contractors, they will assume the number of competing contractors is low; resulting in an unfavorable high bid for the project owner. Rasmusen (2001) labeled this phenomenon the “no news is bad news” thinking.

In contrast to open and selective tendering, a project owner might also consider negotiation to select a contractor. In negotiation, the project owner typically selects one capable contractor and successively comes to a contractual agreement, as can be found in Figure A1. One of the reasons for selecting this method is the desire to preserve continuing business relationships (Murdoch & Hughes, 2008, pp. 133). Moreover, the contractor and project owner can learn from each other, and reapply details from past projects in new opportunities. Interestingly, contracts achieved through tendering often evolve to negotiated contracts over time (Seshadri, 2005). However, although negotiation is being used in practice, further details are outside the scope of this Thesis, as it is not considered tendering (for more information about contracts and relationships read Seshadri & Mishra (2004)).

8. APPENDIX II – SELECTIVE TENDERING

Selective tendering is executed in four consecutive phases, i.e., the preliminary, the initiation, the development, and the evaluation phase. This Appendix will explain each phase in detail.

8.1. Selective tendering - preliminary phase

Before the actual formal tender process starts, the preliminary phase is executed. The project owner often starts by sending out a Request for Information (RFI), in which potential contractors respond to questions from the project owner, either to further specify the project requirements or refine the award/evaluation mechanism (Seshadri, 2005). During the preliminary phase, the contractors' marketing department has to opportunity to increase project owners' awareness of the project range a contractor can deliver (Hicks et al., 2000). Moreover, it allows marketing to present new technologies and improvements for consideration in project requirements. Consequently, at the end of this phase, the project owner pre-qualifies contractors to participate in the tender, mainly selected upon the projects the contractors accomplished in the past (Smith, 1995, pp. 123-126). The preliminary phase is illustrated in Figure A2.

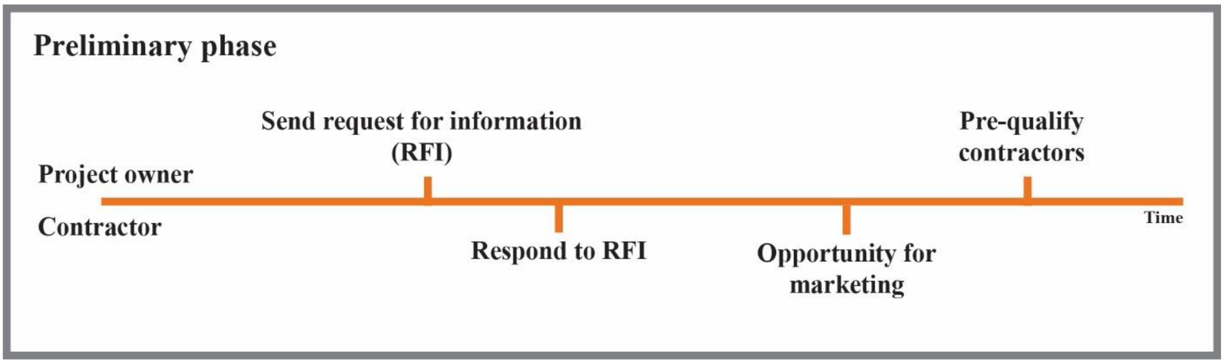


Figure A2. *Illustration of the preliminary phase, including contractor activities below the timeline, and project owner activities above the timeline.*

8.2. Selective tendering - initiation phase

The actual formal tender process is initiated once a project owner sends out an Invitation to Bid (ITB), Invitation to Tender (ITT), Request for Quotation (RFQ) or Request for Proposal (RFP) to several pre-qualified contractors (Eriksson & Westerberg, 2011; Philbin, 2008; Seshadri, 2005; Vincler & Vincler, 1996). Typically, the invitation document includes site plans, details of existing buildings and information of any unusual features or conditions. Although this information is not directly priceable, it gives the contractor context and guidance regarding the scope and associated risk of the project (Smith, 1995). In addition, often the project owner, together with consultants, perform detailed design work in order to create a more solid basis for tendering (Eriksson & Westerberg, 2011). However, in practice the amount of information can vary considerably, some project owners only provide key features, whereas others provide extensive information about design requirements, constraints, contractual conditions and performance (Seshadri, 2005; Smith, 1995). Nevertheless, although documentation might not include all details, it does give contractors an equal base, on which they can base their decision to further develop the bid, or not (Jaques, 2011). An illustration of the initiation phase can be found in Figure A3.

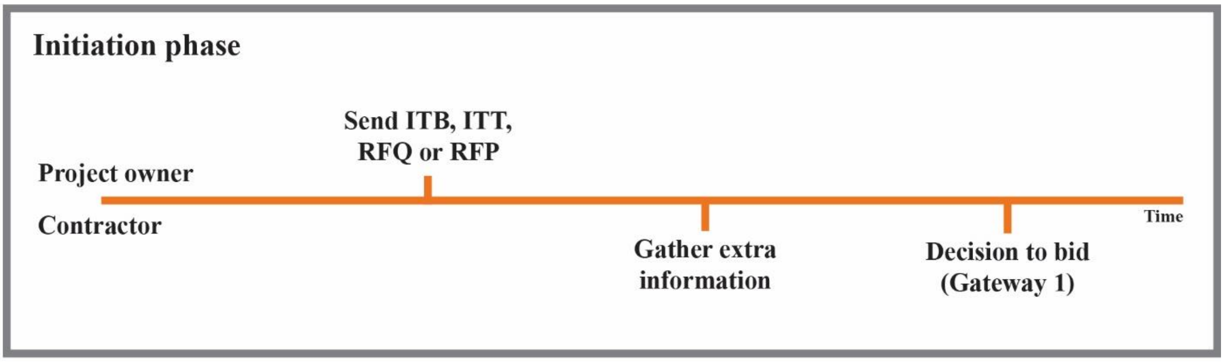


Figure A3. *Illustration of the initiation phase, including contractor activities below the timeline, and project owner activities above the timeline. (ITB = Invitation to Bid, ITT = Invitation to Tender, RFQ = Request for Quotation, and RFP = Request for Proposal)*

Within the initiation phase, the decision-to-bid can be seen as a turning point, since from that moment on, major investments in terms of sales and engineering hours are allocated to the development of the bid (Skitmore et al., 2006).

Factors influencing contractors' decision-to-bid have been identified and prioritized on their relative importance by various scholars (Odusote & Fellows, 1992; Shash, 1993; Wanous et al., 1998). Interestingly, although several studies have identified similar factors, little agreement can be found in their relative importance. For instance, Shash (1993) identified and ranked factors affecting the decision-to-bid, with 'the need for work', 'number of competitors' and 'experience in similar projects' as the most important factors, whereas Wanous and colleagues (1998) reported 'fulfilling the tender conditions', 'financial capability of the client', and 'relation with/reputation of the client' as the most important factors affecting the decision to tender. A more comprehensive overview can be found in Appendix I.
In line with these inconclusive findings, Smith (1995) labeled the search for a definitive list of factors influencing the decision-to-bid as "*a search for the Holy Grail*", as it is more likely that different contractors consider different factors for each project, and that intuitive and subjective judgments change over time. Nevertheless, although finding a definitive list of factors affecting the decision to tender might be impossible, literature does argue that contractors tend to rely on subjective assessments based on past experience and intuition in their decision-to-bid, rather than on quantitative data (Hwang & Kim, 2016).

Furthermore, invited contractors hardly ever refuse to pursue a tender, as they are anxious for not being invited again for future projects (Murdoch & Hughes, 2008, pp. 138). As a result, contractors who do not want to capture a particular project, continue in the tender process, however, they submit a *cover bid*. This cover bid is priced way above current market conditions to be acceptable (Murdoch & Hughes, 2008, pp. 138). By doing so, contractors do not have to decline a tender invitation, and still make a chance being invited in the future.

8.3. Selective tendering - development phase

Once a contractor decides to pursue a tender, it enters the development phase, where the contractor will start creating the bid. An illustration of the various activities within the development phase can be found in Figure A4.

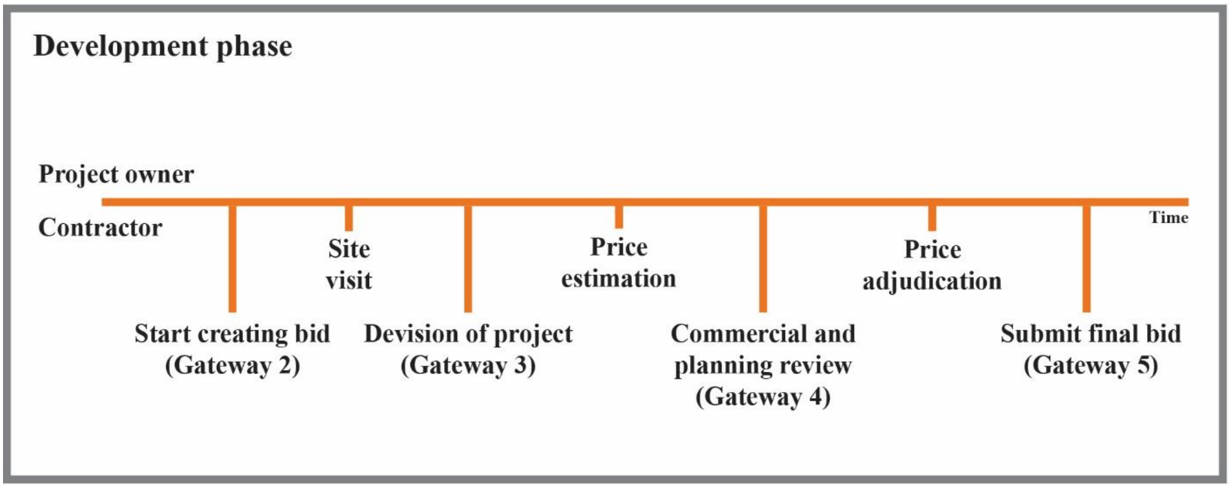


Figure A4. *Illustration of the development phase, including contractor activities below the timeline, and project owner activities above the timeline.*

During the formulation of the bid, contractors need to make many assumptions, as tender documents are often incomplete (Doloi, 2011; Seshadri, 2005). This is worrisome, as incorrect assumptions can turn to serious liabilities after winning the tender (Laryea & Hughes, 2008). To overcome some of this uncertainty, a site visit is often arranged to gain a more accurate impression of the project site before further bid development (Smith, 1995, pp. 22). Reasons for visiting the site may include general site details such as site access, security, the condition and present use of buildings, or the presence of local labor and subcontractors.

Laryea and Hughes (2008) conducted exploratory interviews with UK construction firms to investigate how contractors develop their bid in practice. They found that contractors process their bid through certain internal tendering gateways. The aim of these gateways is to select projects carefully, to avoid spending unnecessary costs and to assign risk in a systematic way. Laryea and Hughes (2008) reported five tendering gateways being used in UK construction contractors:

Gateway 1 (TG1) Board approval to actually pursue a job
Gateway 2 (TG2) Authorization to invest resources in preparing a tender
Gateway 3 (TG3) Division of project into packages for pricing by estimators and subcontractors
Gateway 4 (TG4) Commercial and planning review
Gateway 5 (TG5) Final settlement

Gateway 1 is related to the initiation phase, whereas gateway 2 to 5 are used in the development phase. The tendering gateways can be seen as go/no-go decision points. As a result, not all projects reach the final settlement. Numerous reasons can be found for a 'no go', for instance: the project owner's financial capability is unsatisfactory.

Besides the tender gateways, internal review meetings form a significant part of the contractors' tendering process (Laryea, 2013), and continue until the project is installed (Konijnendijk, 1994). According to Laryea (2013), review meetings are used to coordinate the bid activities, develop the strategy and check the quality of the bid. Laryea and Hughes (2008) further reported that UK project owners typically allow contractors to develop their bid within two to twelve weeks, depending upon the technicality of the project. Even though this time frame seems quite reasonable, in practice there is often not enough time (Smith, 1995, pp. 21).

One of the key activities during the bid development concerns the price setting, which is realized via two separate stages. Firstly, contractors estimate the costs for conducting the project, apply a percentage for overhead and include a mark-up: together this forms the sales estimate (Akintoye, 2000; Carr, 1989; Cassaigne et al., 1997). Second, the sales estimate is converted into a sales price through *adjudication*, in which management takes the prevailing market conditions into account and modifies price accordingly (Smith, 1995). A more extensive description of pricing and its role in tendering can be found in the next chapter.

In addition to pricing, time or lead-time is an important aspect of the bid, as it involves the due-date for the project owner, which is often a stated deadline in the tender document (Konijnendijk, 1994). To make sure lead-time objectives can be accomplished, Konijnendijk (1994) argues to include both sales representatives and engineers in the review meetings, to close the gap between marketing and engineering. This is vital, as high uncertainty and low controllability typify the marketing and manufacturing processes (Konijnendijk, 1994). Besides internal actors, the contractor keeps close contact with its main suppliers to make sure lead-times are appropriate (Hicks et al., 2000).

Last but not least, quality is the last focal aspect during the development of the bid. In project management, the success criteria 'price, time and quality' are often referred to as the 'iron triangle' (Atkinson, 1999;

Basu, 2013; Jha & Iyer, 2006). Where project quality can be defined as meeting project owner's expectations (Jha & Iyer, 2006). This seems like a simple definition, however, in practice it is hard to be compliant with all specifications *and* to document this appropriately beforehand in a contract (Basu, 2013). As a result, it appears that project managers focus more on price and time, while quality is sometimes perceived as 'ticking boxes' or overlooked to achieve the time and cost objectives (Basu, 2013; Jha & Iyer, 2006).

Although the iron triangle concept is widely adopted, there is still an ongoing debate whether the iron triangle is convenient in a project environment (Atkinson, 1999; Lauras et al., 2010). Consequently, over the last decade, several extra criteria have been proposed to be used in project management, however, for the purpose of this literature study – mainly focusing on the tender process – this is not further examined.

Once the project requirements, including the price, time and quality aspects are turned into a final bid, it is submitted and enters the evaluation phase.

8.4. Selective tendering - evaluation phase

In the last stage of the tender process – the evaluation phase –, the submitted bids are evaluated by the project owner. The criteria on which the bids are evaluated, and their relative weights, are known upfront, and usually compromises of: price, technical requirements, management capabilities, earlier experience, environmental and quality management systems, financial stability, collaborative skills, schedule and delivery aspects (Eriksson & Laan, 2007; Lam, Hu, Ng, Skitmore, & Cheung, 2001; Philbin, 2008). Of these criteria, the bid price seems to be a key parameter receiving the most weight, as project owners typically praise the lowest price above other parameters (Jennings & Holt, 1998; Kadefors, 2005; Mochtar & Arditi, 2001). An illustration of the activities within the evaluation phase can be found in Figure A5.

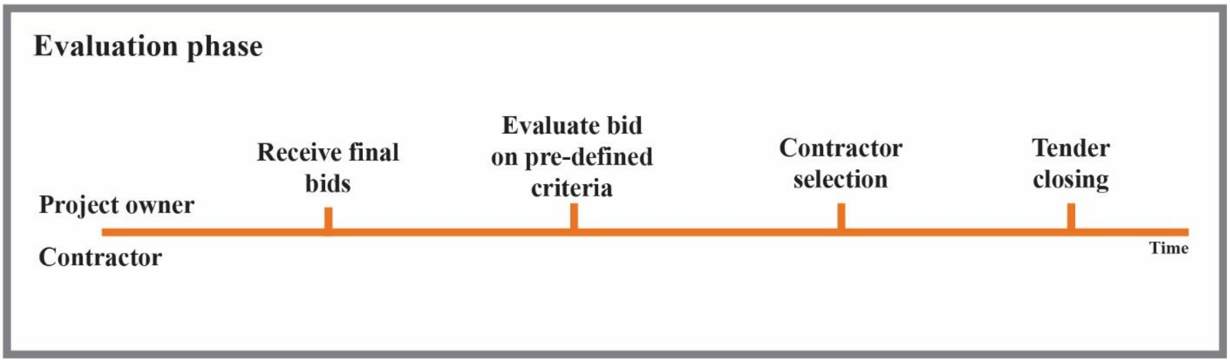


Figure A5. *Illustration of the evaluation phase, including contractor activities below the timeline, and project owner activities above the timeline.*

Recently a trend has been spotted to a more fine-tuned selection method in which the non-price parameters are also involved, especially among governmental tenders (Eriksson & Westerberg, 2011; Kumaraswamy

& Anvuur, 2008; Wong et al., 2000). By giving softer parameters relatively more weight compared to just low price, various benefits have been reported for the project owner, including reduced cost growth, fewer overruns and improved quality of the project (Eriksson & Westerberg, 2011). Unfortunately, in practice, it is still common to select contractors based on the lowest bid price (Kadefors, 2005).

Once a bid is chosen, the tender is 'closed' and it reaches the status of a project. In practice, changes to the project can still occur, however, the contractor has been selected and will not change anymore.

9. APPENDIX III - FACTORS INFLUENCING THE DECISION-TO-BID

Scholars have identified a large number of factors influencing the decision-to-bid, as can be found in the Table below. Identified factors are not merely grounded on contractors' internal incentives, but also market-based. However, as is shown in the table, research findings are inconclusive in regard to the relative importance of the identified factors. Nevertheless, overall, literature does suggest that the decision-to-bid is subjectively assessed (Smith, 1995).

	Odusote & Fellows (1992)	**Shash (1993)**	**Wanous (1998)**
Survey country	United Kingdom	United Kingdom	Syria
Number of identified factors	42 factors	55 factors	38 factors
Survey size	48 contractors	85 contractors	6 contractors
Top 10 factors	1. Client related factors	1. Need for work	1. Fulfilling the tender conditions
	2. Type of work	2. Number of competitors	2. Financial capability of client
	3. Project value	3. Previous project experience	3. Relation with and reputation of the client
	4. Contractor workload	4. Project type and size	4. Project size
	5. Estimated workload	5. Owner identity	5. Availability of time for tendering
	6. Likely profitability	6. Contact conditions	6. Availability of capital required
	7. Project location	7. Past profit on similar projects	7. Site clearance of obstructions
	8. Contract form	8. Tendering method	8. Public objections
	9. Physical resources to do the project	9. Risk involved owing to nature of work	9. Availability of materials required
	10. Identity of consultants	10. Availability of qualified staff	10. Current workload

10. APPENDIX IV – ESTIMATING THE PROJECT PRICE

The project price is estimated using a cost-based pricing approach, after which adjudication (Appendix V) will be applied, prior to submitting the price to the project owner. Typically, cost-based pricing starts by estimating the labor, material and equipment costs, after which a (fixed) contribution margin is added. This appendix will explain the steps in more detail.

10.1. Estimating the variable costs

Cost estimating can be defined as the technical process of assessing and predicting the cost of executing work in a given time, using available information and resources (CIOB, 1997; Kwakye, 1994). Early in the project's lifecycle, conceptual cost estimates are used to support decision making, cost scheduling and resource management (Carricano, 2014; Doloi, 2011; Dysert, 2008; Pickett, 2005). Literature suggests that cost estimating is not merely a precise technical process, but to a large extent a subjective and experience-based process, as estimators consider factors other than quantitative input (Ashworth & Skitmore, 1983; Elhag, Boussabaine, & Ballal, 2005; Martin Skitmore et al., 2006; A. Smith, 1995). An extreme example was uncovered by Whittaker (1970), he found several estimators to intentionally lower their cost estimate because they felt managements' mark-up adjustment -adjudication- was too high. Nevertheless, even though the estimating might be slightly subjective, Smith (1995) and Akintoye (2000) consider the cost estimate of prime importance, as it enables contractors with a bottom-line, i.e., below this value it would be unprofitable to execute the project.

The basic resources that need to be estimated are the labor, material and equipment costs, often termed variable costs or direct costs (Smith, 1995). Each of these will be briefly described.

Labor costs

In the construction and engineering labor is one of the key cost components in the estimate. Nonetheless, estimating labor costs with a high accuracy is difficult (Smith, 1995). In practice, contractors either use their own personnel or employ people via a subcontractor. For own personnel, the estimator uses 'rules of thumb' to calculate the labor costs. While this method may work to a large extent, the rules of thumb only give a single-point average. In practice, every project is unique in nature, as a result, the estimated costs might be far different from the actual costs. In subcontracted labor, the subcontractor gives an overall price for the work to be performed, therefore, when applying a fixed price contract with the subcontractor these costs are more straightforward (Smith, 1995).

Material costs

Material specifications are derived from the drawings or requirements from the project plan. Accordingly, key suppliers are informed to provide a price for the material. To obtain appropriate prices, often some form of competition is imposed by contractors' buying team (Smith, 1995, pp. 48).

Equipment and machinery costs
Contractors have a great variety of general equipment, however, specialized or large items are regularly hired (Smith, 1995, pp. 49). The decision to hire equipment or machinery might be affected by two prime motives. First, for contractors, it is uneconomic to own specialized equipment or machinery, as they generally cannot use the specialized equipment at full capacity over a long period. And second, the distance between the location of the project site and contractors' equipment might be considerable, hence due to transportation costs, it is more economic to hire local gear (Smith, 1995, pp. 49).

10.2. Contribution margin

After the estimation of the variable costs, a contribution margin is added. This contribution margin covers the overhead, risk/contingency, and profit.

Overhead
Overhead costs are fixed costs for the contractor, as they incur irrespective of whether the contractor executes work (Smith, 1995). Overhead may include the salaries of own personnel, office rental, and operating expenses. Typically, a percentage of the yearly overhead is included in the project price, where the percentage depends on the scope of the project.

Risk
Construction and engineering projects are, by their very nature, risky activities. Contractors mark these risks either as quantifiable risks or as unquantifiable risks. The latter called uncertainty and deals with the unknown. Importantly, when historical data is sufficiently available, the risk is mathematically assessable while uncertainty is not (Smith, 1995, pp. 156).

Weather, for instance, is a typical risk in construction projects, as bad weather can have a negative influence on the execution of a project (Chan & Au, 2007). Fortunately, weather statistics make it possible to incorporate appropriate costing for the weather risks in the estimated project price.

In recent years, academic researchers have proposed a rich array of analytical risk models that assess risk during the tender process (Samuel Laryea & Hughes, 2011), for example, by developing a fuzzy set model (Zeng, An, & Smith, 2007) or an artificial neural network model (Li, 1994). However, although some analytical models have received much attention in literature, empirical studies show that these models are hardly ever used in practice, and even if so, the calculated risk may be excluded from the final bid to enhance competitiveness in the tender (Samuel Laryea & Hughes, 2011).

On the other hand, uncertainty cannot be calculated, as information is unknown. For instance, in estimating the project costs, a contractor needs to make many assumptions, as tender documents are incomplete (Doloi, 2011; Seshadri, 2005). This unknown information might be made available or obtained, through research or contact with the project owner, and as a result, be labeled as a risk and calculated. However, in the absence of any data or previous occurrence of an event, a calculation is impossible and a subjective managerial assessment is necessary to incorporate the uncertainty in the estimating process (Smith, 1995, pp. 156).

Profit

To complete the estimated project price, a profit margin is added. As with the previous price elements, deciding the amount of profit margin to be added to a tender project is not a straightforward job, as this is directly linked with winning the tender, or not (Cassaigne et al., 1997; Paul & Gutierrez, 2005). The bandwidth for selecting a profit margin is minimal, as the construction industry is characterized by low-profit margins (Mochtar & Arditi, 2001).

Profit is either administered via ‘net’ estimating or ‘gross’ estimating. In net estimating, contractors do not include profit elements in the project price calculation. The profit is added during the adjudication stage (Smith, 1995, pp. 40). In gross estimating, the project price includes predefined percentages for profit on various project components, which will later be adjusted in the adjudication stage. Both methods are prevalent in use, and according to Smith (1995, pp. 136), the usage of either method is a matter of contractor preference.

11. APPENDIX V – ADJUDICATION

Once the estimated project price is established, the first stage of forming a final bid is completed. Subsequently, in the second stage, adjudication is applied, in which the estimated project price is converted into a final bid price (Smith, 1995, pp. 144).

11.1. Adjudication

The adjudication process starts by establishing the true commercial value of the project, which is the price the contractor would preferably charge for the project. Second, the price is adjusted to sell the project for the highest possible price while still have a chance of winning (Smith, 1995, pp. 144). Obviously, the final bid price should be stated above the costs to make a profit, but below perceived value to have a chance of winning (Hackett et al., 2007). Adjudication generally affects the contribution margin (risks and profit), and not the variable costs of executing the project (Laryea & Hughes, 2008). In practice, however, to make sure the project is captured, the adjustment can even result in a final bid price that is below the estimated costs of executing the project (Smith, 1995, pp. 145).

On the other hand, an upward adjustment – above the contractors' commercial value - is also possible (Smith, 1995, pp. 145). Consider, for instance, the altered example from Smith (1995, pp. 145):

> *Contractor A submitted a final bid price of $2.000.000, while contractor B submitted $ 2.100.000 and contractor C $2.200.000. Contractor A submitted the lowest bid and, all other bid elements being comparable to B and C, contractor A will win the tender. However, contractor A priced the project $100.000 below the next lowest contractor. Contractor A could have submitted up to $100.000 more and still be able to win the tender, capturing an additional profit of almost $100.000.*

Both the downward as well as the upward adjustment demonstrate that project prices become unrelated to the actual project costs, contrary to the 'pure' cost-based pricing strategy, an illustration can be found in Figure A6.

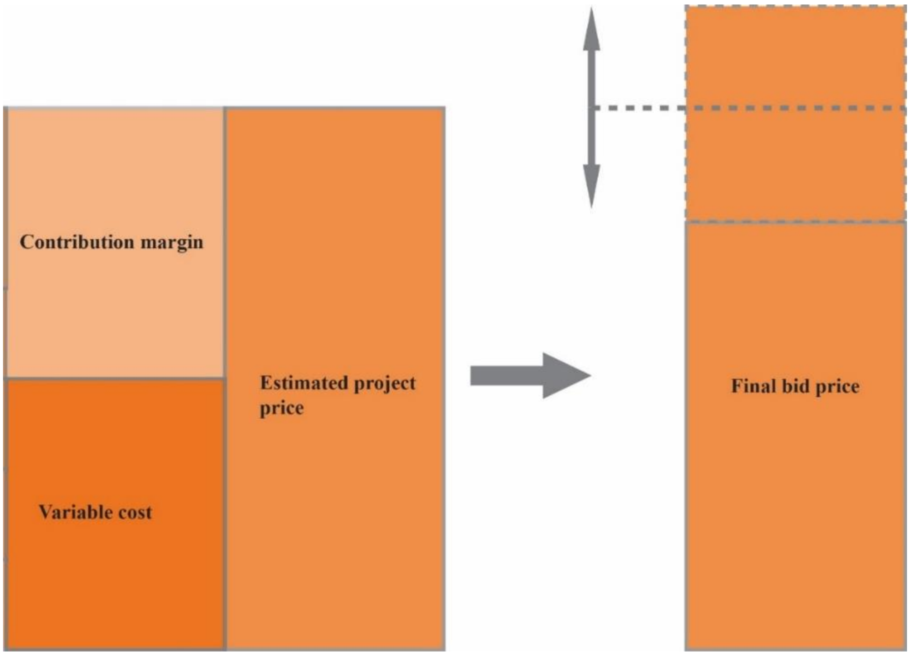


Figure A6. *Realization of final bid price by means of adjudication. The estimated project price is either decreased, increased of kept unchanged.*

To investigate the adjudication process, Mochtar and Arditi (2001) conducted a survey of 400 US contractors and found that competition and intuition seem to be key factors influencing the final bid price. They reported that more than 60% of the surveyed contractors assessed their competition. Where competitor characteristics may contain (1) the number of competitors in the tender, (2) their bidding history, (3) their commercial situation, (4) their present capacity and (5) their expansion strategy (Mochtar & Arditi, 2001). In addition, more than 50% of the participating contractors stated they used intuition in deciding upon the final bid price, which is in line with findings from Smith (1995).

After adjudication has been applied, the final bid price is ready to be submitted to the project owner. Note that the underlying budget submitted in the final bid is still a *prediction* at that particular time. The *actual* costs, and associated profits, are only clear after execution of the project.

Taken together, creating a final bid price is not a straightforward job, as is shown by the various steps a contractor has to go through. Where the first stage of estimating can be conducted in a relatively static and sheltered environment, the second stage of adjudication is more dynamic. Consequently, due to the adjudication process, the tender price is not established using a 'pure' cost-based pricing approach, but rather market-oriented.

12. APPENDIX VI – LOCATION OF THE DATA WITHIN VANDERLANDE

In Figure A7, an abstract life cycle of Vanderlande's BHS projects is shown, including the most relevant IT systems during the input, process and output phase.

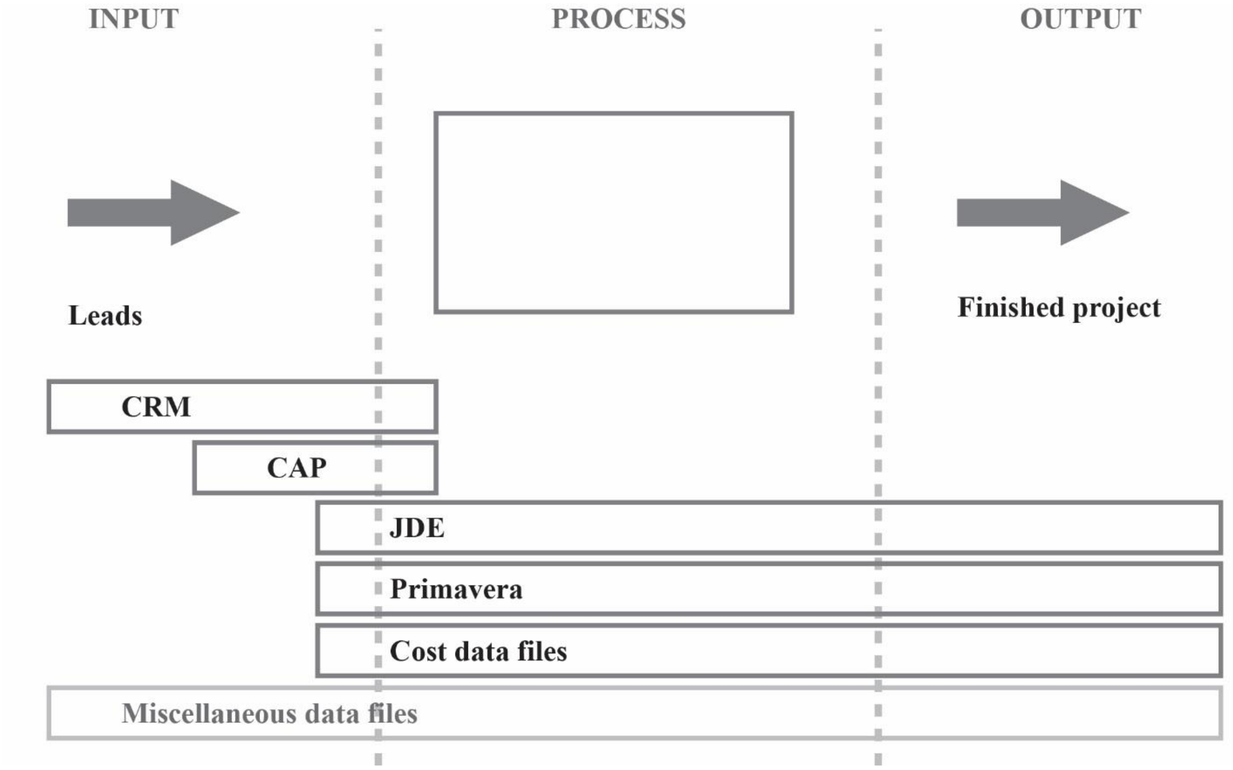


Figure A7. *Abstract life cycle of BHS projects within Vanderlande, including the relevant IT systems.*

- *CRM – Customer Relationship Management system, stores the leads, the project types, sales status, sales value and sales activities including timestamps.*
- *CAP – Computer Aided Pricing system, a software tool that calculates and stores the project price.*
- *JDE – JD Edwards, is an enterprise resource planning (ERP) system, storing various operating and planning activities.*
- *Primavera – an enterprise portfolio management system, used for project management, control capabilities and integrates with the enterprise resource planning system*
- *Cost data files, financial data files containing costs, e.g., the quotation costs made during the development of the bid.*
- *Miscellaneous data files, reports and loose files from various departments.*

13. APPENDIX VII – IDENTIFIED PREDICTOR VARIABLES

Various predictor variables have been mentioned by employees of Vanderlande, however, we were unable to incorporate all the mentioned variables in the analysis. Either because the data was not stored, not stored for a decent period, or stored inappropriately. Therefore, this Appendix will provide an overview of the mentioned variables. In addition, recommendations will be provided regarding the storage of the data.

- *New products*

  During the sales phase of a BHS project, some parts of the –to be built- system might be considered new. However, projects with relatively many 'new products' relate to more complexity as they have never been executed in practice before, which could turn out in extra man-hours and less profit.
  Recommendation: Store the number of new products by means of a percentage of the projects' total sales value in the ERP system. For instance project X has '1.5% new products'.

- *Risk*

  The amount of risk (in Euros) for a project, is implemented within the contribution margin. Since we only had access to the contribution margins, we were unable to acquire the real risk values as these were baked in the contribution margin values. It would have been of great use to have separate columns with the 'as sold' risk values, and the 'latest estimate' risk values.
  Recommendation: Store the amount of Risk in Euros per project in a separate field in the ERP system. For instance, project X has an 'as sold risk budget of € 200.000'.

- *Greenfield / brownfield*

  Greenfield projects are not constraint by prior work or conditions of existing systems. In contrast, brownfield projects relate to projects which are constrained by existing buildings or systems. Typically, in a brownfield setting, the old system needs to be adjusted, demolished or renovated, prior to implementing a new system. Therefore, brownfield projects are more complex and usually relate to extra man-hours, compared to greenfield projects. At the moment, the variable greenfield/brownfield is logged in Vanderlande's IT systems by means of a percentage, e.g. project X is 30% brownfield. Unfortunately, we were unable to incorporate this variable in the analysis, as it was only stored for the last few years. Additionally, based on interviews with several operational departments, it would be of great advantage to know which features relate to the brownfield percentage, for instance 'Project X, brownfield 70%, small building with existing system', or 'Project Y, brownfield 30%, many structural columns'.
  Recommendation: Store the percentage of greenfield/brownfield in the ERP system, *and* add some (special) features of the brownfield.

- *Currency rates*

  Within the ERP system, the sales values of the projects are stored including the currency type, for instance 'USD 7.654.321'. Unfortunately, no currency rates were added, which made it hard to transform the various currencies to EUR for the analysis. Note that the projects were conducted over a time period of seven years, and currency rates can fluctuate heavily over time. One way of dealing with this currency rates would be to find the original contract date and apply a specific currency rate for every project. Unfortunately, we were unable to gather and process the original contract within the timeframe of this study. Therefore, we applied a single currency rate, per currency type, to transform the sales values of various currencies to EUR.

  Recommendation: In addition to storing the currency type of a particular project it would be of great use to also store the currency rate (to EUR) in the ERP system at the time of the contract.

- *Screening machines*

  To screen the baggage, X-rays are used in the systems. X-rays are often in the scope of the project and are expensive products, however, based on budget, they consume relatively few man-hours during installation. Therefore, to take X-rays into account, it would be of great value to have separate columns with the number of x-rays and the related budget in the ERP system.

  Recommendation: Store the number and budget of X-rays per project in the ERP system.

- *Tender methods*

  In practice, different tender methods are used in the sales phase, for instance 'open', 'selective' and 'negotiation methods', as can be found in Appendix I. Unfortunately, the systems did not store any of this information, making it hard to incorporate this variable in the analysis.

  Recommendation: Store the tender method in the CRM system, for instance 'Project X - Open tender method'.

- *Consultants*

  Many project owners make use of a consultant. A consultant can, for instance, assist the project owner with the design and requirements of a system. However, the impact of a consultant on the project can reach much further. For instance, some consultants demand a very detailed documentation or have a specific process to be followed by the contractor which can be very different from the standard process of the contractor.

  Fortunately, since a few years, the systems store whether a consultant was present or not. Unfortunately, additional information regarding the consultant is missing, for instance, whether Vanderlande has worked with the consultant in the past, and whether the collaboration was positive.

  Recommendation: Store whether a project owner makes use of a consultant in the CRM system, and if so, how this particular consultant is expected to impact the project (positively and negatively).

14. APPENDIX VIII - GENERAL DATA RECOMMENDATIONS

➢ *Timestamps and events for process mining*

In addition to making predictive models using multiple linear regression methods, it would be of great advantage to rather insights regarding the process of the BHS projects, during tendering and execution. For this purpose, process mining has emerged, which enables companies to discover process knowledge from event logs (Van Der Aalst et al., 2005). Contrary to data mining, which often only analyses a specific topic in the process, process mining is focused on end-to-end processes, and thus gives a valuable overview of the business process (van der Aalst & Weijters, 2012).

Process mining can be used on CRM systems, as they typically store events including timestamps, for instance: sales activities and meetings. Process mining uses this data to discover, check conformance, and improve real processes (van der Aalst & Weijters, 2012). Process mining seems a promising method for dealing with real problems using event logs and has been applied in a wide variety of organizations, such as municipalities, government agencies, insurance companies, banks, hospitals, multinationals and high-tech system manufacturers (van der Aalst, 2016). Process mining provides fact-based insights to support process improvements, by extracting valuable, process-related information from event logs (van der Aalst, 2016). To be suitable for process mining, the event logs should contain cases, events, and timestamps. When these are available, three types of process mining can be applied on event logs:

- *Discovery* techniques produce process models by means of Play-In, as can be found in Figure A8. Play-In uses event logs with the goal to construct a model (van der Aalst, 2016).

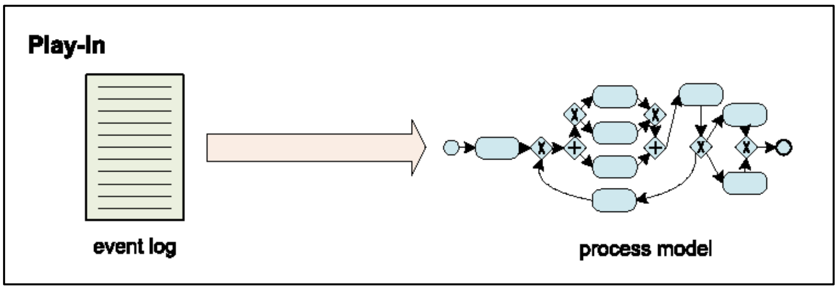


Figure A8. *Discovery of the process model by means of a Play-In. Figure adapted from Van der Aalst (2016).*

- *Conformance* compares an existing process model with a real life event log of the same process, by means of a Play-Out, as can be found in Figure A9. By applying conformance techniques, one can check whether the model captures reality and the other way around (van der Aalst, 2016).

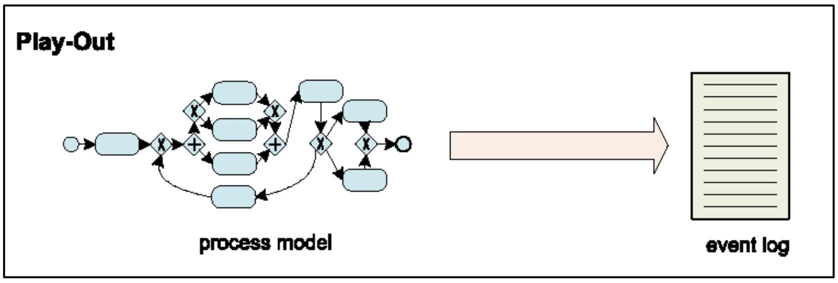


Figure A9. *Conformance checking by means of a Play-Out. Figure adapted from Van der Aalst (2016).*

- *Enhancement* involves the improvement of existing process models by Replaying actual event logs on a process model, as can be found in Figure A10. By doing so, insights can be gathered to repair, or extend the process model (van der Aalst, 2016)

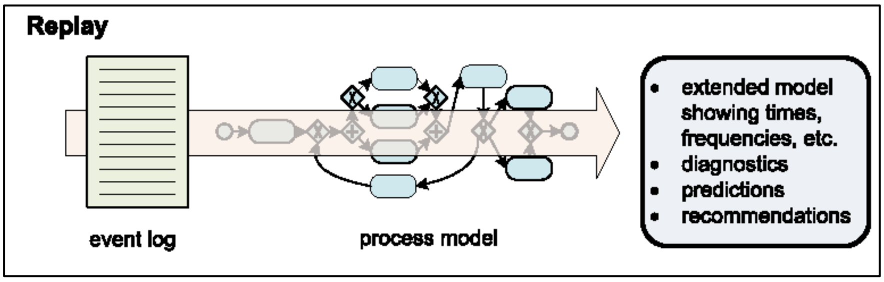


Figure A10. *Enhancement by means of a Replay. Figure adapted from Van der Aalst (2016).*

Based on a prior investigation into the Vanderlande event logs, concerning the tender process and execution, it was found that the IT systems did not log all events, although the steps are mapped on the process map. Furthermore, timestamps are often not logged or logged inappropriately. This inconsistency was found frequently. Therefore, the resulting event logs are too incomplete to be used for process mining at the moment.

Recommendation: To be able to apply process mining tools in the near future, it is important to store cases, events, and timestamps correctly. Fortunately, Vanderlande has recently updated its process by means of a new process map. This process map could be used as a starting point to select the relevant events, to be included in future analyses. In addition, it is of key importance that the timestamps are logged correctly. A suitable location for the timestamps of the events might be the planning system (Primavera) or the ERP system (JDE).

- *Quotation hours / costs*

During the development of a bid, various departments support the sales team. Every employee who works for a certain bid, stores their man-hours in the enterprise resource planning (ERP) system. Logging of this data is done per project number. Hence we were able to obtain the man-hours and costs per project number.

However, we were unable to see which (functional) roles worked on the bid, i.e., we were only able to capture the total man-hours and costs, and not which roles were involved. For instance, for project X we found 500 man-hours during the development of a bid, however, we were not able to find whether this were 400 man-hours for sales, 50 man-hours for legal and 50 man-hours for the operational department.
Recommendation: Next to logging the man-hours, store the functional role in the ERP system.

- *Information of lost projects*

  We were able to find a great deal of data regarding the won projects, however, for ‘lost’ projects this was a different story. (*Where ‘lost’ refers to projects which have spent sales resources but have not been captured*). Typically, when we were able to find a ‘lost’ project, only the project number was found with the invested amount of quotation costs, and some categorical characteristics such as the project type or customer center. However, the (potential) sales value was often not logged, making it hard to gather insights regarding these lost projects. In addition, since we had insufficient data regarding the ‘lost’ projects, we were unable to calculate the hit rates during tendering for BHS projects.
  Recommendation: Store the (potential) sales value of ‘lost’ projects in the CRM system. In other words: log what has been done for the spent quotation costs.

- *Linking project numbers*

  Once the main contractor has been selected by the project owner, the project execution formally begins. However, in practice, after the tender phase, many (small) variations can still be sold to the project owner. For instance, an Airport buys a BHS system for €50 million, after which the contractor sells additional variations, providing an additional sale of €7 million. Within Vanderlande, these variations have separate project numbers and can sometimes be seen as separate projects. In addition, for some large airports, multiple follow up projects might be executed. By doing so, some (large) airports can have over 20 separate project numbers. Nevertheless, these project numbers all belong to the same airport. Unfortunately, by making use of the systems, it is hard to obtain the correct list of project numbers belonging to a single airport. For instance, when we try to obtain this list of project numbers by means of the airport name, we find several abbreviations per airport. In addition, different names are used for the same project number in different systems. Together, this makes it really hard to obtain a definite list of project numbers belonging to a single airport.
  Recommendation: Store the project number of the main project number, e.g. XXX. Then, add the variations to the main project number, e.g., XXX-1. By applying this method, one can easily link all the variation projects to the main project. Furthermore, one can easily calculate the total sales value of a single customer and gather insights for future decision making.

- *Update system data*

Projects within Vanderlande go through several stages, starting in the lead phase and ending with a takeover. During these stages, project data is stored in different systems, as can be found in Appendix VI. Unfortunately, for the same project numbers, we frequently found different data points. For instance for the variables ‘project type’ and ‘sales value’:

- Project type: Project number X was labeled as project type ‘A’ in the CRM system, while in the ERP system the same project number was labeled as project type ‘B’. This was not an exception: in the raw data set 1430 of the 1504 project numbers (95%) were labeled inconsistently, i.e. different labels were attached in the CRM and ERP system to the same project number.
- Sales value: Project X was found to have a sales value of €7 million Euros in the CRM system, while the same project number in the ERP was found to have a sales value of €10 million Euros. Again, this was not an exception: in the raw data set 1247 of the 1504 project numbers (83%) had different sales values or the cell values were empty. Obviously, the fact that there are differences between the systems is not odd, as during the sales phase the sales value is logged in the CRM system and can still change considerably. Though, it would be useful if the sales value, after the signed contract, would be the equal in the different systems.

Recommendation: Cross-validate the various systems with each other, and update the errors /differences where necessary.

15. APPENDIX IX – VARIABLE SELECTION METHODS

Variable selection methods aim to select the best subset of predictor variables. Including too many predictor variables is called over-fitting, while incorporating too few predictor variables is called under-fitting (Sheather, 2009). There are two approaches for selecting the best predictor variables, namely the 'all possible subsets' method and the 'stepwise selection' method. Both methods are clarified in more detail in this Appendix.

15.1.1. All possible subsets method

This approach considers all 2^n possible subsets of predictor variables, and evaluates the subsets based on:

1. R^2_{adj}. Where R^2 is defined as the proportion of the total sample variability explained by the regression model (Sheather, 2009). While the R^2_{adj} ($R^2 - adjusted$) takes the number of predictor variables into account, i.e., adding inappropriate predictor variables does lead to a higher R^2, but not to to the same increase of R^2_{adj}.
2. *Akaike's Information Criterion (AIC).* This criterion balances the goodness of fit with a penalty for model complexity. Smaller AIC values are related to better models (Sheather, 2009).
3. *Corrected Akaike's Information Criterion (AIC_{Cor}).* The AIC tends to over-fit for small sample sizes. Therefore, Hurvich & Tsai (1989) developed a corrected AIC version, such that AIC can be used for smaller sample sizes.
4. *Bayesian Information Criterion (BIC).* Similar to AIC, though the BIC function penalizes model complexity more fierce, favoring models which are simpler over complex ones.

Scholars argue that, for model selection purposes, there is no clear preference whether to use R^2_{adj}, AIC, AIC_{Cor} or BIC (Sheather, 2009). Instead, they compute the R^2_{adj}, AIC, AIC_{Cor} or BIC of the models, and then compare the models with the minimal AIC, AIC_{Cor} and BIC values to the models with the greatest R^2_{adj} , and make a decision accordingly.

15.1.2. Stepwise selection

The stepwise selection methods are based on inspecting only a subset of the 2^n possible subsets, either by backward elimination, forward selection or a combination of both approaches. As the stepwise approaches do not examine all possible subsets (2^{n}), they are a computationally efficient alternative to the all possible subset method (James et al., 2013).

1. *Backward elimination* starts by placing all predictor variables in the model, from which it starts eliminating predictor variables, until the information criteria (AIC, AIC_{Cor} or BIC) values are minimized.

2. *Forward selection* begins with an empty model (Y ~ 0) from which it will start adding predictor variables per step, such that the information criteria decreases (Sheather, 2009). The process stops once the information criteria is minimized, or all predictor variables have been added.
3. *Hybrid selection* is a mixture of forward and backward selection. It starts with an empty model (Y ~ 0) from which it will start adding predictor variables such that the information criteria decreases. However, it also incorporates backward steps, such that all predictor variables in the model have sufficiently low p-values, while the variables outside the model have large p-values once they are added to the model (James et al., 2013).

The forward approach might include variables early in the process that later become redundant, while the backward selection is inappropriate for models with relatively many predictor variables (James et al., 2013). The hybrid selection method overcomes both of these drawbacks. In addition, the hybrid stepwise method approach is closely related to the all possible subset method, while having the computational efficiency of forward and backward stepwise selection (James et al., 2013). For that reason, we will incorporate the hybrid selection method for the predictive models.

16. APPENDIX X – CORRELATION MATRIX

In the three tables below, the correlation matrix (Pearson method) is shown. We used 'pairwise' complete observations to calculate the correlation coefficients. Some coefficients are based on over 200 observations, while others might only have 5 or 0 observations.

	Motors	Elec. Connections	Length	Quotation Costs	CRM Sales Value	Actual Sales	Actual Costs	Actual CM
Motors	1	.966**	.926**	.816**	.763**	.794**	.803**	.746**
Elec. Connections	.966**	1	.837**	.827**	.668**	.640**	.651**	.593**
Length of conveyor belts	.926**	.837**	1	.738**	.779**	.874**	.879**	.837**
Quotation Costs	.816**	.827**	.738**	1	.651**	.554**	.571**	.498**
CRM Sales Value	.763**	.668**	.779**	.651**	1	.830**	.826**	.815**
Actual Sales Value	.794**	.640**	.874**	.554**	.830**	1	.996**	.979**
Actual Costs	.803**	.651**	.879**	.571**	.826**	.996**	1	.958**
Actual CM	.746**	.593**	.837**	.498**	.815**	.979**	.958**	1
Actual CM percentage	-.230**	-.220**	-.185**	-.289**	-.079	-.102	-.149*	.013
delta sales	.428**	.278**	.549**	.164*	.191**	.706**	.704**	.689**
PM hours	.731**	.620**	.741**	.665**	.767**	.806**	.836**	.716**
PLE hours	.790**	.631**	.833**	.556**	.896**	.960**	.967**	.922**
IM hours	.672**	.627*	.546*	.509	.532	.644*	.709**	.410
PD hours	.b	.b	.b	.b	.b	.b	.b	.b
ME hours	.871**	.762**	.858**	.671**	.814**	.892**	.898**	.854**
SM hours	.581**	.482**	.547**	.495**	.582**	.599**	.630**	.504**
LLC hours	.716**	.758**	.636**	.727**	.478**	.387**	.434**	.265*
HLC hours	.364**	.342*	.260	.397**	.302*	.297*	.423**	.003
PTD Sales	.794**	.640**	.875**	.557**	.830**	1.000**	.996**	.979**
PTD Costs	.784**	.628**	.871**	.547**	.825**	.999**	.996**	.978**
PTD CM	.815**	.670**	.880**	.582**	.838**	.993**	.989**	.974**
PTD CM percentage	-.207**	-.182**	-.217**	-.137	-.091	-.156*	-.166*	-.129
LE CM	.747**	.595**	.837**	.502**	.815**	.979**	.957**	.999**
LE CM percentage	-.221**	-.211**	-.176*	-.280**	-.091	-.097	-.142*	.016

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).
b. Zero observations

Abbreviations:
CRM = Customer Relationship Management, CM = Contribution Margin, PM = Project Management, PLE = Project Leader Engineering, IM=Integration Management, PD = Project Director, ME= Mechanical Engineering, SM = Site Management, LLC = Low Level Controls, HLC = High Level Controls, PTD = Project-To-Date, LE = Latest Estimate

	Actual CM percentage	delta sales	PM hours	PLE hours	IM hours	PD hours	ME hours	SM hours
Motors	-.230**	.428**	.731**	.790**	.672**	.b	.871**	.581**
Elec. Connections	-.220**	.278**	.620**	.631**	.627*	.b	.762**	.482**
Length of conveyor belts	-.185**	.549**	.741**	.833**	.546*	.b	.858**	.547**
Quotation Costs	-.289**	.164*	.665**	.556**	.509	.b	.671**	.495**
CRM Sales Value	-.079	.191**	.767**	.896**	.532	.b	.814**	.582**
Actual Sales Value	-.102	.706**	.806**	.960**	.644*	.b	.892**	.599**
Actual Costs	-.149*	.704**	.836**	.967**	.709**	.b	.898**	.630**
Actual CM	.013	.689**	.716**	.922**	.410	.b	.854**	.504**
Actual CM percentage	1	-.079	-.329**	-.110	-.470	.b	-.099	-.505**
delta sales	-.079	1	.605**	.775**	.267	.b	.541**	.478**
PM hours	-.329**	.605**	1	.886**	.490	.b	.787**	.770**
PLE hours	-.110	.775**	.886**	1	.861**	.b	.889**	.918**
IM hours	-.470	.267	.490	.861**	1	.b	.844**	.493
PD hours	.b	.b	.b	.b	.b	.b	.b	.b
ME hours	-.099	.541**	.787**	.889**	.844**	.b	1	.549**
SM hours	-.505**	.478**	.770**	.918**	.493	.b	.549**	1
LLC hours	-.432**	.161	.610**	.440**	.665*	.b	.515**	.327*
HLC hours	-.422**	.102	.713**	.422*	.995**	.b	.427**	.160
PTD Sales	-.102	.705**	.806**	.960**	.643*	.b	.892**	.599**
PTD Costs	-.103	.710**	.802**	.955**	.616*	.b	.886**	.601**
PTD CM	-.099	.683**	.811**	.967**	.703**	.b	.902**	.591**
PTD CM percentage	.586**	-.159*	-.130	-.101	-.185	.b	-.106	-.044
LE CM	.013	.687**	.714**	.922**	.403	.b	.854**	.502**
LE CM percentage	.935**	-.054	-.337**	-.111	-.473	.b	-.094	-.509**

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

b. No observations

Abbreviations:
CRM = Customer Relationship Management, CM = Contribution Margin, PM = Project Management, PLE = Project Leader Engineering, IM=Integration Management, PD = Project Director, ME= Mechanical Engineering, SM = Site Management, LLC = Low Level Controls, HLC = High Level Controls, PTD = Project-To-Date, LE = Latest Estimate

	LLC hours	HLC hours	PTD Sales	PTD Costs	PTD CM	PTD CM percentage	LE CM	LE CM percentage
Motors	.716**	.364**	.794**	.784**	.815**	-.207**	.747**	-.221**
Elec. Connections	.758**	.342*	.640**	.628**	.670**	-.182**	.595**	-.211**
Length of conveyor belts	.636**	.260	.875**	.871**	.880**	-.217**	.837**	-.176*
Quotation Costs	.727**	.397**	.557**	.547**	.582**	-.137	.502**	-.280**
CRM Sales Value	.478**	.302*	.830**	.825**	.838**	-.091	.815**	-.091
Actual Sales Value	.387**	.297*	1.000**	.999**	.993**	-.156*	.979**	-.097
Actual Costs	.434**	.423**	.996**	.996**	.989**	-.166*	.957**	-.142*
Actual CM	.265*	.003	.979**	.978**	.974**	-.129	.999**	.016
Actual CM percentage	-.432**	-.422**	-.102	-.103	-.099	.586**	.013	.935**
delta sales	.161	.102	.705**	.710**	.683**	-.159*	.687**	-.054
PM hours	.610**	.713**	.806**	.802**	.811**	-.130	.714**	-.337**
PLE hours	.440**	.422*	.960**	.955**	.967**	-.101	.922**	-.111
IM hours	.665*	.995**	.643*	.616*	.703**	-.185	.403	-.473
PD hours	.b	.b	.b	.b	.b	.b	.b	.b
ME hours	.515**	.427**	.892**	.886**	.902**	-.106	.854**	-.094
SM hours	.327*	.160	.599**	.601**	.591**	-.044	.502**	-.509**
LLC hours	1	.747**	.386**	.368**	.437**	-.076	.266*	-.431**
HLC hours	.747**	1	.296*	.276*	.345*	.000	-.002	-.422**
PTD Sales	.386**	.296*	1	.999**	.994**	-.155*	.979**	-.096
PTD Costs	.368**	.276*	.999**	1	.989**	-.165*	.978**	-.098
PTD CM	.437**	.345*	.994**	.989**	1	-.122	.974**	-.092
PTD CM percentage	-.076	.000	-.155*	-.165*	-.122	1	-.127	.588**
LE CM	.266*	-.002	.979**	.978**	.974**	-.127	1	.017
LE CM percentage	-.431**	-.422**	-.096	-.098	-.092	.588**	.017	1

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

b. No observations

Abbreviations:
CRM = Customer Relationship Management, CM = Contribution Margin, PM = Project Management, PLE = Project Leader Engineering, IM=Integration Management, PD = Project Director, ME= Mechanical Engineering, SM = Site Management, LLC = Low Level Controls, HLC = High Level Controls, PTD = Project-To-Date, LE = Latest Estimate

17. APPENDIX XI – MLR MODELS MAN-HOURS

Similar to the MLR models for the Mechanical Engineering hours in the main body, the following tables provide the MLR models for the other roles. Unfortunately, we were not able to obtain valid MLR models for Integration Management (7 obs.), High Level Control (14 obs.) and Project Director Hours (0 obs.).

Summary of MLR Analysis for variables predicting log ***Project Management Hours*** *(N=80). Both the full models, as the (hybrid) stepwise regression models are shown.*

	Full MLR models			MLR models after hybrid stepwise variable selection (per model)		
Model 1	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	0.78		B_0	0.78	
log Sales Value CRM	B_1	0.06	<0.000	B_1	0.06	<0.000
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.728	0.724		0.728	0.724	50%
Model 2	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	1.14		B_0	1.06	
log Sales Value CRM	B_1	0.12	<0.000	B_1	0.11	<0.000
log Nr of Motors	B_2	0.12	0.917	-	-	-
log Length of Belts	B_3	0.14	0.034	B_3	0.11	0.004
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.756	0.746		0.756	0.749	49%
Model 3	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	1.22		B_0	1.06	
log Sales Value CRM	B_1	0.12	<0.000	B_1	0.11	<0.000
log Nr of Motors	B_2	0.37	0.434	-	-	-
log Length of Belts	B_3	0.15	0.084	B_3	0.11	0.004
log Nr of Elec. Connections	B_4	0.31	0.427	-	-	-
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.720	0.705		0.756	0.749	49%
Model 4	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	1.18		B_0	1.01	
log Sales Value CRM	B_1	0.11	<0.000	B_1	0.11	<0.000
log Nr of Motors	B_2	0.35	0.634	-	-	-
log Length of Belts	B_3	0.15	0.075	B_3	0.10	0.016
log Nr of Elec. Connections	B_4	0.29	0.503	-	-	-
log Delta Sales	B_5	0.01	0.002	B_5	0.01	0.002
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.787	0.772		0.785	0.776	49%

Notes:

- *Delta Sales was calculated by subtracting the 'Sales Value CRM' from the 'Signed Contract Value'.*
- *The MAPE has been calculated by splitting the dataset (80 obs.) at random into a trainset (64 obs.) and a testset (16 obs.). The data from the trainset was used to estimate the regression coefficients using the stepwise regression method. Then data from the testset was used to predict the response variable, followed by calculating the MAPE. This process was repeated ten times and averaged.*

Summary of MLR Analysis for variables predicting log ***Project Leader Engineering Hours*** *(N=44). Both the full models, as the (hybrid) stepwise regression models are shown.*

	Full MLR models			MLR models after hybrid stepwise variable selection (per model)		
Model 1	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	1.76		B_0	1.76	
log Sales Value CRM	B_1	0.12	<0.000	B_1	0.12	<0.000
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.653	0.644		0.653	0.644	101%
Model 2	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	2.19		B_0	2.15	
log Sales Value CRM	B_1	0.20	<0.000	B_1	0.19	<0.000
log Nr of Motors	B_2	0.26	0.142	B_2	0.17	0.037
log Length of Belts	B_3	0.27	0.875	-	-	-
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.688	0.665		0.688	0.673	80%
Model 3	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	2.42		B_0	2.15	
log Sales Value CRM	B_1	0.20	<0.000	B_1	0.19	<0.000
log Nr of Motors	B_2	0.67	0.488	B_2	0.17	0.037
log Length of Belts	B_3	0.29	0.855	-	-	-
log Nr of Elec. Connections	B_4	0.56	0.899	-	-	-
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.688	0.656		0.688	0.673	80%
Model 4	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	2.23		B_0	1.54	
log Sales Value CRM	B_1	0.19	<0.000	B_1	0.11	<0.000
log Nr of Motors	B_2	0.61	0.729	-	-	-
log Length of Belts	B_3	0.27	0.620	-	-	-
log Nr of Elec. Connections	B_4	0.50	0.783	-	-	-
log Delta Sales	B_5	0.02	0.004	B_5	0.01	<0.000
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.751	0.718		0.740	0.728	74%

Notes:

- *Delta Sales was calculated by subtracting the 'Sales Value CRM' from the 'Signed Contract Value'.*
- *The MAPE has been calculated by splitting the dataset (44 obs.) at random into a trainset (36 obs.) and a testset (8 obs.). The data from the trainset was used to estimate the regression coefficients using the stepwise regression method. Then data from the testset was used to predict the response variable, followed by calculating the MAPE. This process was repeated ten times and averaged.*

Summary of MLR Analysis for variables predicting log ***Site Management Hours*** *(N=41). Both the full models, as the (hybrid) stepwise regression models are shown.*

	Full MLR models			MLR models after hybrid stepwise variable selection (per model)		
Model 1	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	1.96		B_0	1.96	
log Sales Value CRM	B_1	0.14	<0.000	B_1	0.14	<0.000
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.509	0.499		0.509	0.499	136%
Model 2	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	2.73		B_0	0.85	
log Sales Value CRM	B_1	0.26	0.920	-	-	-
log Nr of Motors	B_2	0.25	0.152	B_2	0.23	0.136
log Length of Belts	B_3	0.30	0.079	B_3	0.27	0.051
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.646	0.617		0.646	0.627	90%
Model 3	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	2.75		B_0	1.04	
log Sales Value CRM	B_1	0.25	0.827	-	-	-
log Nr of Motors	B_2	0.70	0.009	B_2	0.55	<0.000
log Length of Belts	B_3	0.31	0.384	-	-	-
log Nr of Elec. Connections	B_4	0.60	0.022	B_4	0.55	0.005
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.695	0.661		0.683	0.667	73%
Model 4	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	2.82		B_0	1.04	
log Sales Value CRM	B_1	0.26	0.741	-	-	-
log Nr of Motors	B_2	0.71	0.009	B_2	0.55	<0.000
log Length of Belts	B_3	0.32	0.440	-	-	-
log Nr of Elec. Connections	B_4	0.60	0.022	B_4	0.55	0.005
log Delta Sales	B_5	0.01	0.624	-	-	-
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.697	0.653		0.683	0.667	73%

Notes:

- *Delta Sales was calculated by subtracting the 'Sales Value CRM' from the 'Signed Contract Value'.*
- *The MAPE has been calculated by splitting the dataset (41 obs.) at random into a trainset (33 obs.) and a testset (8 obs.). The data from the trainset was used to estimate the regression coefficients using the stepwise regression method. Then data from the testset was used to predict the response variable, followed by calculating the MAPE. This process was repeated ten times and averaged.*

Summary of MLR Analysis for variables predicting log **Low Level Controls Hours** *(N=61). Both the full models, as the (hybrid) stepwise regression models are shown.*

	Full MLR models			MLR models after hybrid stepwise variable selection (per model)		
Model 1	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	1.46		B_0	1.46	
log Sales Value CRM	B_1	0.10	<0.000	B_1	0.10	<0.000
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.582	0.575		0.582	0.575	102%
Model 2	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	2.47		B_0	0.43	
log Sales Value CRM	B_1	0.25	0.382	-	-	-
log Nr of Motors	B_2	0.36	0.001	B_2	0.09	<0.000
log Length of Belts	B_3	0.28	0143	-	-	-
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.667	0.649		0.652	0.646	81%
Model 3	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	2.52		B_0	0.43	
log Sales Value CRM	B_1	0.27	0.442	-	-	-
log Nr of Motors	B_2	0.76	0.129	B_2	0.09	<0.000
log Length of Belts	B_3	0.30	0.171	-	-	-
log Nr of Elec. Connections	B_4	0.68	0.892	-	-	-
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.668	0.644		0.652	0.646	81%
Model 4	Beta	Std. Error	p-value	Beta	Std. Error	p-value
Constant	B_0	2.49		B_0	0.43	
log Sales Value CRM	B_1	0.26	0.415	-	-	-
log Nr of Motors	B_2	0.76	0.219	B_2	0.10	<0.000
log Length of Belts	B_3	0.29	0.212	-	-	-
log Nr of Elec. Connections	B_4	0.67	0.714	-	-	-
log Delta Sales	B_5	0.01	0.098	B_5	0.01	0.100
	R^2	$R^2_{Adj.}$		R^2	$R^2_{Adj.}$	MAPE
	0.684	0.655		0.668	0.657	78%

Notes:

- *Delta Sales was calculated by subtracting the 'Sales Value CRM' from the 'Signed Contract Value'.*
- *The MAPE has been calculated by splitting the dataset (61 obs.) at random into a trainset (49 obs.) and a testset (12 obs.). The data from the trainset was used to estimate the regression coefficients using the stepwise regression method. Then data from the testset was used to predict the response variable, followed by calculating the MAPE. This process was repeated ten times and averaged.*