Analysis of costs in new terminals investments

MSc. Thesis
Coos van Buuren

Civil Engineering
Quantitative Finance

Strictly private and confidential

January 2011

Delft University of Technology

ERASMUS UNIVERSITEIT ROTTERDAM

APM TERMINALS

Table of contents

List of figures

Abstract

This thesis studies cost estimation and investment decisions under cost uncertainty of large construction projects. The combination of these two topics intends to satisfy the Master's theses of Civil Engineering at the Delft University of Technology and Quantitative Finance (Econometrics) at the Erasmus University Rotterdam. The first part relates to the Delft University of Technology, the second to the Erasmus University Rotterdam. Both parts are interrelated but can be read separately.

Part A develops a model for APM Terminals to estimate costs of new terminals investments. Present-day estimates of APM Terminals insufficiently incorporate risk. Therefore, the study formulates and analyses six (new) estimation models. Differences in the model include the use of error distribution function or the incorporation of interdependency. The analysis subjects the models to various constraints and selects the most appropriate model for APM Terminals. The selected model requires little input information, uses normally distributed error distributions and accounts for shocks. Moreover, the study points out shocks are of great importance in the estimation of costs. Shocks increase expected costs and mainly determine cost uncertainty. The change of estimation model and occurrence of shocks implies that APM Terminals changes its estimation process. The new approach requires an estimate of both expected cost and uncertainty to estimate construction costs.

Part B studies investments of projects subjected to cost uncertainty. Prior to construction investors have an idea of the value but not of the costs. Academic research assumes that cost uncertainty is composed of technical and input uncertainty (Pindyck (1993)). Technical uncertainty covers the physical difficulty to complete the project and is only known after completion. Input uncertainty relates to the pricing uncertainty of the required commodities to complete the project and is known beforehand. This study adds shocks to uncertainty because of the significant contribution to uncertainty (Part A). The research argues that shocks are a special form of technical uncertainty. Shocks occur after investing and complicate physical completion. But where technical uncertainty can accelerate construction, shocks solely delay progress. The study uses option theory to examine investments subjected to the different types of uncertainty. The analysis defines optimal investments and shows shocks increase the aversion to invest.

Part A – Delft Study

A probabilistic approach to estimate costs as part of new terminals investments

Preface

On the 17th of August 2009 Prof. Vrijling send me an email. He wrote:
"A former graduate who works within the Danish liner company AP Moller passed by. He looks for a model on investments in harbor infrastructure to account for uncertainty and risk. That might be of your interest. Please walk by and I will tell you more."
That is how we started off.

This thesis describes the journey to develop a realistic and reliable model that forecasts the costs of investments in harbor infrastructure. The model is designed to map the risk and uncertainty involved in the civil infrastructural related investments of the design phase of a new terminal.
The model is developed for APM Terminals. APM Terminals is part of the liner shipping company A.P. Moller-Maersk and designs, constructs and operates container terminals.
This section is part of the studies Civil Engineering at the Delft University of Technology.

During the thesis I was reviewed, supported and challenged by the graduation committee. The committee consists of:

Ir. Paul van Weert	Head of civil engineering - APM Terminals
Ing. Peter Biemond	Principal civil engineer - APM Terminals
Prof. drs. ir. Han Vrijling	Professor probabilistic design – Delft University of Technology
Dr. ir. Pieter van Gelder	Associate professor – Delft University of Technology
Dr. Michel van der Wel	Researcher econometric Institute – Erasmus University

I would like to thank the graduation committee for their support and feedback.

The information within this report is strictly private and confidential. Under no circumstances or conditions than the written approval of APM Terminals details of the report may be shared with anyone or any organization.

Introduction

This first part of the Master's thesis develops a cost estimation model for APM Terminals. Investors use costs estimates to judge profitability of investment opportunities. Costs estimates have a strong empirical character. There is not one approach that exactly forecasts the costs. The reason is the deviation between the various projects that changes cost drivers and cost development. But basic statistical principles underlie all estimates.

This part of the thesis aims to develop a flexible, accurate and user friendly tool to estimate costs of civil infrastructure as part of new container terminal investments. In particular, this thesis develops a model that presents the inherent risk of new investments. The thesis first analyses six different estimation tools and thereafter proposes a model for APM Terminals. The model helps APM Terminals within the process to review investment profitability and the inherent risk of the opportunity.

This study develops the estimation tool in eight chapters. The first chapter introduces APM Terminals, their investments and the research topic. It defines the research question and gives the chapters that provide the answers. The first few succeeding chapters provide a solid foundation to estimate costs and select the estimation model. Chapter 2 analyzes cost estimates and discusses statistics that underlie cost estimates. Chapter 3 develops the different estimation models. Chapter 4 describes the conditions the tool should meet. Knowing sufficient (background) information on cost estimation, the chapters thereafter estimate costs and evaluate the models. Chapter 5 discusses an investment opportunity to estimate. All the different models forecast the investment opportunity showing the differences and the similarities of the different estimates. Chapter 6 presents the estimation results and comes to the proposal of an estimation tool. Chapter 7 discusses the use of the model and chapter 8 concludes all findings and comes to any recommendations.

1 Introduction to APM Terminals, new terminals investments and the research topic

This first chapter gives an introduction to APM Terminals and describes the roadmap towards new container terminals. The information explains the importance of cost estimates. Insight in the current cost estimation process clarifies the demand within APM Terminals' to review its estimation model. This chapter distills a thesis question from the insights. Thereafter this section describes the research methodology as well as the outline of this document.

1.1 Introduction to APM Terminals

APM Terminals (APMT) is a global terminal operator with a network of 50 ports and 19,000 employees in 34 countries. The company provides port management and operations to over 60 liner shipping customers who serve the world's leading importers and exporters of containerized cargo.

The company, headquartered in The Hague, aims to meet constantly changing needs of the international trade community through high productivity operations and port capacity in economically, environmentally and socially responsible ways. APM Terminals, part of the AP Moller-Maersk Group, generated over USD 3 billion revenues in 2009 (USD 738mn EBITDA). To maintain growth APM Terminals constantly aims to secure new profitable investments.

APM Terminals can be split into two parts; Existing Terminals (ET) and New Terminals (NT). The Existing Terminals division operates the terminals within its portfolio. Its main objective is to realize the projected throughput of containers and deliver high service to its customers. High quality of service is provided by handling containers fast, safely and accurately at low costs.

The New Terminals division reviews, develops and invests in infrastructure and terminal opportunities to further optimize the terminal portfolio. Investment in port and inland transportation access remain essential to the continuous development of the global economy. The company globally pursues appropriate projects and partnerships, currently focusing on construction and expansion projects in Africa, Southern America and Asia.

1.2 Investing in new terminals

The financial performance of a terminal is determined by the difference between the revenues and the costs. The revenues of a container terminal are determined by the amount of handled containers and the rates per container move.

The costs of a container terminal are determined by the required investments to enable container handling (Capital Expenditures or CAPEX), the costs during operations (Operational Expenditures or OPEX) and the concession fees.

1.2.1 Project stages

The process to develop a new terminal is divided into several stages. APM Terminals defines this process as the Project Excellence Life Cycle. Five different stages are identified; from the initial exploring study to operating the terminal. These stages are shown in Figure 1.

Figure 1; Project Excellence life cycle

At the end of every stage several documents or operations are finalized to add a new terminal to the portfolio. In the early stages the project is progressed to the next stage based on the opportunity and feasibility. At later stages the progress is dependent on permitting and contracting documents or physical progress of construction works. The different stages classified within APM Terminals are;

- Probe; a first exploring phase to identify the business opportunity. The opportunity is reviewed roughly. The results of the feasibility study drive further analysis of the business opportunity.
- Develop; the opportunity is carefully analyzed to create an attractive business case. Documents are prepared to come to an internal approval. A bid is submitted at the governing authority of the terminal area.
- Secure; the bid documents are negotiated and legal documents are signed. An implementation team is selected and installed to prepare the terminal construction works.
- Implement; the project is constructed and the terminal is prepared for costumers signed up.
- Operate; operate the terminal and evaluate the process to come to the new terminal.

1.2.2 Evaluating the business case

To analyze and evaluate a new business opportunity, the costs and revenues of new investments are forecasted and combined into a cash flow model. As the value of money changes over time the cash flows are discounted.

The rate to discount the cash flows (i.e. discount rate) is the expected rate of return on investments. Academically the rate is called opportunity cost of capital. An investor may have multiple opportunities to invest in, but can only select one. The opportunity selection is based on the expected return versus the project risk. The expected return is balanced with the investor's risk appetite; the discount rate.

Within APM Terminals the discount rate is determined by the APMM group to align the project risk to the group risk profile. For more information on the discount rate and risk appetite, Ross (1) is recommended. A method to determine the discount rate is shown in Annex 2.

In Figure 2 the cash flows of an APMT example project are shown. The revenues and investment cash flows are shown, both in face and discounted value.

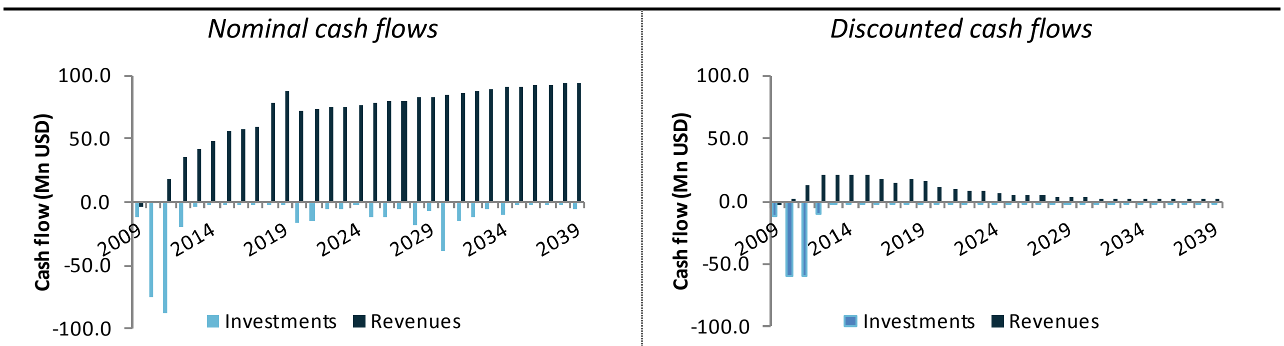


Figure 2; Nominal and discounted cash flows

The sum of the discounted cash flows is called the Net Present Value(1). A positive NPV indicates a profitable project aligned to the expected return. The example shows a NPV of 83.1mn based on an expected return of 17.5%.

In addition the Internal Rate of Return (IRR) is calculated. The IRR is the discount rate at which the NPV is zero. The IRR of the example is 25.8%. The difference between the IRR and the discount rate indicates the likelihood of the project to become unprofitable. This example project shows a gap of 8.3%. Experience has shown this gap is sufficiently large.

One may intuitively understand this idea. The expected return and IRR represent a certain risk profile. A high(er) risk premium is rewarded with a large(r) return. If one discounts the project with expected return, the project represents less risk instead of discounting the project with the IRR. One expects the project to contain a risk profile reflected within the expected return. Possibly the project risk profile could be represented with the IRR. A large difference between both rates reflects such a different risk profile unlikely to happen.

1.2.3 The civil CAPEX

The graph in Figure 2 clearly shows that the cash flows in the early stage of the model have a large impact on the financial performance of the project, as they are greatly negative. These early cash flows are mainly capital expenditures (CAPEX); investments are required first to generate profit

In Figure 3 the CAPEX of a project are shown in detail. The left bar chart shows the scheduled expenses for both the equipment and the civil CAPEX. The equipment investments contain the cranes and other equipment to operate the terminal; the civil CAPEX comprise the permanent terminal structures.

The pie chart shows the split of the civil and the equipment CAPEX in face value over the entire project. The bar chart at the right shows the discounted value of the CAPEX. A split of the discounted CAPEX cash flows would show an even larger influence of the civil CAPEX.

These figures clearly show that the CAPEX of civil related infrastructure account for 60% of the total investments. The figures also show that the majority of the 'civil CAPEX' are spent in the early stage of a project. From the above information it is likely to conclude civil CAPEX have a large impact on the financial performance of a terminal. The forecast of civil CAPEX is therefore one of the main drivers of the new terminal valuation.

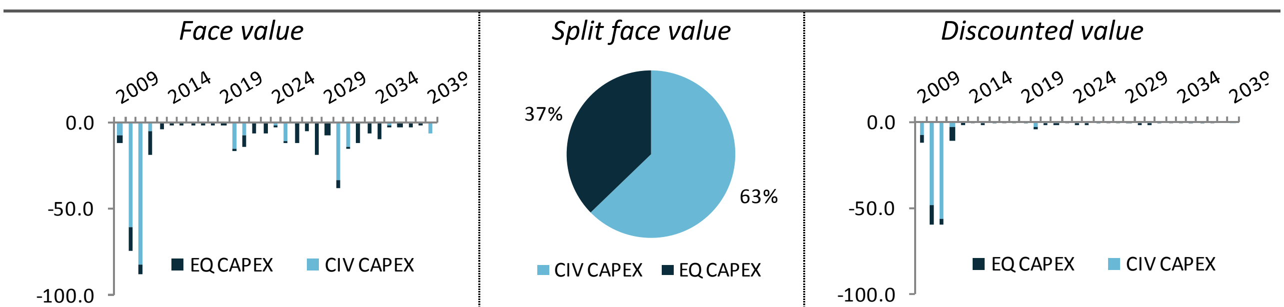


Figure 3; CAPEX of a project

1.2.4 Estimation of the civil CAPEX

The costs of the civil infrastructure are estimated by the Civil Engineering Department. Once a business case is selected the department is asked to provide an estimate. The estimate is a forecast of the final price to construct the terminal. An approximated present day value of the terminal construction costs is provided. A construction escalation rate is estimated to take increasing construction prices within the financial model into account. The estimate is often verified with the contractor bid price. This is an indicator of the final construction costs. Unspecified scope and exceptional events will ultimately result in a higher price.

The process to estimate the construction costs of the terminal is:

1. Identification of the terminal layout
2. Investigate the terminal components and process flow of containers within the terminal to determine structural design
3. Estimation of the dimensions (i.e. quantities) of the terminal parts
4. Estimation of the unit rates to determine the costs per element
5. Sum of the elements is the cost estimate
6. A safety factor, named contingency, is added to achieve a sufficiently low probability level of cost overrun

A rough outline is provided in the early stage to develop a terminal. This results in a rough estimate as well. As the level of detail on the design increases over time, the estimate also becomes more accurate over time. The rough estimates are provided by APM Terminals' employees. In a later stage an external consultant provides the estimate. A detailed estimation requires specialized knowledge and is time consuming. The estimate is generally based on market contacts, experience from previous projects and expert judgment to estimate price and quantity of the terminal elements.

In Table 1 the cost classifications and the evolving accuracy defined within APM Terminals are shown(2)

Class	Name	Typical Accuracy	Comments
Class 5	Ballpark Estimate	+/-40%	Rough estimate without design sketches except for draft terminal layout
Class 4	Feasibility Estimate	+/-20%	In-house produced estimate with limited support from external consultancies for sketches and/or drawings. Local cost benchmarks are available and a site visit is required.
Class 3	Budget Estimate	+/-10%	Extensive support from external consultants to prepare; 20-50 engineering drawings and related documents such as construction plan, detailed project schedule, contracting strategy etc.
Class 2	Control Estimate	+/-5%	This estimate is mainly used for monitoring and accounting purposes as the terminal is already under construction.
Class 1	Detailed Control Est.		

Table 1: Estimation classification APM Terminals

The cost estimates are mainly relevant in the development phase of a terminal; Class 3 to 5. These estimates are provided in the first two stages of the APMT Project Excellence Life Cycle; probe and development phase (see Figure 1 in section 1.2.1).

A Class 5 estimate is generally provided in the probe phase. A Class 4 estimate is usually provided in the early development phase. The civil engineering department provides the Class 5 and 4 estimates.

The Class 3 estimate is provided at the end of the development phase. The Class 3 estimate is provided by an external consultant and supervised by the civil department.

The Class 1 and 2 estimates are used within the implementation phase to monitor the actual costs. The actual costs are verified with the estimated costs. The verification shows the development of the costs compared to the estimate and provides information on the final costs. Adjustments can be made at an early stage.

An example of a civil cost estimate without the contingency factor is presented in Figure 4. The estimate only includes the price and quantity of the different terminal components. The project estimate may present the costs reflecting the largest probability.

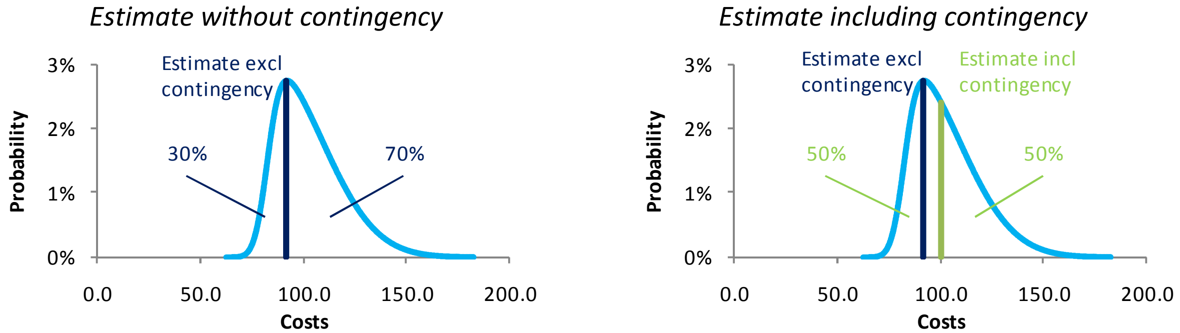


Figure 4; Project cost estimate

The graph presenting the estimate without contingency also shows the probability of cost overrun. The probability the actual costs will be within budget is lower than the probability of cost overrun. Eventually one expects higher costs. The expected value of the cost density function equals the probability of cost overrun and control and is beyond the project estimate. Additional budget is reserved to safeguard the estimate from overrun; contingency. The estimate including contingency is shown in Figure 4.

If the project variability increases, the density function for the project cost estimate widens. If an equal probability of cost overrun is accounted, the contingency budget increases. It is therefore concluded that the contingency factor depends on the variability of the project. A higher variability drives a higher contingency ratio. The APM Terminals' estimates are assumed to represent an expected value (i.e. P50) of the project including a contingency factor. The range indicating the confidence interval is assumed to be symmetric. The typical ranges used are already shown in Table 1. The contingency is added to the raw estimate; i.e. the most likely value.

The applied contingency is driven by various sources. The APM Terminals' engineers classify contingency to three different categories;

1. Development allowance; risk of (under-)estimation, i.e. construction items/quantities/quality/specifications that are not included in the scope of work due to lack of detail; known-unknowns.
2. Construction contingency; risk of events, i.e. events that may or may not suddenly occur or which happen by chance, particularly during construction; unknown unknowns.
3. Management reserve; covering possible expenses caused by unexpected and/or unknown events. Whereas Construction Contingency serves primarily for downward governance and target setting, the Management Reserve serves for upward governance and adjusting to the company's project risk appetite.

To come up with a profitable business case the cost-estimate is discussed between the Project Manager and the Civil Engineer. An underestimate of the CAPEX will lead to cost overrun, but an overestimate will not result in a winning tender. Therefore an iteration process is started to balance the CAPEX with the scope of works and the commercial assumptions.

1.2.5 Limitations and improvements on current civil CAPEX estimation process

At present the civil CAPEX estimate is presented as a fixed value or point estimate. The estimation process has its limitations. Two projects with equal civil CAPEX estimate but different variability are now reviewed equally. Both projects will show a difference in probability of cost overrun. Above that, the actual costs may differ if cost overrun takes place. The probability of cost overrun or consequence of cost overrun are neither known nor taken into account. But this information influences the decision to construct a new terminal.

One could argue the contingency factor is incorporated within the estimate taking the variability into account. Thus, the two projects with an equally estimated civil CAPEX can be compared. Unfortunately this is not true. The preceding showed contingency is a budget to account for the variability and prevent cost overrun. The civil engineers estimate the contingency on expert judgment with limited insight in the project variability. As a result, the contingency estimate is prone to inconsistencies.

The estimate is provided within the design phase of a terminal. It might be useful to determine the cost and variability drivers. The present methodology does not present these drivers. This makes it difficult to change and align the civil CAPEX estimate with the scope of works and commercial assumptions.

The preceding makes it likely to conclude additional value is provided by incorporating uncertainty in the cost estimate model. The insight on the uncertainty provides APMT additional information during the development of the business case in various ways;

- By developing a model showing outcome of the estimate and inherent uncertainty, the likelihood of the terminal construction costs can be identified. This provides additional information in the possible realization of the costs and feasibility of the business case.
- Within the current estimation process the costs of the defined scope are estimated. To account for uncertainty and lack of information, contingency is added to compensate the project uncertainty. No link is present between the estimate and the contingency ratios. By computing the project variability, the contingency ratios can be based on project uncertainty.
- From the present estimate the cost and variability drivers are hardly apparent. Taking uncertainty into account it is also possible to determine the cost and uncertainty driving variables. The identification of these driving variables makes it possible to come up with measures to reduce the project cost and risk.

1.3 Research objective and questions

The previous section clarified the limitations of the current process within APM Terminals to determine and present an estimate of the civil construction cost for a new container terminal. Taking uncertainty into account provides support to APM Terminals in developing and evaluating a new business case.

Therefore the following research goal has been developed:

> *"Currently, CAPEX for civil infrastructure in APM Terminals is estimated and reported in a purely deterministic manner and presented as a point-estimate. The objective of the study is to* ***propose a practical model that can be used by APMT's civil engineers to estimate and present the estimate including uncertainty in a clear and consistent manner****"*

1.3.1 Research questions

From the goal of the research the main questions are filtered;

1. Which issues can be identified from the estimation process?
2. What needs to be improved and why?
3. What are the requirements a new model to estimate civil CAPEX has to meet?
4. Which tools are available and which of these can be used to model the estimate and estimate uncertainty and how to use these tools?
5. Do these tools present the engineering judgment of uncertainty on the projects clearly and correctly?
6. Can the tools be used in practice?
7. How to present the civil CAPEX estimate?
8. How to identify the main cost and variability drivers?

The first three research questions focus on the explication of APM Terminals' question on cost estimation difficulties and the requirements for the solution. Question 4 focuses on the application of academic literature in practice. Questions 5 to 8 focus on the functionality and performance of the future model.

1.4 Research outline

To answer the questions the research is built from several parts. In this first section the root cause and possible improvements have been clarified; APM Terminals' civil engineers present estimates of new terminals' construction costs as a fixed point estimate. The estimate does not reflect the inherent uncertainty to identify the likelihood of construction costs. A model that allows for uncertainty establishes a link between the contingency ratios and identifies the cost and contingency drivers of the construction costs.

In the second chapter an academic view on cost estimation, uncertainty and risk is described. Insight in estimation processes show the challenges of cost estimation. The definition of estimates is given and it is shown that an estimate is based on a probability density function of the project. The density function shows the project risk and uncertainty but also introduce risk and uncertainty.

Being able to present and deal with uncertainty and risk, these concepts are clarified. The process from uncertainty to risk is shown. Several classifications of risk bearing elements are shown based on the process. The classification will be the input of the model to eventually present the project density function.

Several tools to compute the estimate and the associated variation from the input are described in chapter 3. Three different probabilistic tools will be introduced to estimate the construction costs. Above that, the pros and cons of the different tools are discussed. Special attention will be paid to the incorporation of interdependency between the different elements. Describing the different tools will answer question 4.

Chapter 4 is an intermediate chapter to define the APMT requirements for a new estimation tool. Based on the requirements the most suitable tool will be selected. Thus, this will answer question 3 by defining the APMT requirements.

In chapter 5 an example business case is introduced. This business case is used to verify and determine the most suitable model. Before the tool is selected, the business case is introduced. The scope of works is clarified and possible exceptional events are presented.
In the first part of Chapter 6 the performance of the tools is presented. The preceding has shown that APMT civil engineers mainly provide rough estimates (Class 5 and 4) and supervise detailed estimates (Class 3). Therefore, two different estimates will be provided; a rough estimate (Class 5) and a more detailed estimate (Class 3/4). The current estimation practice of APM Terminals will be presented as a base case aside from the selected tools.
The output of the estimates using the different tools will be reflected with engineering judgment. The engineering judgment is used to verify the correctness of the output and practicality of the different tools. This selection is made in the second part of Chapter 6 and thereby answers question 5 and 6.
Chapter 7 proposes a method to present the civil CAPEX estimate to the end users and decision makers. Different numbers arising from the probabilistic estimate are presented. This makes it possible to present the civil works, the uncertainty and risk of the civil CAPEX estimate. This enables the end users to take a solid investment decision.
In the last chapter the main conclusions of this MSc thesis are presented and recommendations for future research are proposed.

2 Introduction to estimation, uncertainty and risk

This chapter will introduce the process to come to cost estimates and the risk and uncertainty involved within estimating. In general cost estimation involves a lot of complexity. A thin line has to be walked as on the one hand the actual costs should not overrun the estimate. On the other hand, too conservative estimation does not result in winning tenders or profitable results. Based on the available information a competitive estimate should be provided to profit from the investment opportunity. Thus the estimate should be large enough to prevent cost overrun and sharp enough to win the tender.

The process to estimate and actively deal with the consequences of the estimate is part of risk management. Risk management comprehends the identification, assessment, and prioritization of risky elements and events. This analysis is followed by coordinated and application of resources to minimize, monitor, and control the probability and/or impact of unfortunate events or to maximize the realization of opportunities.

One will find many articles searching for risk management within large infrastructural projects. The search result at Science Direct (search portal of Elsevier Science) of the tag 'risk management of large infrastructural projects' is 28,820 articles.

The majority is related to qualitative risk management tools. Qualitative methods describe and mark risk to actively manage the exposure. The methods are valuable to plan and control a project. These methods provide no quantitative or measured proof. This makes the tools sensitive to errors or inconsistencies.

Only a minor part of the articles is on quantitative risk management. Quantitative risk management (QRM) aims to quantify the risk. The quantitative risk management is a supportive tool of the qualitative risk management. The impact of risky elements can be measured to identify measures. As a result the right- or wrongness of the risk management is proven.

The availability of little literature might be due to the following reasoning. First, one needs to have a firm understanding of technical risk. Secondly a firm understanding of statistical modeling is required to convert the technical risk of a project to a probability density function. Wang and De Neuville (3) address that many analysts do not possess knowledge on both fields. Therefore, a small amount of articles is available.

This chapter describes and positions the theory within present literature. First the academic view on cost estimation is presented including the use of contingency and the evolving estimate accuracy. Thereafter risk and uncertainty are clarified. Knowledge on these items makes it possible to identify the model input.

2.1 Process of cost estimation

Cost estimation within APM Terminals is described within the first chapter. This process is in line with literature(3). A design brief is required to define terminal requirements. Examples of APM Terminals' requirements are the amount of cargo per year or the size of the design vessel. A terminal layout is based on these requirements.

From the terminal layout, the terminal components and process flow are determined. The terminal components differ between projects, but the new terminal will likely consist of a container yard, quay wall and several administration buildings. The process flow describes the flow of cargo within the terminal and determines the position of the different components.

From the terminal components and process flow, the structural design is derived. The structural design is the result of a strength analysis. The design specifies the dimensions, structure and construction materials of the components.

Knowing the dimensions, structure and construction materials, the amount of building materials and construction method can be determined. First the quantities are estimated; thereafter the unit rate is estimated. The product of both quantity and price determine the price per terminal item. The cumulative costs of the items will set the raw estimate.

One or more contingency factors are added to the raw estimate. It is added to achieve a certain probability the actual costs do not overrun the estimate. In mathematical terms the estimate is defined as;

$$Estimated\ Costs = \left(\sum_{i=1}^{n} P_i * Q_i\right) * (1 + \gamma_k)$$

Q_i represents the quantity of element i, P_i is the price of element i per unit of quantity, and the product is the price per element i. The sum of the different elements, shown with the Σ-mark, is the raw estimate. The contingency factors γ_k are added additionally to the estimate.

2.1.1 Contingency factor

The contingency factors account for the uncertainty on the estimate. The incorporation of uncertainty by contingency is two-fold. First, information and scope are lacking within the estimate. One shall face cost increase and a budget should be incorporated. In the first project stage overhead, supervisory or tax may be incorporated as an additional factor. One could define this contribution “known-unknowns”. It is known by experience these unknown costs will come to the estimate.

Secondly, contingency is a safety factor to lower the probability of cost overrun. Chapter 1 showed the application of contingency to account for project variability. A budget is reserved to protect the company against cost overrun. This part of the factor can be seen as “unknown-unknowns”. From a protective perspective a budget is reserved against unknown costs. It is also unknown whether the budget has to be spent.

The two-folded meaning of contingency introduces confusion. One factor is applied but different sources drive the factor. Dysert(4) has pointed out the various drivers of contingency. Examples are;

- Not everything is incorporated or correctly incorporated within the estimate.
- Variability in pricing due to changing conditions over time within the labor market, availability of construction materials or construction market.
- Variability in quantities; the design is a sketch in the early project phase making it difficult to approximate the exact value. Otherwise, exact quantification of variables (like the amount of dredging material) may be impossible.

Dysert(4) stated contingency does not include exceptional events that might happen during construction. Exceptional events reveal information that is not incorporated within the cost estimate. Examples of exceptional events are the discovery of contaminated soil or an earthquake. These events should be accounted for within a special budget; the management reserve. Research by Boschloo and Vrijling (5) has shown exceptional or unforeseen events happen at every construction project. Therefore one should incorporate these events within the cost estimate.

2.1.2 Estimate accuracy or the confidence interval

During the development process of a terminal more information becomes available. Additional information on the local conditions is present and the project scope becomes more detailed. This results in increasing estimate accuracy. In the early stage of the business case little information is available and the estimate is prone to uncertainty. The estimate is therefore used to evaluate project feasibility.

As the business case progresses over time, the information and scope becomes more detailed. As a result, the estimate accuracy increases. The final estimate is used to set the project budget and forecast the contractor tender bids.

It is general practice within the construction industry (6)(4) to assign an estimate accuracy range to the estimate. The range bounds a confidence interval (e.g. 80%). This confidence interval represents a probability (e.g. 80%) the final costs will be within the accuracy range.

The upper and lower limits bounding the interval are usually expressed as a percentage of the point estimate; the limits equal the point estimate including plus or minus 20% of the point estimate. The limits are chosen in relation to the confidence interval and reflect the probability the cost do not overrun the boundary. In the example of an 80% confidence interval, the lower boundary generally represents a threshold the costs do not overrun with 10% probability; the P10. The upper boundary is usually the P90; the 90% probability the costs remain within the boundary.

In the left graph of Figure 5, the raw estimate and final estimate including contingency are shown. Additionally the confidence interval and the boundaries are given.

Industry and academic practices propose to separate the design process to different stages. A cost estimate is provided at each stage to evaluate project feasibility.

As the project passes several stages and the design layout will be developed into more detail, the accuracy will improve over time. The spread of the confidence interval determines the estimate accuracy. Improving estimate accuracy tightens the spread. The different project stages, the evolving estimate and the improving estimate accuracy are shown in the right graph of Figure 5.

The evolving accuracy also has an impact on the contingency factor. More known elements are known and priced within the estimate. The evolving accuracy reduces the project variability. The contingency factor should be reduced to maintain the same probability of cost overrun.

The APMT process to come to a new terminal is described in the first chapter. The process is divided into different stages as was shown in Figure 1. An estimate is provided at every stage. The estimate accuracy improves over the different stages as the design outline and scope of work becomes more detailed. This is also shown in Figure 5 and is in line with academic literature.

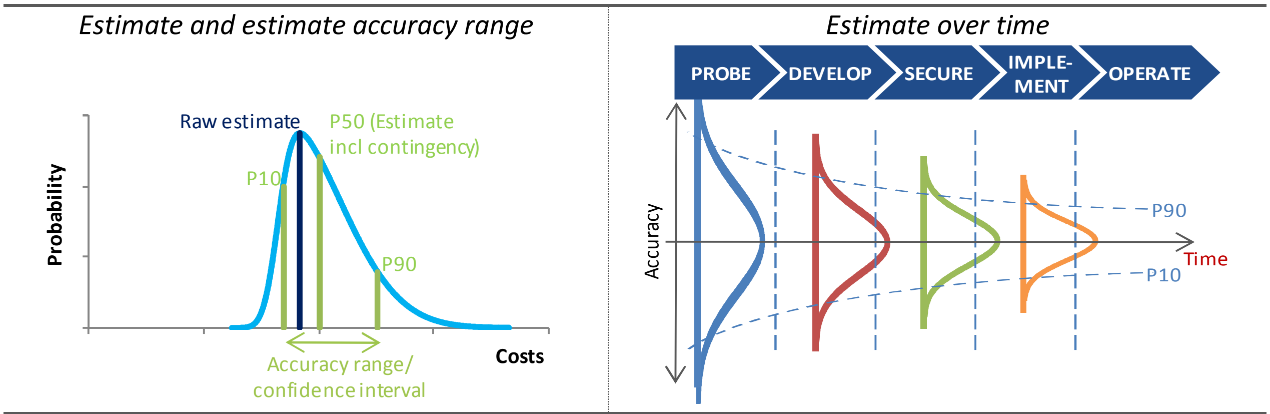


Figure 5; estimate accuracy range

The preceding described the process to come to an estimate. Based on the design brief and terminal layout an estimate is provided. During the process to develop a new terminal an estimate is provided several times. As the information on the terminal design, layout and scope improves over time, the estimate also becomes more accurate.

One is now familiar with the process to come to an estimate. But in order to provide an estimate one should know what an estimate really is. This is treated next.

2.2 Defining cost estimates

Crucial to provide an estimate is to know what an estimate is. The Association for the Advancement of Cost Estimating (AACE) describes an estimate as: *"an evaluation of the elements of a project or effort as defined by an agreed-upon scope".* Dysert(4) points out that the definition should address more explicitly the uncertainty involved; *"a prediction of the probable costs of a project, of a given and documented scope, to be completed at a defined location and point of time in the future"*.

The definition of both the AACE and Dysert state the estimate is a prediction (of the construction costs) conditioned on information. As an estimate involves uncertainty, the estimate is (by its definition) a probability density function. The probability density function shows the uncertainty on the outcome.

The probability density function is developed using a set of information. Dysert addresses the information includes project scope, location and time. The information content changes over time. One spends time acquiring information by e.g. extensive desk research, site visit or soil investigation.

Based on the preceding it is likely to conclude the mathematical reproduction of an estimate is;

$$Estimated\ costs = P(X|\mathcal{J}_t)$$

In the above formula X represent the costs and $P(X|\mathcal{J}_t)$ the conditional probability density function of the costs X conditioned on the set of information at a point in time $\mathcal{J}_t$.

The estimate is distilled to a number with a certain probability of cost overrun. Examples are the expected value (i.e. P50) or an upper limit P70 or P90.

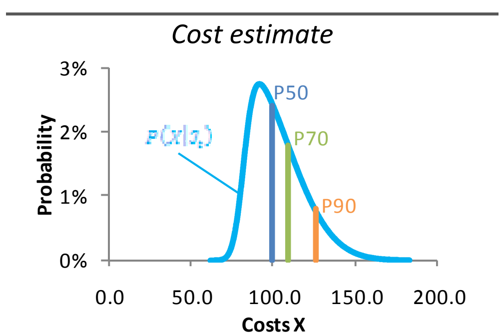


Figure 6; Definition cost estimation

Within the first paragraph of this chapter the estimated costs were defined differently;

$$Estimated\ Costs = \left(\sum_{i=1}^{n} P_i * Q_i\right) * (1 + \gamma_k)$$

It was also addresses this definition is commonly used in practice. The costs are estimated and based on experience contingency factors are added. A fixed point estimate is the outcome of this approach. The second definition shows the estimate as a probability density function. The probability density function is the foundation to determine the costs.

Both approaches have the same objective; provide an estimate of the construction costs. But both approaches are different as the probability density presents the risk involved with the estimate. Within the other approach contingencies are added to ensure a similar risk profile. One should notice it is unknown whether or not the risk level is achieved.

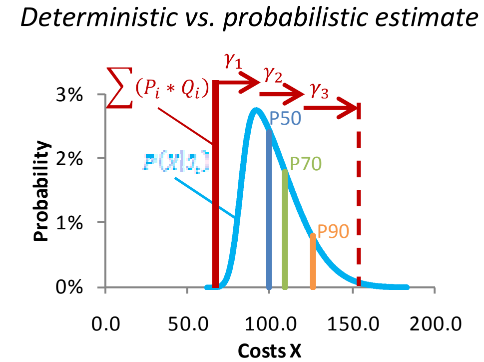


Figure 7; Difference between estimates

This process is shown in Figure 7. Various estimates can be provided accordingly a certain risk level (using the probability density function). The other approach attempts to achieve a similar estimate by applying different contingency factors. It is unknown whether this is achieved. The estimate presented with the probability density function is called a probabilistic estimate; the other (using contingencies) is called a deterministic estimate.

This section showed what an estimate means. It showed two forms of estimates; a probabilistic and deterministic estimate. Additionally, an estimate was linked to risk. To improve the understanding of risk, the next section will treat risk in detail.

2.3 Risk and uncertainty

Based on uncertainty one can determine the risk involved with the cost estimate. Uncertainty is the doubt on the (to be realized) costs. The variability or spread of the distribution function presents this doubt. Risk is some measure of uncertainty on the outcome. As the probability density function shows the estimated likelihood of construction costs, it also presents the risk.

As described by Dysert, the probability density function reflects the possible costs of a project, given the amount of information $\mathcal{I}_t$. However not all uncertainty and risk is shown; one does not possess all information. Therefore it can be concluded that $\mathcal{I}_t$ is biased. The result is a probability density function deviating from its true shape.

In addition, the known information should be modeled correctly. Changing variables and the joint behavior of the variables introduce uncertainty. Often there is limited insight on the (joint) behavior of variables. And above that, the inaccuracy of the approximation to describe the variables' behavior creates even more uncertainty on the probability density function.

To come up with the best estimate one should have all information available and correctly modeled. Because the information is not complete, correct or present, an estimate will involve uncertainty.

Despite the correctness of the information and the incorporation of the information, the probability density function presents the risk and uncertainty on the construction costs.
This shows that an estimate is linked to risk and uncertainty. Being able to present and actively deal with risk and uncertainty, one should have a clear idea of risk, uncertainty and the inherent link. The link between risk and uncertainty troubles (and often confuses) the meaning of risk and uncertainty.
In 1921 Frank Knight (7) separated risk from uncertainty. He described uncertainty as something random and immeasurable. Risk can somehow be quantified and is therefore different from uncertainty. The Knight's approach is questioned by others though. Hubbard(8) stated a likelihood or probability of occurrence is attached to risk. Risk involves uncertainty but uncertainty does not necessarily involve risk. This approach makes it likely to conclude risk is a measure of uncertainty.

In order to present risk, one should identify uncertainty and quantify the likelihood. Applying this process to terminal construction projects, one should identify the scope of work and events. Additionally, the uncertainty on the scope of work and events should be identified; thereafter the probability of occurrence and consequence should be estimated. In general the process from risk to uncertainty can be shown by;

Identification scope of work and events→ Uncertainty on the outcome→ Probability of occurrence →Consequence→ Risk

This process is mathematically summarized by the formula;

$$Risk = probability * consequence$$

This formula computes risk as an expected value of the scope and events, often expressed in some amount of money. Based on a set budget, one can determine the risk on cost overrun as a clear expected value;

$$Expected\ Utility = \int_{budget}^{\infty} X * P(X|\mathcal{I}_t)\, \partial X$$

This conditional expectation represents the expected costs in case of cost overrun. The formula above is derived from the estimate definition; X show the costs and $P(X|\mathcal{I}_t)$ the project density function.
It has to be remarked that the presented formula values risk neutral; a 20% chance of losing 10 million is equally valued as the 20% probability of winning 10 million. In reality, one does not value both outcomes equally. From an economical perspective one should therefore convert the formula to;

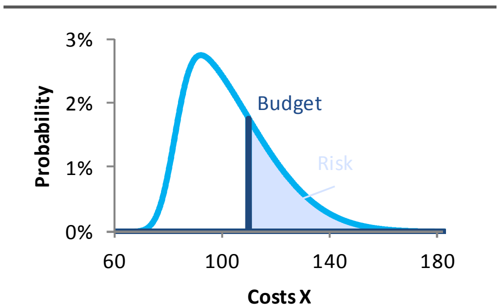


Figure 8; Risk

$$Expected\ utility = \int U(X) * P(X|\mathcal{I}_t)\, \partial X$$

$U(X)$ is a utility function quantifying the risk appetite. In the risk neutral situation the utility function is linear; $U(X) = X$. Risk aversion is shown with a convex (i.e. increasingly upward) utility function. Other examples are the exponential or quadratic function. Within the process to show the project density functions, all risk bearing processes have to be valued equally. This justifies using this formula; the decision maker has to evaluate the project construction costs and not the model itself. I will therefore use this definition to come to the project density function.

The previous section discussed the meaning of risk and uncertainty. But it is still pretty conceptual. One of the objectives of this thesis study is to determine risk of construction costs. Any additional literature may be helpful. Vrijling(3) focuses on reliability of systems. A system is defined as a coherent set of matching objects and their parts. An example is the cash flows of a new terminal.
Reliability is related to failure. Failure means the system (i.e. coherent set of matching objects) do not meet the intended objective. Referring to the terminal example; failure means the costs overrun the revenues.
Reliability is the rate of (no) system failure. In case of the terminal, reliability is the rate of profit making.
Risk management focuses on the probability of failure and how to avoid failure. Engineering examples are strength analyses. The system can be described with:

$$System's\ strength = Resistence - Load$$

Economic examples relate to profitability analysis. The financial performance of an APM Terminals project is described by the system;

$$Profit = Revenues - Costs$$

The profitability is to review the entire project. This thesis focuses on cost estimating and is just a part of the project evaluation. The aim is to estimate the construction costs and present the inherent estimate uncertainty. A probability density function can be used to estimate the costs (i.e. calculating the expected value) and present project uncertainty and risk.
One can also use the estimate as a budget to construct a new terminal. A budget (i.e. cost level) is defined with a certain probability of overrun. Therefore the system Z can be defined:

$$Z = C_0 - Costs$$

In the system Z, C_0 means the cost estimate or construction budget and $Costs$ the estimated cost density function. In the preceding paragraph the density function was defined as $P(X|\mathcal{I}_t)$. Z is denoted as some amount of money presenting the result of the estimate versus the actual costs.
This thesis focuses on the development of the cost density function $P(X|\mathcal{I}_t)$ to determine the estimate C_0 and present the estimate uncertainty. After $P(X|\mathcal{I}_t)$ and C_0 are determined, an iterative loop is run to review the system reliability and cost budget. Reliability of systems is treated in detail in the appendix.

Based on this system approach one can determine the risk. The costs of a project should be estimated to set a budget and derive the risk. In order to estimate the costs it may be helpful to formalize the estimation process. I call this process classification. Based on this process one determines the risk. It is an important step to present the estimate and the inherent estimate uncertainty. This important process is similar to cooking. Without the right ingredients one will never end up with something tasteful. Therefore the classification process is treated in more detail within the next paragraph.

2.4 Classification of uncertain elements and events

Showing the risk of construction costs, one should identify the project scope and the inherent uncertainty. This will cover all the works that are planned to be constructed. In addition to the project scope, several exceptional events have to be identified. Boschloo(5) showed that during every project of Rijkswaterstaat (Dutch ministry of infrastructure) several exceptional events happen. In any case not all events will happen, but a few will.

Having identified the events, the probability and consequence of all risk bearing items should be estimated. Now we can set the budget to finalize the system. This makes it possible to determine the risk. The foundation of a probabilistic model is therefore:

- Identify the scope elements
- Estimate the price and quantity of the elements
- Identify the possible events
- Estimate the probability of occurrence and the consequence of the events

A practical approach is presented by Vrijling (3). Within the design layout, the construction elements are specified. Vrijling defines the construction of these elements as normal elements. These elements will be built certainly, but the price and quantity are uncertain. This price and quantity uncertainty is reflected in the probability density function.

The left graph in Figure 9 shows normal events. The probability density function shows the uncertainty on the output. The surface of the density function is equal to 1 (or 100%), stating that the output will materialize within the project density function.

Vrijling classifies exceptional events like Boschloo(5). These events happen rarely, but have a major impact on the construction costs and the variability of the costs.

An exceptional event is shown in the middle graph of Figure 9. The bar reflects the probability the event will not happen; the resulting costs are zero.

The bell curve presents the outcome if events happen. The surface covered by the bell curve is the probability of occurrence and the bell curve itself shows the uncertainty on the output.

Additionally Vrijling identifies design uncertainty. In the early stage of a project several options are possible. Between two or more alternatives may be chosen, but one will be selected. Several distribution functions are selected representing the alternatives. The total surface covered by the density functions equals one.

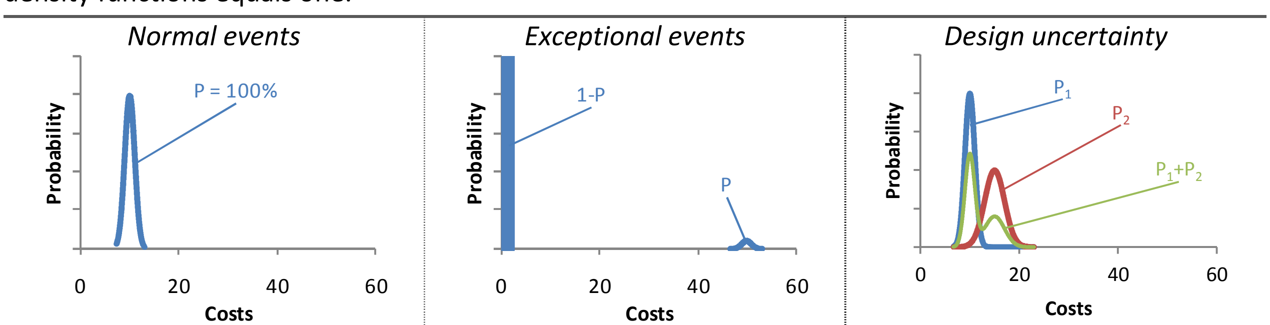


Figure 9; event and element classification

After estimating the scope of work and the various exceptional events, one can determine the project density function. This step is taken using statistics. A brief introduction to statistics is provided in the appendix. The coherent behavior of various elements is an important factor determining the system's reliability. The joint behavior can be measured with correlation. Correlation drives project variability (measured with standard deviation) and variability is an important factor setting the budget and determining project risk. Therefore joint behavior will be treated in the next section.

2.5 Dependency

Dependency between, for example, a quay wall and container yard means that the costs of both are related. Dependency is present as the structure of the quay wall and the container yard depend on the soil conditions (i.e. both are constructed on similar soil conditions). Above that, both terminal components are constructed by the same contractor and consist of similar construction materials. This makes it likely to conclude mutual drivers contribute to the costs causing an interrelationship. This section will explain interdependency. In a later stage of this thesis it will be used as part of one of the estimation tools. Interdependency is often measured with correlation. Correlation is a measure of linear dependency between two variables C_1 and C_2. It is the quotient of the covariance and the product of the standard deviations;

$$\rho = \frac{cov(C_1, C_2)}{\sigma_{C_1} * \sigma_{C_2}}$$

The correlation can best be explained graphically. In Figure 10 the simulated results of four scatter plots are presented. Each dot of the scatter plot represents a simulation of the costs to construct the wharf and container yard. The scatter plots show from left to right increasing correlation between wharf and yard costs.

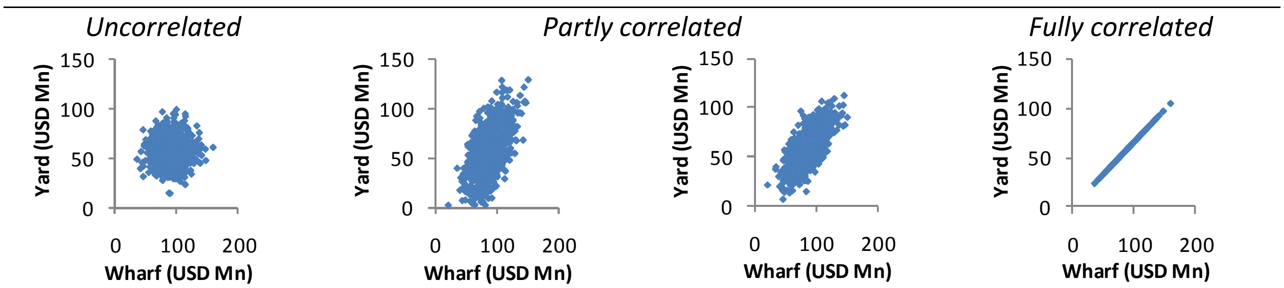


Figure 10; Differences in correlation

Real time historical data will not present these clouds at all. This is shown in Figure 11. The historical data is scattered without any coherence. This lack of joint behavior is due to the heterogeneous data. Heterogeneity is introduced by the different requirements, structural designs and market conditions.

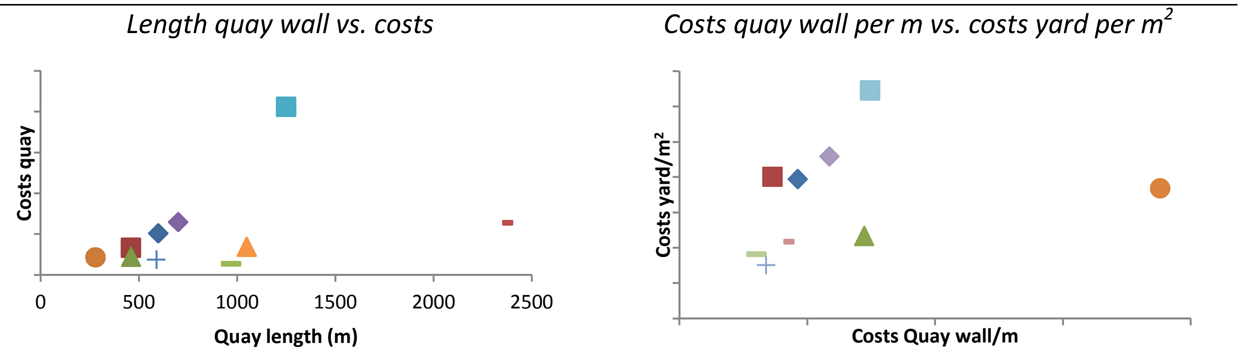


Figure 11; Joint behavior

To explain (linear) dependency I return to Figure 10. The simulated behavior of the left graph (in Figure 10) shows the independent (i.e. uncorrelated) situation in the left graph. All the data points are scattered within a circle. This means the the wharf costs do not influence the yard costs at all.

The straight line at the right of Figure 10 shows the situation of linear dependency (i.e. full correlation) between the two terminal components. If the situation of linear dependency is true, the costs of the container yard are given by;

$$C_{yard} = \alpha + \beta * C_{wharf}$$

The two graphs in the middle of Figure 10 show two examples of an intermediate situation. A straight line can be drawn through the cloud with an error term ε around the straight line;

$$C_{yard} = \alpha + \beta * C_{wharf} + \varepsilon$$

The error term ε is the vertical deviation between the data point and the line $\alpha + \beta * C_{wharf}$. The linear equation is usually determined by computing the minimum of the squared error term. More information on this analysis and the academic background can be found in Heij et al. (9)

A small error term shows closely scattered dots around the linear equation. This shows a strong relationship and a high correlation. A weak relationship is characterized by large(r) deviation from the linear equation (i.e. large error term) and low correlation.

Thus one may conclude the spread plays an important determining the (linear) relationship. From the linear equation one can identify three different standard deviations; the standard deviation of C_{yard}, C_{wharf} and the error term ε. The standard deviation of C_{wharf} is composed of the standard deviation of C_{wharf} and ε. As C_{wharf} and ε behave independently, the relationship of the standard deviations is given by;

$$\sigma^2_{C_{yard}} = \beta^2 * \sigma^2_{C_{wharf}} + \sigma^2_{\varepsilon}$$

This relationship is graphically shown with a triangle in Figure 12. One should remark the triangle is the same as the Pythagorean Theorem triangle. This proves the standard deviation of the yard is composed of the orthogonal (i.e. perpendicular projected) vectors $\sigma_{C_{wharf}}$ and σ_{ε}. The vectors are perpendicular projected by their independent behavior.

Based on this relationship one can determine the percentile contribution of the wharf's variance (i.e. squared standard deviation) to the yard's variance. This is presented below;

$$\frac{\sigma^2_{C_{yard}}}{\sigma^2_{C_{yard}}} = \beta^2 * \frac{\sigma^2_{C_{wharf}}}{\sigma^2_{C_{yard}}} + \frac{\sigma^2_{\varepsilon}}{\sigma^2_{C_{yard}}}$$

The contribution of the wharf's variance to the yard' variance is called the Sum Squared of Residuals (SSR) or R^2; the other term is called the Sum Squared of Errors (SSE). This relationship is shown in the middle graph of Figure 12. The graph also shows R^2 is the orthogonal projection of the wharf's variance on the yard's variance.
The R^2 is also related to the correlation coefficient; the squared root of the R^2 equals the correlation. The correlation is the projection of $\sigma_{C_{wharf}}$ on $\sigma_{C_{yard}}$; $\rho = \frac{\sigma_{C_{wharf}}}{\sigma_{C_{yard}}}$. The relation between R^2 and ρ as well as both projections are shown in the right graph of Figure 12.

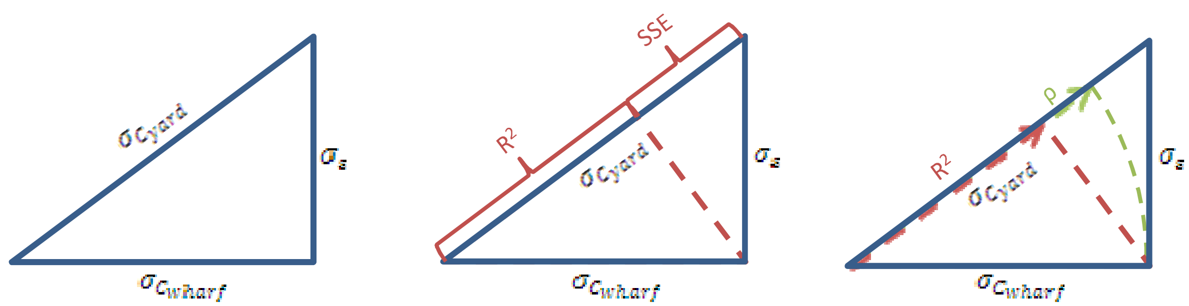


Figure 12; Explanation of R^2 and ρ

This paragraph introduced interdependency and its (linear) measures; correlation and R^2. Some linear relationship between two variables is the basis of interdependency. The linear formula is the basis to determine the explained and residual variance. The R^2 is the contribution of the explained variance to the total variance. Additionally, a link is established between R^2 and correlation.
As interdependency is explained, one can start incorporating dependency within a model. One of the models will be based on various drivers mutually driving the costs of the scope. The insight in dependency and the R^2 measure are important to understand the model and its output.

2.6 Positioning within other studies

This study is related to other several earlier studies. Vrijling has proposed a method to estimate and present risk of construction projects. This is described in the post-academic course "Probabilistic design". While constructing the models this approach will also be the starting point for this study.
Closely related to the Probabilistic Design course is the paper by Vrijling and Boschloo. They point out the occurrence of exceptional events at construction projects. This addresses the need to incorporate these events within the estimate.
As the estimation procedure is based on expert opinion this study is also related to Tversky and Kahneman. Their (Nobel prize winning) study on expert elicitation pointed out the difficulties of expert elicitation. The heterogeneity of construction projects makes historical data difficult to use. Therefore one relies on expert opinion. Their study pointed out why expert elicitation may be biased, inconsistent or wrong.
Other studies propose more sophisticated modeling to improve the quantitative risk management. McCabe uses Bayesian Belief Networks to forecast or estimate construction time of projects. The progress of construction projects is often conditioned on the progress of other components. The conditional property of Bayesian statistics is a powerful tool to model this behavior.

Pindyck, Wang and De Neuville use financial option theory to value construction projects or define optimal stopping rules. This approach takes the irreversibility of investments and the opportunity of waiting into account.

2.7 Summary

This chapter introduced the construction practice to estimate costs of construction projects. An estimate is provided several times while developing new terminals. As the uncertainty on the design and other conditions decrease during the process, the estimate becomes more accurate.

Several definitions of an estimate are described. It is shown the estimate is the expected value of the project costs given a set of information. The estimate (that is deduced from the project density function) presents the risk and uncertainty on the construction works. The set of information covers the terminal design, scope of works or site information. Possibly, this set of information is biased and the information is not always well incorporated. This introduces uncertainty and risk on the estimate itself.

Thereafter risk and uncertainty are clarified. Uncertainty is some random process whereas risk is the quantification of this process. The process from risk to uncertainty is described in more detail and the mathematical formula derived from this process is presented.

Having clarified risk and uncertainty, the input to construct a probability density function was described. Two different classifications of events and elements are presented; normal events and exceptional events. This input is the foundation to derive the probability density function.

Some additional statistics is explained to incorporate dependency within the estimate. Correlation measures the linear dependency between elements. Based on the linear relationship, the R^2 measures the contribution of the dependent part to the total variance. This insight in dependency consists of the last step to start constructing various models. In the last paragraph the study was positioned within other literature.

3 Constructing a model to estimate cost

This chapter introduces various models to estimate costs. The preceding chapter introduced any background information on cost estimates and risk. Taking the information into account, various models are developed to come to the cost density function. Based on this cost density function, budget and risk can be determined.

Step by step, I will progress towards a model presenting a cost estimate as realistic as possible. The starting point is theory developed by Vrijling. He defined three different levels of cost estimates; level I, II and III. A level I method is a purely deterministic estimate. The outcome is a fixed point estimate without any information on the uncertainty or risk. Above that, the deterministic method is similar to the current approach of APM Terminals to estimate costs.

A level II estimate is referred to as semi-probabilistic. Of all the different functions only the mean and standard deviation are taken into account and fitted to the normal distribution. It is a first (rough) step towards a probabilistic approach; applying the method, one will have an idea of the expected value and the project variability.

Level III estimates are purely probabilistic. Simulation is used to come to a probability density function of the costs. I will come up with four different estimates. First I will assume all functions normally distributed. This may be confusing; a level II estimate also assumes normality. It will be shown this level III approach differs from the level II estimate. The second level III method will take different distribution functions into account. Thereafter a model with dependency is introduced. Drivers are identified contributing to the costs of various scope elements. This contribution introduces dependency. The last estimation model will take shocks on the drivers into account.

Wrapping it all up, this chapter will present six different models. The order and the different models are shown below;

1. Level I deterministic method
2. Level II semi probabilistic method
3. Level III normally distributed functions
4. Level III various functions
5. Level III dependency
6. Level III dependency and shocks of drivers

3.1 Level I deterministic method

The deterministic tool is often used within cost estimation. The different scope elements are priced by estimating the quantities and unit rates. The deterministic tool estimates all elements fixed. The estimate does not show variability but the estimate is still prone to variability. To account for the variability a factor is added; contingency. The factor is some sort of additional budget to achieve a certain probability of cost overrun. Mathematically, the deterministic estimate is;

$$Estimate = \left(\sum_{i=1}^{n} P_i * Q_i \right) * (1 + \gamma_K)$$

P_i and Q_i are the unit rate respectively quantity of the terminal elements. The product is the element's price and the sum of the different elements is the raw estimate. The raw estimate is multiplied with contingency factors γ_K to set the estimate. An example project is shown in Table 2.

Element	Unit Rate ($)	Quantity	Unit	Total
Engineering	2,000,000	1	Project	2,000,000
Break water	50,000	1,000	m^1	50,000,000
Quay wall	40,000	1,000	m^1	40,000,000
Dredging	4	3,000,000	m^3	12,000,000
Container Yard	75	100,000	m^2	7,500,000
Administration building	1,000	2,000	m^2	2,000,000
Subtotal				113,500,000
Contingency – CRAF 1				22,700,000
Contingency – CRAF 2				22,700,000
Total				158,900,000

Table 2; Deterministic example

The example reflects a typical Class 5 estimate. Only large components of the estimate are specified. The raw deterministic estimate is according to the above specified scope USD 113.5mn. To account for lacking scope and events contingency is added. The final estimate becomes EUR 158.9mn.

According to the current process APM Terminals assigns the limits of the confidence interval at 40% of the cost estimate; the costs are expected between USD 95.3 and 222.5mn construction costs. Figure 13 clarifies the meaning of the assumption. A possible density function is fitted to the estimated mean and confidence interval.

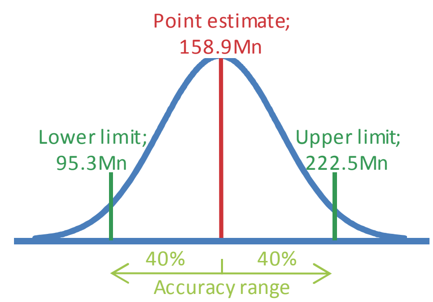


Figure 13; APM Terminals variability assumptions

Within the deterministic estimate contingency is added. The contingency is applied to account for unspecified scope, variability on the unit rates and scope, lack of detail and exceptional events. Additionally the applied contingency ensures a certain probability level of cost overrun. Within this deterministic estimate the unspecified budget and exceptional events ("known-unknowns") as well as the safety factor ("unknown-unknowns") are not explicitly incorporated.

Chapter 1 introduced the decomposition of the contingency factor within APM Terminals;

- Development allowance or CRAF 1; "known unknowns"
  The development allowance accounts for additional costs that are not specified. By experience it is known these costs will incur. In a later stage these costs may be specified.
  Typical examples are the discovery of additional soil that should be dredged or changes in the structural design due to conditions from assumed.
- Construction contingency or CRAF 2; "unknown-unknowns"
  The construction contingency covers possible additional costs. It is neither known (i.e. unknown) whether these costs do (not) occur nor the cause is known. Within this study the construction contingency is therefore defined as a safety factor. Commonly used examples are the occurrence of exceptional events like the discovery of a bomb.

These two factors equal within the class 5 estimate 40% contingency. An additional special budget is incorporated to account for excessive results of exceptional events; management reserve. The management reserve is another contingency factor; the factor is a percentage of the estimate.

A link is present between project variability, confidence interval and contingency factor. The boundaries of the confidence interval reflect a cost limit related to the project variance. A larger variability will result in a broader confidence interval. Above that, the contingency factor is driven by project variability. To ensure a similar risk exposure, a different variability should also change contingency. But the lacking insight of variability does not make this possible. Thereupon, large variability should result in a large confidence interval and contingency factor.

Project variability may vary (heavily) between projects. For instance, a terminal expansion after a few years is less exposed to variability than a new green field project. Thus, different project characteristics cause differences in the variability and applied contingency.

Within the deterministic estimation process contingency (accounting for project variability) is determined by the project stage, not by specific information. This leads to the conclusion that a direct link between the project variability and range lacks. The contingency is based on (unproven) past experience and engineering judgment. Therefore the method is often called 'black box method'.

To show the project variability the variability of the individual line items will be taken into account. The classification of the various elements is described in section 2.4. Knowing uncertainty of elements, it is possible to determine the project density function.

3.2 Level II semi probabilistic method

The semi probabilistic estimation tool takes the uncertainty on the outcome into account. It is the first step to a probabilistic estimate. Of all the different input variables the average and variability are incorporated. The shape of the different distribution functions is not accounted for; all different variables are assumed normally distributed.

The use of normally distributed elements is justified by the Central Limit Theorem. The sum of many (independent) distribution functions converges to a normally distributed probability density function. Knowing the mean and variance of the variables, one can determine the total distribution function.

Above that, the behavior of elements is often approximated with the probability density function. The normally distribution function well describes the behavior of oil or steel prices. These are liquid assets within mature markets leading to normally distributed prices.

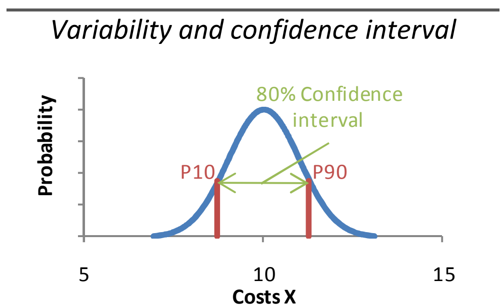


Figure 14; Variability and confidence interval

Using the normal distributed tool, one estimates the mean and standard deviation of the scope elements and exceptional events. This will result in a project density function. The mean is the expected value of the different elements.

The standard deviation is estimated by means of the variability. The variability is the quotation of the standard deviation and the mean; it is the deviation as a percentage of the mean. A variability of 10% on a 10mn element means the element will be between 8.7mn and 11.3mn. With 80% probability the estimate will be within the interval. In case of a 90% probability, the interval widens.

In the table below the example project is shown again. The expected value is assumed to vary according to the variability. Estimating cost more detailed, both price and quantity may vary. The product is assumed to be normally distributed as well. Analysis has shown that if the product of two standard normal distributed[1] functions is non-normal (i.e. χ^2-distribution). As this is not apparent, the assumption of normality is justified. The analysis outcomes are shown in the appendix.

Element	Unit rate ($)	Quantity	Unit	Total	Std deviation	Variability
Engineering	2,000,000	1	Project	2,000,000	300,000	15%
Break water	50,000	1,000	m^1	50,000,000	10,000,000	20%
Quay wall	40,000	1,000	m^1	40,000,000	6,000,000	15%
Dredging	4	3,000,000	m^3	12,000,000	1,800,000	15%
Container Yard	75	100,000	m^2	7,500,000	937,500	12.5%
Administration building	1,000	2,000	m^2	2,000,000	300,000	15%
Total				113,500,000	11,844,784	10%

Table 3; Standard normal example

The probability density function of the example project is shown in Figure 15. The density function assumes no interdependency between the different scope elements; no interrelation is assumed. The total standard deviation decreases to a small value (in this example 10%) which does not reflect the expert judgment.

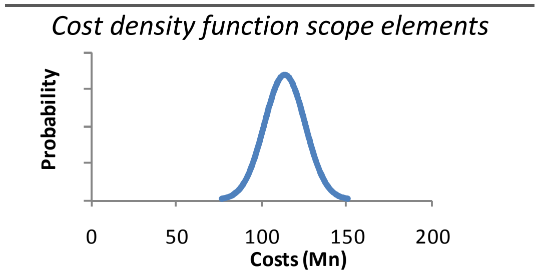


Figure 15; Normally distributed scope elements

Boschloo showed exceptional events happen at every project. Exceptional events are occurring events different from the assumptions and causes additional costs. As it is known exceptional events happen at every project, this is part of the "known-unknowns". Within the estimation example the exceptional events are not incorporated. Three exceptional events are presented below;

Event	Probability	Consequence	Variability of consequence	Mean event	Stdev event
Contaminated soil	10%	50,000,000	20%	5,000,000	15,329,710
Earthquake (Richter scale 7)	1%	100,000,000	25%	1,000,000	10,259,142
Bomb	5%	10,000,000	10%	500,000	2,190,890
Total				6,500,000	18,575,522

Table 4; Event examples

The new cost density function is presented in the left graph of Figure 16. With the introduction of exceptional events, the expected value and the variability increase. The old (excluding events) and new (including events) cost density function is shown in the right graph of Figure 16. The density function shifts to the right as the expected value increases; the widening of the density function shows the increase of the variability.

Within this example it is not expected one event happens. By definition we expect 0.26 events to happen (equal to the sum of all probabilities). The probabilities of all events should sum to 1 and it is expected one event will happen during construction.

[1] Both functions have mean and standard deviation equal 0 and 1

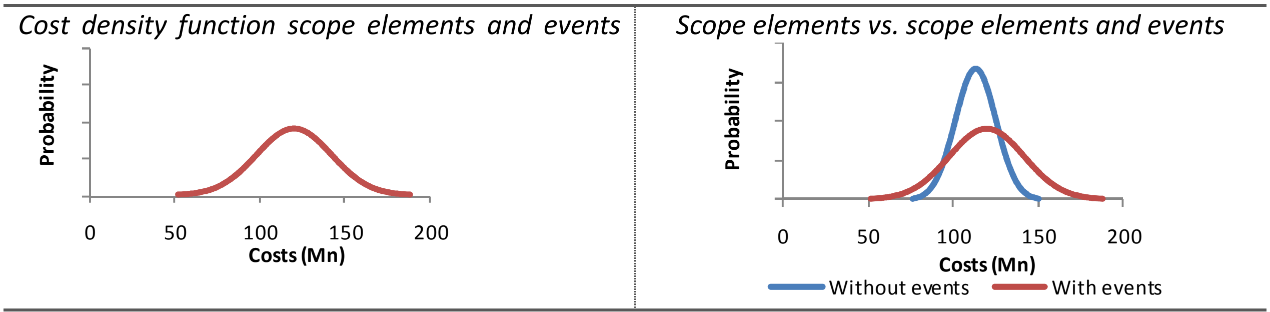


Figure 16; Cost density function including events

The cost density function of the scope elements is significantly different from the deterministic estimate. The cost estimate was set at USD 136.3mn between USD 81.8mn and USD 191.2mn.

The standard normal method estimates the scope elements USD 113.5mn. The 80% confidence interval is bounded between USD 104mn and USD 123mn. The project estimate including the events is USD 120mn. The final costs are expected with an 80% probability between USD 100mn and 140mn.

3.3 Level III normally distributed functions

The level II method is a quick method to show variability of cost estimates; only the mean and variability have to be estimated. This creates a starting point for risk management and presents the costs probabilistically. The behavior of variables is approximated taking mean and standard deviation of all elements and events into account and assuming normality. The method makes it possible to come to an analytical solution without the use of a computer.

Fitting variables to the normal distribution is applicable to continuous (i.e. scope elements) distribution functions. Applying the methodology to discontinuous functions (i.e. events) introduces errors. A sign of the incorrect approximation is shown in the example as the standard deviation of the events is larger than the expected value. This means the costs can be negative.

Figure 17 shows the true probability density function and the approximated density function of a exceptional event. This approximation is inaccurate. The tails are incorrectly incorporated. The upper limit of the modeled cost density function is approximately USD 50mn; the upper limit of the true density function is approximately USD 75mn. The fitted distribution function can produce negative outcomes and outcomes in the discontinuous section (i.e. between no event and the consequence). This does not reflect reality.

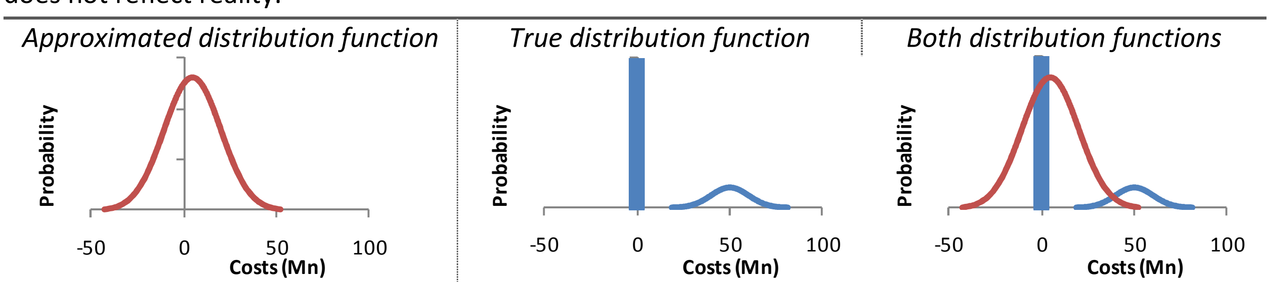


Figure 17; True vs. modeled distribution function

The level II approach is useful to approximate construction costs roughly. One is not interested in the lower tails or limits (cost decline will never be an issue). Incorporating multiple exceptional events reduces the error. This makes it likely to conclude the approach is useful as a first insight.

Taking this discontinuity into account, the solution can easily be obtained by simulation. It should be remarked the discontinuity does not make it possible to determine the analytical solution. Figure 18 shows the simulation of the level II and level III estimate of the example estimate;

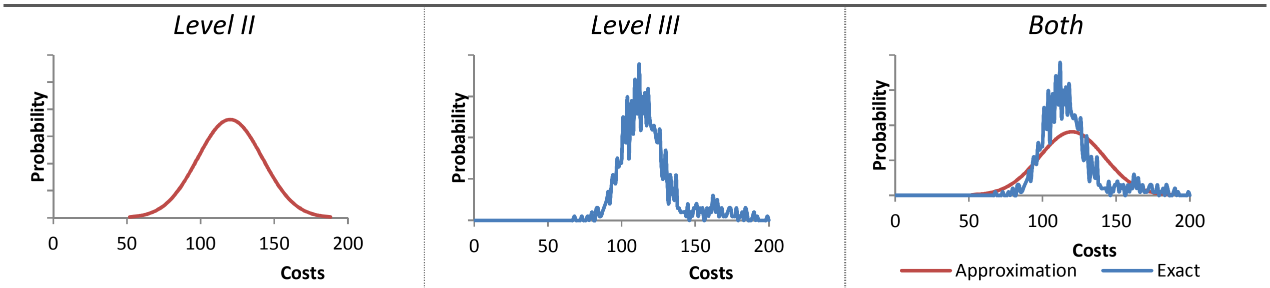


Figure 18; True vs. modeled distribution function

This step introduced the advantage of simulation. A more accurate result can be obtained. Elaborating this approach, the output may become more realistic and accurate using multiple distribution functions. Not all elements or events behave normally distributed. Incorporating various distribution functions may therefore present a more accurate result.

3.4 Level III various functions

Having the opportunity to assign various probability density functions to the estimate may provide additional power. The flexibility may result in a precise estimate of the cost density function. The expert now has the opportunity to estimate the expected value, confidence interval and probability distribution. This creates a flexible tool. The flexibility makes experts assume this method is a flexible tool. One may question whether the flexibility leads to more realistic results or not.

I selected six different probability density functions showing the likelihood of the costs. These functions and the expert opinion are given;

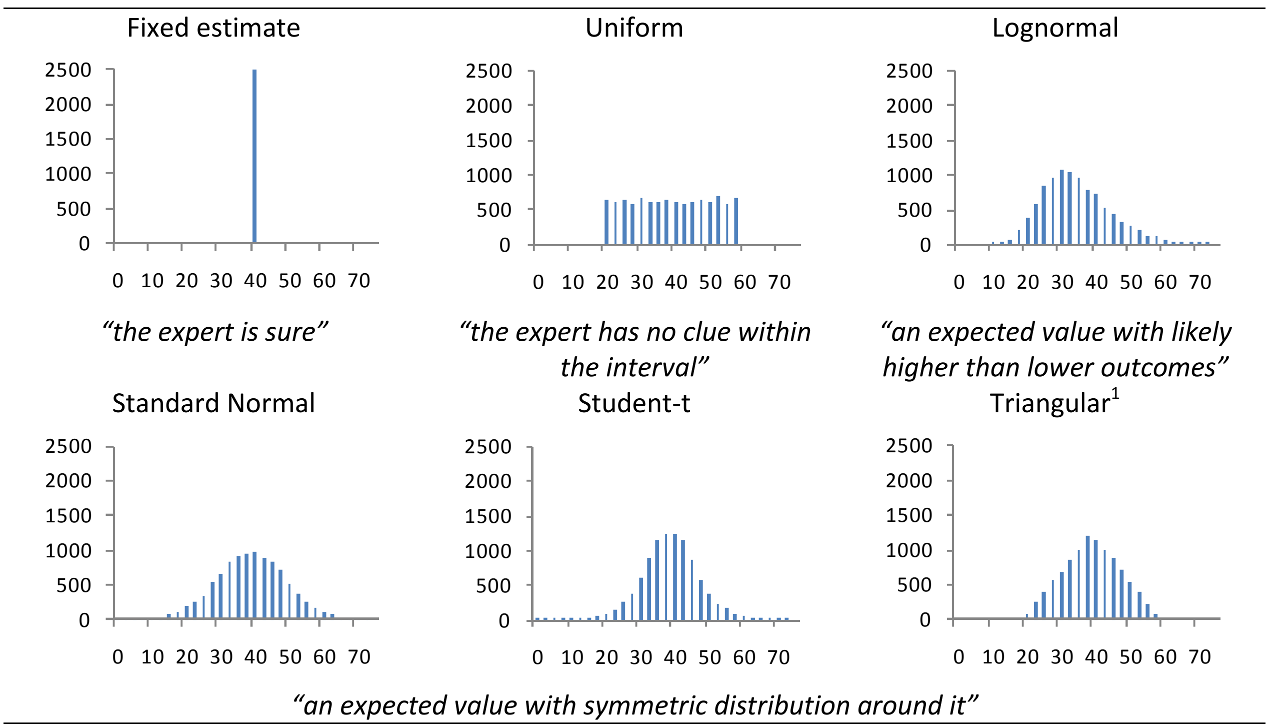


Figure 19; Cost distribution functions, Note (1): A triangular function can also be asymmetric

The Monte Carlo tool exactly describes the behavior of elements. Unfortunately the method is not the perfect estimation tool. The method easily introduces modeling errors;

- Deviations between the estimated most likely value and the expected value
- An incorrect fit of distribution function to the expert estimate
- No incorporation of possible estimation errors

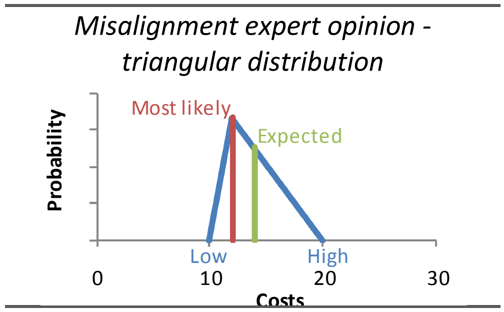


Figure 20; Triangular distribution

Experts often assume to estimate the expected or average of the distribution function. In fact, the experts estimate the most likely value. This value may differ from the expected value. An expert may assign an asymmetric triangular distribution function to a scope element. The expert estimates the low, medium and high value; for instance 10, 12 and 20. The distribution is show in Figure 20. The medium value shows the most likely value as the value is assigned the largest probability. This is not the expected value of the probability density function as the mean is $\left(\frac{10+12+20}{3}=\right)14$. Experts tend to estimate an asymmetric distribution function; the price is more likely to increase than to decrease. This leads to a difference between the expert estimate (sum of the most likely values) and the expected value of the simulated distribution function.

Secondly it may be possible the estimated confidence interval and estimated mean do not correspond with the assumed distribution function. Many distribution functions have a fixed shape. A fixed shape means a fixed ratio between the mean and standard deviation. The distribution function is fitted to the expert estimate, but will not exactly reflect the expert estimate. This fit does introduce an error. Within this thesis the project density function is fitted to the estimated confidence interval. This introduces a difference between the estimated and simulated expected value.

Thirdly, the project density function may be robust to expert errors. An expert may provide an incorrect prediction of the costs; i.e. the outcome deviates from the provided estimate. A probability density function with a bounded confidence interval does not take possible estimation errors into account. A bounded confidence interval restricts the simulation to solely present the estimate. An example of a density function with bounded confidence interval is the triangular function. The surface under the triangular is 1. No deviations from the estimation are possible. A standard normal distribution function has infinite boundaries. The expert estimates within a specified (e.g. 80%) confidence interval. If an 80% confidence interval is estimated, 20% of the simulated output is different from the estimate. Unexpected outcomes are possible making the model more robust.

The Monte Carlo simulation method may provide accurate results with its flexibility. But, the estimation tool may also introduce estimation errors. The expert should take the possible inconsistencies into account.

The Monte Carlo method does not take interdependency into account like the standard normal tool. As dependency is not incorporated, the cost items are valued stand alone. The independent behavior makes cost differences level out; cost increase of an item is backed by cost decrease of another item. The stand alone approach reduces variability. The modeled output presents unrealistic low variability

and small confidence interval, not in line with expert opinion. Dependency should be account for to reflect variability in line with expert judgment.
Physical relations show the presence of dependency; similar construction materials are used among different scope elements. If the prices of the materials change, the prices of the different elements will change as well. Several variables may mutually influence different cost items. The next section will introduce a model to incorporate interdependency.

3.5 Level III interdependency

The previously described tools do not incorporate dependency. A relationship between cost elements is present as corresponding variables drive the costs of the elements. Experts have clarified the price can be decomposed to different variables. As each different element is driven by the same subset of variables interdependency is introduced. The experts also use industry guidelines (13) to support the cost estimate. Within these guidelines the costs of various elements are also decomposed to a subset of drivers. Within this paragraph an analysis of the drivers is explained. This information is used to construct a new estimation tool.

3.5.1 Analysis of the different drivers

The introduction mentioned a subset of drivers that contributed to the price of elements. In general the different common cost drivers are;

1. Terminal requirements and local technical conditions
2. Construction methodology
3. Contractor market conditions and type of construction materials

The structural design of a terminal is determined by its requirements and the local technical conditions. As an example the type of subsoil is taken. The forces on container terminals are transferred from the structure to the subsoil. As a result weak soil will require a different design than the situation with strong subsoil. In the end the entire terminal will be built on the same site. The terminal components are designed to the same soil conditions. This makes soil a common cost driver.
The structural design is designed to the soil conditions. Different soil conditions lead to a different structural design. Based on this observation it is concluded one cannot express the soil conditions as a percentage of the costs. One should identify different soil conditions and come to a concept design for each condition. The costs of different designs should be estimated as well as the probability of the associated soil conditions. The result is accordingly "design uncertainty" classification of Vrijling in Chapter 2. Other (similar) examples relate to the design vessel mooring at the terminal, technical quality of the terminal or forecasted throughput of the terminal. These fundamental changes of requirements and different conditions often cause significant changes. For example, a larger design vessel or expected throughput may lead to a more complicated (i.e. expensive) terminal. This leads to the conclusion changes of these fundamental assumptions lead to great cost variability.

Having finalized the design contractors are approached and asked to state a price of the construction costs. While determining their bid price, contractors think of a construction methodology and the use of equipment. The methodology determines the use of equipment and the manning. This makes it an important driver of construction costs. Mostly general practice and experience lead to a common method. But, many examples are known of unconventional construction methods driving lower prices.

The market conditions and investment appetite influence the pricing. Contractors aiming to enter a new market will be more willing to acquire a project and lower the prices. In contrast; many projects in execution lead to a heated market and substantial price increases. One may conclude the balance between supply and demand is reflected in the scope elements. As only one contractor wins the tender, the market conditions are a common driver within the terminal elements.
The contractor premium expresses the market conditions and investment appetite. This premium is usually known with the experts. This makes it possible to reflect the market conditions within the pricing.
Each terminal component is constructed using similar materials. The components are constructed out of concrete and steel. People and oil are used to operate the machinery. The materials account for the costs of each element. The use of materials and prices are known with the experts.
Above that, industry tools (13) that are used to estimate decompose the costs. The costs of various elements are decomposed to various drivers. This makes it possible to estimate the contribution of various drivers to the price.
Variability caused by market conditions and type of construction materials generally has a smaller impact on the costs with respect to requirements, local technical conditions and construction methodology. But, it is possible that increasing prices may lead to other designs. Other designs may become more attractive because of a strong cost increase of commodities. But, the allowance for other designs is still dependent on local conditions and terminal requirements.

Based on the above analysis two conclusions are drawn. First a triangular graph is constructed. This graph is shown in Figure 21. The three different drivers have been transformed to layers. The width of the layers indicates the drivers' contribution to the cost variability. Above that, the layers are placed in chronological order. Local conditions and requirements are important in the first stage of developing a new terminal; later on the market conditions and construction materials become important.

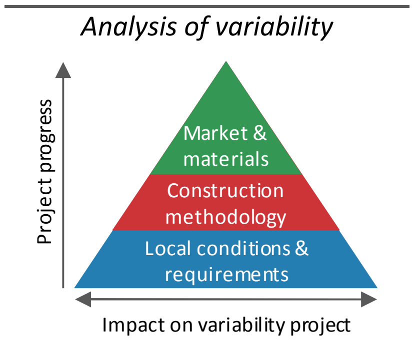


Figure 21; Variability

Secondly the analysis allows describing the costs in mathematical terms. The local technical conditions and terminal requirements result in various possible designs. The construction methodology, construction market and construction materials directly contribute to the price. An example of an element's price can therefore be expressed with;

$$C_{element}|soil = \left(C_{methodology} + C_{market} + C_{materials} + C_{other}\right)|requirements$$

$C_{element}|requirements$ represent the cost of an element conditioned on the requirements and local conditions, $C_{methodology}$, C_{market} and $C_{materials}$ are the contribution of the construction methodology, market conditions and materials to the costs and C_{other} reflect the unidentified part of the construction costs. This decomposition introduces joint behavior that we measure with correlation. It is easy to identify the correlation by generalizing the costs and cost decomposition to $y_1 = \alpha_1 + \beta_1 x_1 + \varepsilon_1$and $y_2 = \alpha_2 + \beta_2 x_1 + \varepsilon_2$. The covariance and correlation are;

$$cov(y_1, y_2) = cov(\beta_1 x_1, \beta_2 x_1)$$
$$cov(y_1, y_2) = \beta_1 * \beta_2 * var(x_1)$$
$$\rho(y_1, y_2) = \frac{\beta_1 * \beta_2 * \sigma_{x_1}^2}{\sigma_{y_1}\sigma_{y_2}}$$

Within this thesis assumptions are made on the behavior of the different drivers. The different designs (originating from different local conditions and requirements) should be estimated separately; the other drivers are described by time series. The time series are derived from the US department of labor statistics[2]. The time lag from mid 2007 to mid 2010 presents the behavior. The variability and correlation are computed and the time series is assumed normally distributed. It is assumed the behavior is applicable to the case study.

Within this thesis joint behavior between the various drivers is taken into account. Neglecting this joint behavior may lead to inaccurate results. A simple example points this out. Oil may be described by the spot prices of Gasoil (Reuter's data spot prices); steel may be described by the spot prices of Mediterranean Steel (London Metals Exchange spot prices). Both selected time series start January 1st to May 28th 2010. The variability of oil and steel are 14% respectively 7%. Additionally, both time series show a not negligible interdependency; the correlation (i.e. ρ) is 75%.

This rough analysis may present an incorrect presentation of correlation. A drift term in both time series may lead to spurious correlation. To take out the drift, return series are analyzed. The return series of both series show a joint behavior of 30%. To clarification the analysis, both time series are presented in Figure 22.

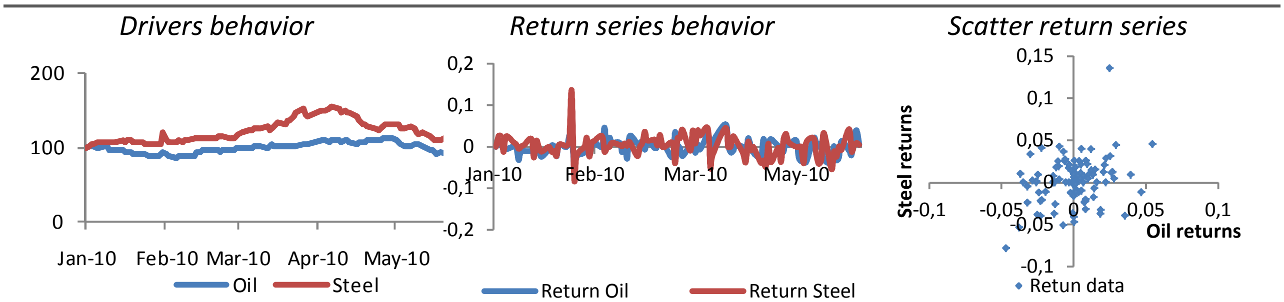


Figure 22; Oil and steel behavior

A presentation of dependency caused by the construction methodology, common construction materials and market conditions is shown in Figure 10 (Chapter 2). The oval clouds show the dependency between different scope elements. Changing costs, caused by different local conditions and requirements show a different behavior. Of two elements, a container yard and a wharf, we can simulate the costs. The behavior is assumed independent (i.e. no influence of , $C_{methodology}$, C_{market} or $C_{materials}$). But for both elements a different structural design may be applied. This design leads to different costs of both scope elements. Of course, the cost change differs between the two scope elements. The scatter plot of both elements is shown in Figure 23. The left graph show the costs if the first design will be constructed; the middle graph shows the costs if the other design will be built. The elements behave independently in both situations. The right graph shows the combined behavior of the

[2] The Oil time series is based on the Reuters' spot gasoil prices from mid 2007 to mid 2010.

scope elements. The change of costs clearly establishes a link between the two scope elements. A straight line can be drawn through the two clouds. This proves the various designs contribute to dependency.

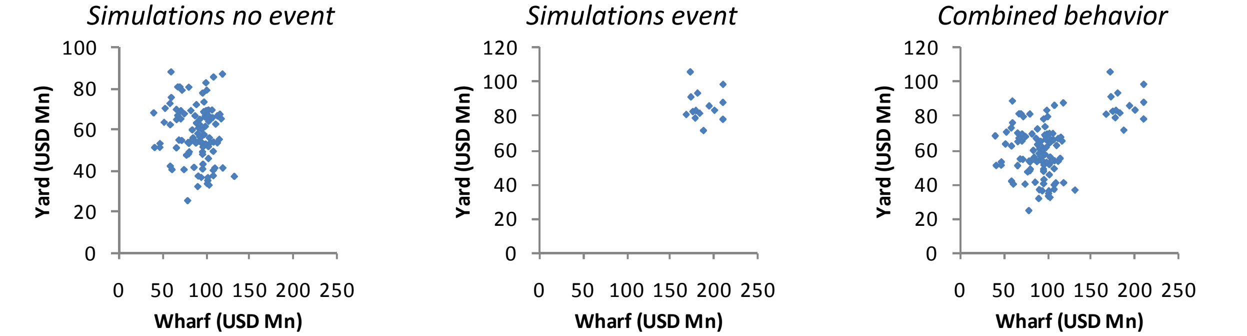


Figure 23; Scatter plot event

3.5.2 *Application of the model including dependency*

The preceding section presented the pathway to determine the existence of a subset of drivers contributing to the costs. To illustrate the use of this approach, an example is provided. This example of the price decomposition is presented in Figure 24. The expert estimates the percentile contribution of drivers to the unit rate price; P_i is decomposed. Quantity Q_i is determined by the structural design making decomposition not possible. In Figure 24 an example of decomposed elements is presented. The identified scope is equal to the earlier presented example. Different designs are not incorporated within this example.

Decomposition scope elements

			Cost drivers			
Scope	*Oil*	*Steel*	*Concrete*	*Labor*	*Contractor*	*Other*
Engineering	0%	0%	0%	40%	15%	45%
Breakwater	10%	0%	15%	10%	10%	55%
Wharf	5%	25%	10%	15%	10%	35%
Dredging	40%	0%	0%	10%	10%	40%
Yard	10%	5%	10%	20%	10%	45%
Admin Building	5%	10%	10%	10%	10%	55%

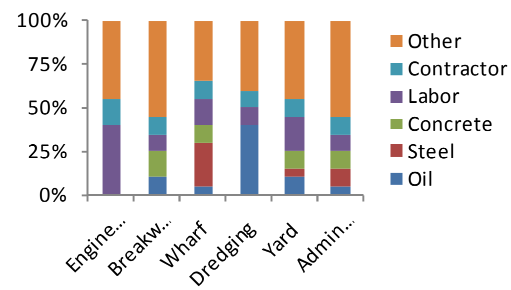


Figure 24; Decomposition scope elements

The project density function resulting from the analysis is shown below. It is shown clearly the additional effort does not provide any difference with the independent cost density function.

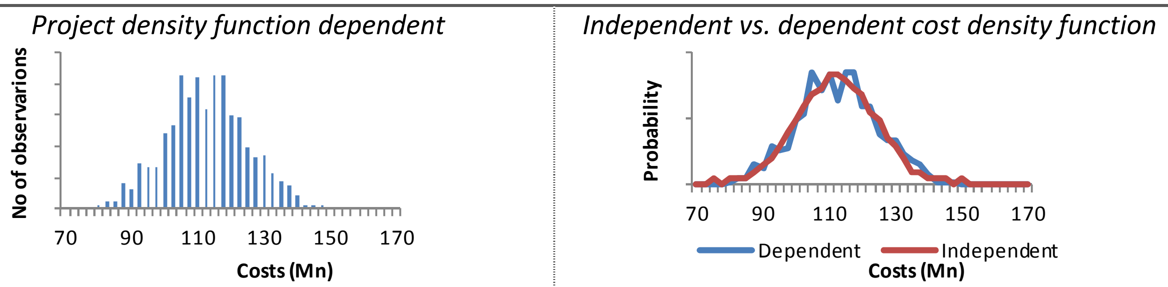


Figure 25; Project density function with component analysis

The answer to the reason of this minor difference is rather straightforward. In Chapter 2 it was explained the standard deviation of the total project density function consists of two orthogonal components; the identified and error term. If the error term is larger than the identified term, the dependency does not provide additional power. Dependency only plays a role if the standard deviation of the identified term contributes over 70% of the total standard deviation. This is because the squared term accounts for the total variability. Then, the correlation between the total and identified part is 70%; $\rho = \frac{\sigma_{identified}}{\sigma_{total}}$. The error term accounts for less than 30%. The squared contribution is measured by R^2. The R^2 measures the identified variance vs. the error variance. If the correlation equals 70%, the standard deviation of the identified part equals the error term. The relation between the correlation, R^2 and indentified versus error standard deviation is shown in Figure 26.

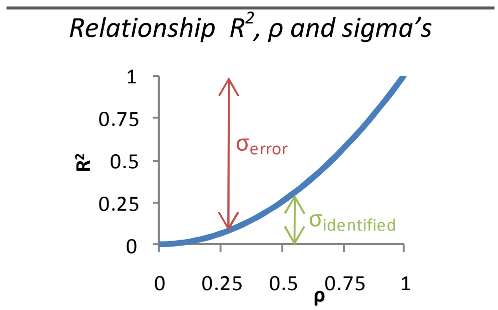


Figure 26; Relationship R^2 and ρ

This fifth paragraph proposed an approach to decompose scope elements and incorporate the behavior of the elements. A subset of drivers is identified and described. This example shows the additional effort may not provide significant different output. If the correlation between the identified drivers and the scope elements is below 70%, the variability of the project density function is hardly influenced. Therefore a present day estimate decomposing the price may not provide additional value. The project density function may not differ between independent and dependent scope elements.

Even though a direct improvement may not be provided, the approach of decomposed prices may provide additional value. By forecasting the subset of drivers it is possible to predict future costs more accurate. The inclusion of time is beyond the scope of this thesis subject. The future development of the estimated costs will not be analyzed. But, in emerging markets cost increase may develop rapid. To account for this development shocks of the drivers are incorporated. The variability of the explained part may also increase and the result may be significant dependency.

3.6 Level III interdependency and shocks

Boschloo and Vrijling have pointed out the importance of exceptional events. As events happen at every construction site, these events should be incorporated within estimates. The development of container terminals may take considerable time. The design and tender process take on average two years. In the meanwhile cost increase may play a substantial role. For example, the growth of emerging markets may result in heated markets and construction prices.

To account for this price increase shocks of the different drivers are introduced. Shocks will only be incorporated within the drivers described by time series. These shocks drive additional variability. As a result, the shocks increase explained variance and dependency may become important. Figure 27 shows the increase of variability and R^2.

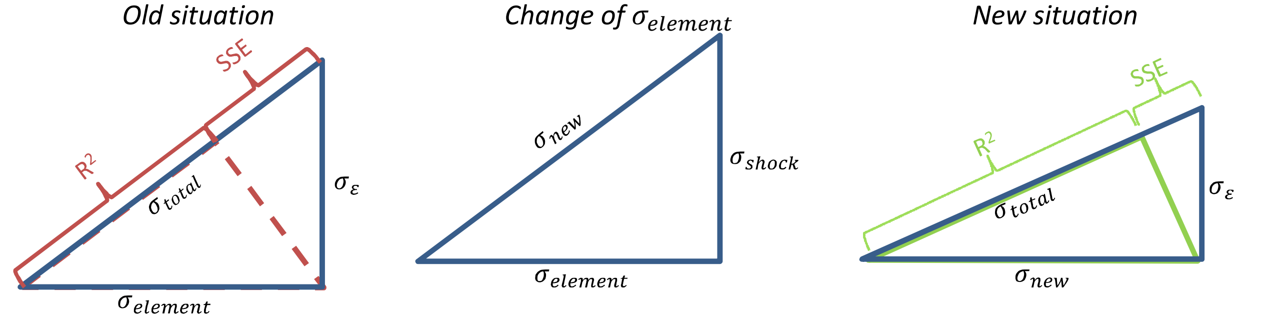


Figure 27; Explanation of R^2 and ρ

3.7 Conclusion

This chapter has presented six different tools to estimate the construction costs;

1. Level I deterministic method
2. Level II normally distributed method
3. Level III normally distributed functions
4. Level III various functions
5. Level III dependency
6. Level III dependency and shocks of drivers

The deterministic tool is used within APM Terminals at present. This method is quick, but dirty. The quantities and price ratios of all the elements are estimated roughly. A contingency factor is applied to account for lack of information. The contingency factor is determined on the project stage without specific information on the project uncertainty and risk.

The standard normally distributed tool fits all elements to the normal distribution. The mean and standard deviation of the different elements should be computed to apply the tool. Assuming all different variables to be normally distributed is justified by the Central Limit Theorem. CLT assumes that the sum of many density functions converges to the standard normal distribution. Also, the individual behavior of variables is often well approximated with the standard normal distribution.

Fitting the mean and standard deviation of an event to the standard normal distribution function provides a rough idea of the variability. But the approximation provides inaccurate results applied to discontinued functions. Using numerical simulation, a more accurate solution is provided.

The level III various functions approach allows all elements to vary differently. The flexibility of incorporating multiple functions should provide an accurate result. This accurate result is obtained by simulating the various individual elements and events. The tool is very flexible, but also prone to inconsistencies. Deviations between the estimated and the simulated expected value may be present. The fit of various distribution functions to the estimate introduce additional errors. Some distribution functions may not account for possible estimation errors.

The described models do not take dependency between the various scope elements into account. However, costs are driven by mutual drivers causing interdependency. Within section 3.5 an approach is proposed to decompose the scope elements. Drivers should be identified driving the elements' costs. These different drivers are fully dependent among the scope elements. This makes it possible to incorporate dependency. This report approaches the different drivers two-fold. Changing fundamental assumptions (i.e. requirements and technical conditions) leads to different designs. Other drivers can be described by historical behavior. The example showed the incorporation of variability show a different output from the independent example. This is caused by the contribution of the explained variance to the total variance (measured by R^2).

The use of constant variability assumes constant demand and supply. Up and downward shocks (based on expert judgment) incorporate the effect of changes in demand and supply. While developing new terminals, changing prices may play an important role. Local pricing may increase substantially in emerging countries.

Therefore shocks are incorporated within the price development of the drivers. This will not only provide a more realistic view, but also increase the explained variance. This may have an impact on the application of dependency and increase the variability of the total project density function.

Before applying the different tools, the APMT requirements of a model are listed. This makes a selection of the different models possible. The requirements are described in the next chapter. The performance and selection of the tools is presented in chapter 6.

4 The APM Terminals' model

This thesis focuses on the development of a new model to estimate civil CAPEX and present the uncertainty of this civil CAPEX estimate within APM Terminals. A model presenting a realistic output is obliged within the process to estimate the construction costs. Estimating an inaccurate or meaningless number does not add any value. Above that, the model should also be used in practice. The model input should be aligned to the available information, the time spent and the engineers' capabilities.

To come to a practical model, the APM Terminals's requirements have been documented. During various sessions with the civil engineers, business development managers (responsible for new terminal development) and finance managers (holding the financial model and collecting all financial data) a lot of information was acquired. This section presents the information. I will start with an outline of the civil engineers, their capabilities and playing field. Thereafter the particularities of the different estimate classes are presented. The analysis will provide a solid foundation to determine the practical requirements. These requirements are verified with the various estimation tools to determine the most suitable estimation tool.

4.1 APM Terminals civil engineers

The APM Terminals' civil engineers are senior project managers. The engineers are very experienced in managing and/or supervising large infrastructural projects. They are well familiar with the practices on the construction site. This experience provides a vast amount of expertise on project costs and market contacts to verify expectations and judgment. The contractor market may not always be transparent and market intelligence is important to benchmark the expectations.

The engineering background does put its limits on other skills; one cannot know all. It is of their disadvantage within the cost estimation process that they have limited familiarity with statistics. This may introduce (significant) estimation errors. To prevent these errors, the objective is to clarify the estimation input of the model[3]. In general, it is known the engineers estimate price and quantity differently. Independency between unit rate and quantity is assumed; the two variables are estimated separately with different stochastic behavior.

Additionally, the engineers should know the implications of the answer they provide; the confidence interval and its boundary values as well as the expected or most likely value. Knowing these assumptions provides experts clarity on what they estimate and the result of their estimate. This will eventually help the expert to come to an accurate estimate and correct procession of the information.

The practical experience from the construction site also introduces limits on software or programming skills. There is limited experience with software like Microsoft Excel or Microsoft PowerPoint. Mathematical software like Matlab may not add value within the process to estimate the civil CAPEX. This may require software programming skills. The aim is to develop an easily usable and robust model without complicated software programming and easy access.

A new business opportunity is split to several stages. During the process to develop a new terminal, the APM Terminals' civil engineers provide multiple estimates. The costs of civil infrastructure are estimated each time the project progresses to a new stage.

[3] More information on expert elicitation can be found in the article by Tversky and Kahenman (17)

In the early project stages the engineers estimate the civil costs themselves; in a later stage they supervise an external consultant who provides the estimate. The civil engineers provide a class 5 and a class 4 estimate. These estimates and the conditions to provide an estimate are treated in detail next.

4.2 The class 5 estimate

The class 5 estimate is provided in the first stage of a project. This first stage is called a probe phase as shown in Figure 1 (Chapter 1). The estimate is used as a (quick) feasibility check before analyzing the project in more detail. The estimate is described as the "back of an envelope" or "order of magnitude"; the estimate is more of a sanity check before further analyzing the project.

As the project is in an early phase of the project, many input variables are prone to uncertainty. The opportunity is often the result of a general market survey. Little specific information is known. This introduces uncertainty on the design conditions (i.e. design vessel, throughput forecast, etc.). An accurate estimate can only be obtained developing multiple scenarios. This requires a lot of time which is not meaningful at this stage.

It is concluded limited time is available to roughly determine project feasibility. Internal guidelines show the engineers spend at maximum 4 days. An in depth research would require a lot of time which is not meaningful at this stage. In contrast, the engineer may have an idea of the structural design, but this may even not be the case. If the engineer spends 4 hours, the provided estimate will be very rough. After 4 days of work the engineer will have a significant better idea of the structural design and the project.

Based on the preceding description it is shown an estimate is provided using little information. One may even not have a clear idea of the business case. The business case, terminal (structural) design and costs are prone to changes resulting in large cost variability. Therefore the class 5 estimate is a very rough indication of the civil construction costs.

4.3 The Class 4 estimate

After an interesting project has been identified, a project team starts to elaborate the business case. The team will develop a solid investment proposal based on a detailed design, financial outlook and risk assessment. This information determines the decision to investment and is taken by the Maersk board.

The civil cost estimate is an important driver within the financial outlook. The details are already explained in chapter 2. Two cost estimates are presented within this stage; a class 4 and a class 3 estimate. Early on the civil engineers provide a class 4 estimate; in a later stage an external consultant (under supervision of the civil engineers) provides a class 3 estimate.

Providing an estimate, the engineer should have an idea of the local conditions. Therefore, the expert visits the site to become familiar with the local conditions. The gained information is combined with expert experience and judgment. This process determines a first structural design. Possibly, several meetings with contractors are held to become familiar with the local market, available materials and pricing.

The engineer focuses on clarification of the local conditions and structural design. This is the basis to estimate the scope of works and ask contractors to issue a price. Little attention is being paid to deviations from the assumptions or occurrence of exceptional events; the objective is to clarify and fix

the design or any other risk exposure. Thereafter a qualitative risk analysis is provided describing the main risk drivers without quantification.
The described process already makes clear a considerable amount of time is spent to the estimate; the engineers spend 4 days to 4 weeks to provide a class 4 estimate. The amount of time is determined by the scope of works, the required level of detail and required estimate accuracy.
All in all, the engineers provide an estimate based on a first inquiry on the terminal area, the scope of works and local market conditions. There is an outline of the structural design and deviations are driven by new information on the site and local market developments.

4.4 Model requirements

The preceding three paragraphs have characterized the civil engineers and the different estimates. The engineers' capabilities and available information and time were addressed. From these characteristics the requirements are distilled. The requirements will be used as evaluation criteria to recommend the proposed tools (i.e. deterministic, normal, Monte Carlo and interdependent tool).
The model requirements are grouped under;

- Input information
- Output figures
- Use in practice

4.4.1 Input information

The engineers have a certain amount of information to their possession while producing an estimate. This information should be maximally used to come to a correct estimate of the construction costs. In general, it should be possible to the engineers to incorporate;

- Scope of works
- Unit rates and quantities
- Engineering judgment

To come to an estimate the expert should be able to incorporate the available information. A (brief) insight on the scope of works is present. The expert estimates the unit rates and quantities of the identified scope.
Based on the engineering judgment the engineer determines the variability. In fact, the engineering judgment is the experience based tradeoff between market conditions, structural design and historical data. The expert judgment includes the incorporation of exceptional events as exceptional events present the possible deviations from the design assumptions.
In general the engineering judgment is determined by the available amount of information. Therefore a small summary of the available information of the two classes is shown below.
The class 5 estimate is characterized by little information. There is a general idea on the business case, but all the different variables are heavily prone to changes. The engineer determines his cost estimate on several assumptions which may not be applicable and estimates the scope of works as an order of magnitude. An engineer is able to identify a rough outline of the scope of works with limited insight on possible events.

To come to a class 4 estimate additional effort and time are spent. The engineer visits the site to become familiar with the terminal area and other local conditions. This is the basis to the outline of the structural design and scope of works.
A general overview of the main risks is provided within a (qualitative) risk report. Based on this analysis it is possible to determine any exceptional events.

4.4.2 *Output figures*

The resulting project density function of the estimation process is verified with internal APM Terminals' documents(2). These documents prescribe the civil CAPEX estimate (including contingency) is the expected value of the project density function (i.e. P50).
The estimate is bounded by an 80% confidence interval. (The outcome will be within the confidence interval with 80% probability.) The boundaries (i.e. P10 and P90) are expressed as a percentage of the (deterministic) point estimate. The class 5 confidence interval is set at +/- 40% of the estimate; the class 4 interval is set at +/- 20%.

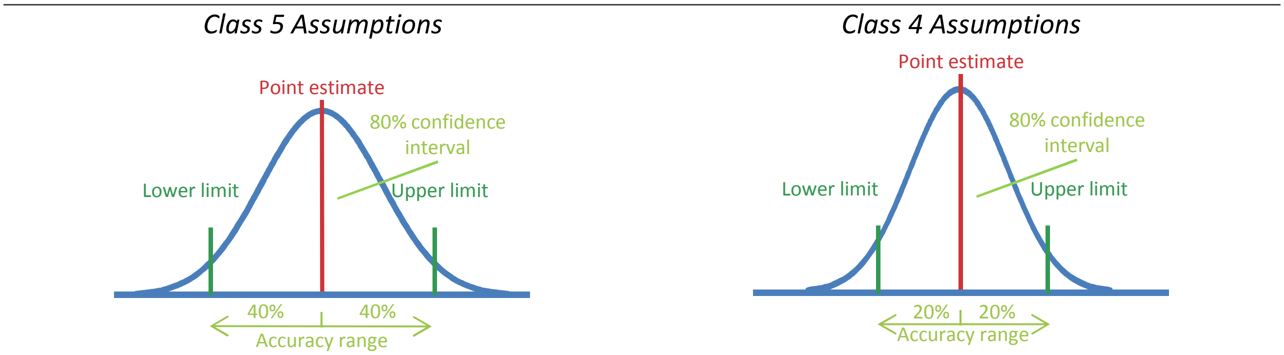


Figure 28; Estimation assumptions

4.4.3 *Use in practice*

The probabilistic estimation tool should be used by the engineers within daily practice. Using the model, the engineers should clearly present the estimate and its inherent uncertainty. The estimate itself should be derived from the model as well as the involved uncertainty.
The AP Moller Maersk group approach is to keep risks simple and understandable to all involved. The group is predominantly interested in the probability of cost overrun and its consequence. The model should therefore clearly show the provided estimated. The result of the incorporated safety factors (i.e. contingencies) should be presented as an upper limit of the bounded cumulative confidence interval. The remaining risk involved with the estimate should be presented with the expected value in case of cost overrun (i.e. expected shortfall). The uncertainty can additionally be presented with the graph of the cost density function.
The output of the estimate should be aligned to the financial model. Within the financial model all the various cash flows are combined to determine the (expected) project profitability. The estimated civil CAPEX should be distilled into a Microsoft Excel sheet to be transferred into the financial model.
To distribute and verify the input with other parties, the model size should remain within 6Mb. Additionally, the model should be based on Microsoft Excel to fit best to the APMT practices.

4.5 Evaluation of the requirements

Based on the practical and performance requirements it is possible to select a tool to use. The different tools will be used to estimate the construction costs of a project. Thereafter it is possible to select the best tool per estimation class. The different tools are scored accordingly the selection criteria from 1 to 5; 1 is the worst result, 5 the best. The tool with the highest score will be proposed to APM Terminals.

An example of a scoring matrix is shown below;

Class 5 estimate	**Estimation tool I**	**Estimation tool II**	**Estimation tool III**	**...**
Input information				
Scope of work	...	...	...	...
Unit rates/quantities	...	...	...	...
Engineering judgment	...	...	...	...
Output figures				
Expected value	...	...	...	...
Confidence interval	...	...	...	...
Practical requirements				
Easy approach to statistics	...	...	...	...
Model size	...	...	...	...
Use of software	...	...	...	...
APMT template	...	...	...	...
Total	...	...	...	...

The above table is used to determine the best suitable tool. This selection is made by judging the tools on above criteria by using them to estimate the civil costs of an example project. In the next chapter a project in Moin, Costa Rica is introduced as an example case. In Chapter six the output results of the tools applied to this case will be presented and compared.

5 Terminal business case example; Moin – Costa Rica

APM Terminals is a worldwide container terminal operator; the company establishes new terminals all over the world. The terminal business case shows an opportunity in Moin, Costa Rica. An introduction of the project will be presented followed by an outlay of the structural design. The possible exceptional events follow from the introduction and the design outline. Interdependency between the scope and events will be explained. The analysis makes it possible to estimate the construction costs in Chapter 6.
The project is within the APM Terminals' development phase. Therefore the information is confidential and may not be distributed without written approval by APM Terminals.

5.1 Introduction to the business case

Costa Rica is situated in Central America between Panama and Nicaragua. Approximately 4.3 million people live in Costa Rica. Costa Rica is as large as the Netherlands, but has a more mountainous landscape with its highest point at 3810 meters above sea level.
The Costa Rican economy is characterized by cultivation of traditional agricultural products and tourism. The people are well educated due to government policy, but a poor state of the infrastructure limits economic growth.
Export of bananas, pineapples and coffee are the main drivers of the economy apart from tourism. Bananas and pineapples are transported by ship with refrigerated containers (reefers). The current capacity for transporting bananas and pineapples is inferior. Improving port infrastructure would stimulate export volume growth and drive growth of the Costa Rican economy. Therefore the opportunity of a new terminal in Costa Rica is being developed.
The terminal will be close to the city Moin. Moin is a large city at the Atlantic coast of Costa Rica. At present some terminal infrastructure is present. As mentioned, the state of this terminal infrastructure is insufficient to drive economic growth.

Costa Rica global positioning *Costa Rica country view* *Costa Rica terminal positioning*

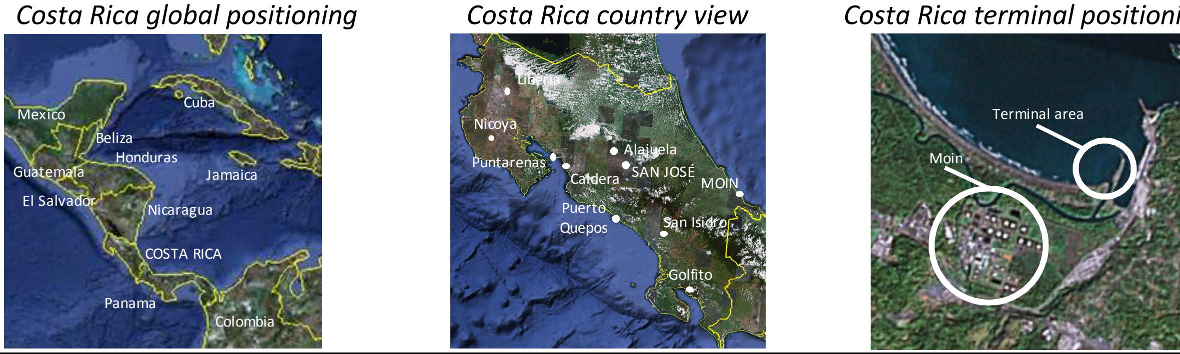


Figure 29; Costa Rica Outline

5.2 Moin design outline

The proposal is to construct a new container terminal from scratch; i.e. a green field project. Figure 29 shows the position of the terminal; in Figure 30 the terminal layout is presented.
Along the coast, a new area will be reclaimed to situate the container terminal. A sheltered area is chosen to prevent the new structure and moored ships from incoming waves. The selected terminal area is situated a few meters from the coast. A road will be constructed connecting the terminal area to the land.

An access channel will be dredged to provide container vessels a safe entry. The layout of the access channel is based on the main wave direction and the dimensions of the design vessel. The main wave direction determines the position of the access channel. Ships are difficult to control in case of perpendicular arriving waves. The dimensions of the channel are based on the size of the design vessel; the channel will be 200 meters wide and 16 meters deep. In front of the wharf additional space is cleared to ensure safe mooring of the ships.

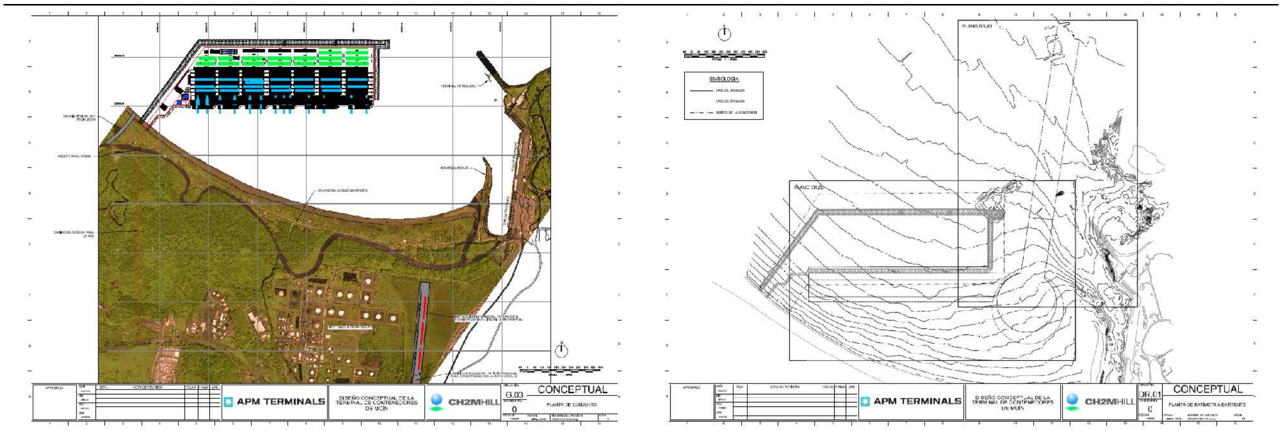


Figure 30; Costa Rica terminal overview

The terminal will consist of a wharf of 1,500 meters long, approximately 35 meters wide. In front of the wharf the sea bed is at 15 meters below sea level. The quay structure will be a deck on piles. Piles will be put into the ground. On top of the piles, a concrete plate will be constructed. The length of the piles depends on the strength (i.e. resistance) of the soil. The gap between the piles and the forces on the deck determine the thickness of the deck. In Figure 31 the cross section of the quay and the area in front of the quay are presented. The deck on piles is shown clearly as well as the mooring area for the container vessels.

Cross section wharf and mooring area

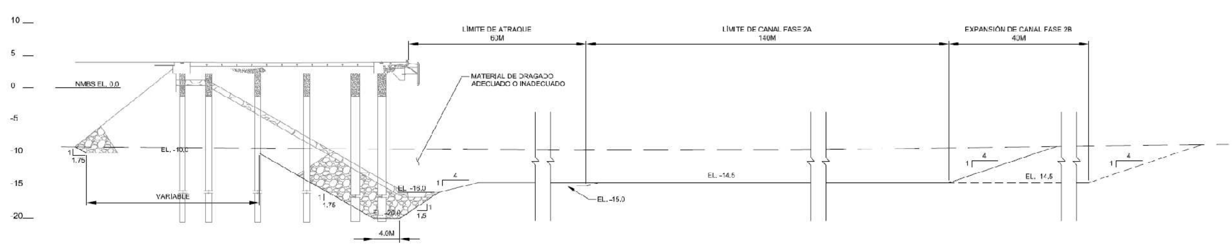


Figure 31; Cross section wharf/mooring area

On the wharf cranes and other vehicles are situated to unload the containers from the ship. The containers are stored at the container yard. To minimize the operating time, the container yard is directly situated behind the wharf. The yard will be circa 450 meters wide. In general, the dimensions of the wharf and the container yard are dependent on the dimensions of the arriving ships, the forecasted throughput and dwell time of the containers.

Around the yard a breakwater will be constructed to protect the reclaimed area against incoming waves. A cross section from the quay to the breakwater is presented in Figure 32;

Cross section terminal

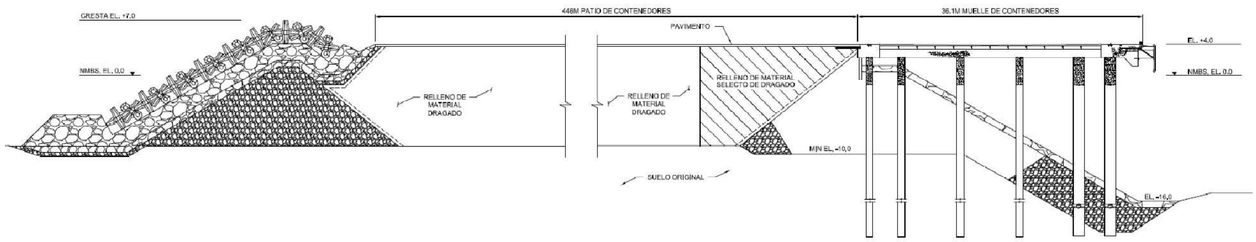


Figure 32; Cross section container terminal

From the cross sections it is clearly shown that different materials are used to construct the terminal. In general soil will be removed to 10 meters below sea level, at some places deeper dredging is required. On top of the dredged area new material will be placed. Under the breakwater and at the toe of the wharf rock core material will be used. These sections are subjected to large forces. The breakwater will be further strengthened with a revetment layer and a layer of concrete units. The slope under the quay will be reinforced with an additional revetment layer.

The earlier dredged material will be used to reclaim the container yard. This saves money as dumping soil is costly, as well as winning new soil. The principle of using dredged soil to reclaim the new area depends on the quality of the soil. The soil has to resist the loading due to the operations on the new terminal. Weak soil contains many fine particles. Many fine particles make drainage difficult resulting in a weak structure. The soil particles might compact insufficient to resist the loads.

The dredged soil is assumed to possess sufficient strength. The foundation layer should contain at maximum 30% of fine particles, not finer than 75 micro meters. The shaded area aside the wharf is subjected to larger loading. There, the foundation layer contains at maximum 10% of fine particles. Soil improvement will ensure safe silt content. The container yard will be paved with blockers on top of an under layer.

Apart from the main civil structures, other cost items should be taken into account to construct the terminal;

- Several buildings are constructed to operate the terminal. Examples are administration building, repairmen shop and power houses
- Basic provisions like a power grid, water supply and a communications network
- Additional infrastructure for the refrigerated containers
- General expenses;
    - Mobilization and demobilization costs
    - Temporary facilities
    - Consultant to supervise the contractor

Within the preceding a short summary of the Moin project is provided. The structural design and design assumptions are provided. The identified scope of work, the estimated costs and the contribution of the different scope elements to the estimate is shown in Figure 33.

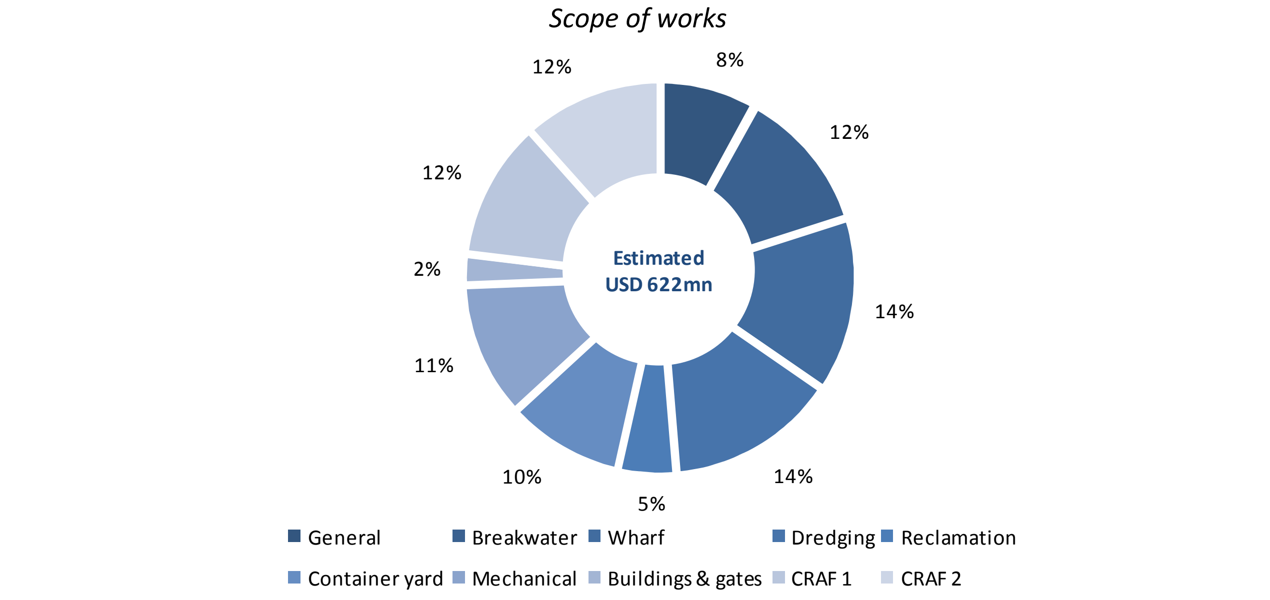


Figure 33; Scope of works

5.2.1 *Construction method*

The construction of the terminal is a large investment. Therefore it has been decided to build the terminal in three phases called 2A, 2B and 3. Figure 34 shows the different construction phases including the scope of works per phase. The terminal will be built to the right and the length of the planned wharf construction per phase is shown in red. In phase 2A 600 meters will be constructed and in 2B and 3 respectively 300 and 600 meters.

The cost estimate within this thesis does not take the different construction periods into account; a present day cost estimate is provided. In practice the APMT engineers provide a construction schedule and an escalation rate. The escalation rate is used to extrapolate the present day estimate to the future. Future research may extend the model to account for the time effect of construction cost estimates. The decomposition tool of Chapter 3 may be useful. The construction costs are determined by various drivers. Forecasting these drivers may provide accurate results to determine future construction costs.

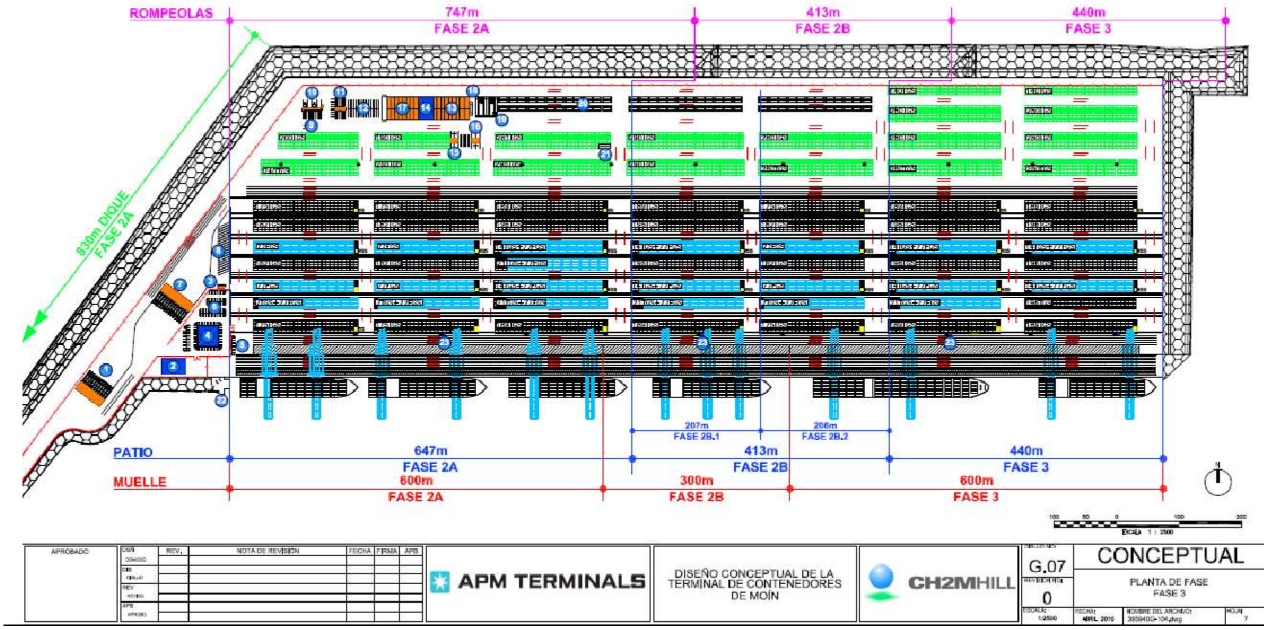


Figure 34; Terminal outline and construction phases

5.3 Possible exceptional events

The preliminary design is based on several assumptions. Without several assumptions one cannot come to a design. In the early stage the assumptions are rather rough. As more time and effort are spent, the assumptions will be further benchmarked to the real conditions. But, deviations from the assumptions are still possible. This introduces uncertainty and risk to the project. Additional costs are incurred to comply the terminal structure with the conditions and requirements.

At this (design) stage the risk of deviating assumptions is mainly caused by the uncertainty on the soil conditions. The strength of the soil directly influences the strength of the structure. Piles are put into the soil requiring a solid foundation. The soil under the reclaimed area as well as the reclaimed soil itself should able to resist the loading. Above that, it is assumed dredged soil can be used to reclaim the terminal area. This assumption puts requirements to the soil conditions. Weak soil may lead to a (too) weak foundation of the structure (and the structure itself) to resist the loading due to the operations.

One can imagine the presence of mutual dependence between the soil strength and the usefulness of the soil. The soil conditions influence both the strength and usefulness similarly. An increasing loading capacity of the soil is accompanied by improvement of soil quality to reclaim the area. This leads to positive correlation. Peat is very weak and not useful to reclaim. Sand has sufficient capacity to build the terminal and reclaim the area. The relationship is not infinitely sustainable. At some point the ground is that hard, unsuitable for reclamation. Rock is a (very) solid foundation, but its processing is expensive.

It is rather difficult to incorporate the exact relationship between the soil strength and the usefulness of the soil. First, it is difficult to quantify the relationship. Usefulness can be interpreted dependent on the business case. Secondly, the design will have certain robustness. The design can cope with different soil properties like strength and composition. The robustness is quite important as homogeneous soil conditions will not be present. To overcome these difficulties and still have a realistic estimate, the soil conditions are modeled discretely. Multiple scenarios show the result of the associated soil conditions. In this case three different scenarios are identified;

1. A 15% probability is estimated finding soft soil. The soil cannot be used to reclaim the area. This leads to;
    a. Dumping of dredged fill leading to 50% increase in operational costs
    b. Winning of soil to reclaim the area is priced 15 USD/m^3
    c. Reduction of soil improvement costs at the container yard with 50%
    d. Increasing soil preparation costs by 20% to ensure sufficient strength of the subsoil
    e. 25% longer piles under the deck of the wharf; the construction and material costs increase by 30%.
2. The soil conditions are accordingly the assumed conditions; 70% probability
3. There is a 15% chance finding hard soil. This soil is very useful to reclaim the area;
    a. The dredging costs increase to 18 USD/m^3
    b. The piles are sufficiently long, but the piling costs increase with 15%

The impact of the soil conditions is clarified in Figure 35 at the right. Not only the strength (or type of) soil introduces risk on the costs. Two other events that may occur are introduced.

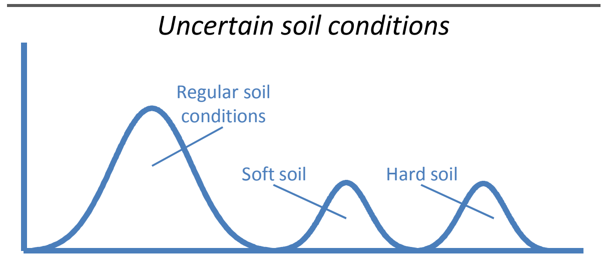


Figure 35; Uncertain soil conditions

First, there is an additional risk of discovering contaminated soil with 1% chance. If polluted soil is found, 5% of the dredged soil is assumed to be polluted. Removal and dumping the soil will cost 50 USD/m^3. On top of that, additional site supervisory is required and the delay results in a contractor claim. Extra supervisory leads to 25% increasing costs. The contractor claim will probably be USD 5mn.
Secondly, wave conditions may be different from expected with a 1% probability. These higher waves lead to 25% increased costs of the breakwater. A larger breakwater is required as well as heavier revetment to resist the forces. Within this stage the identified events comprehend the majority of risks.
A last note to the estimates is necessary. One should remark the probabilities of the different soil conditions and the other two events do not sum to 1. Less than one event is expected to happen. If one expects one event to happen, the probabilities of occurrence should be raised or other events should be thought of. No information on historical performance of projects is known. This shows the need to do an additional survey on historical performance to benchmark current estimates.

5.4 Conclusion

This chapter has introduced the Moin project. The design has been clarified and the scope of works was introduced. Several possible events are shown. One should notice the probabilities of these different events do not sum to 1. That means less than 1 event is expected to happen. A study on the occurrence of exceptional events is recommended to benchmark the estimate.
The project will now be simulated using the different methods. In the next section the results of the different simulations are presented and described. Based on the simulation, a probability density function will be present and a cost estimate can be provided. A verification of the APM Terminals' requirements with the model leads to the selection of the best model.

6 Cost estimation results for the Moin case

This chapter reports and evaluates the estimation results. This is the final stage towards a probabilistic model. The steps that have been set are summed to clarify the path. Chapter 1 gave an introduction to the problem and set out an approach to solve this issue. In Chapter 2 any background literature was described before proposing several tools to estimate the costs of civil construction works. Chapter 3 presented six different estimation tools;

1. Level I deterministic method
2. Level II semi probabilistic method
3. Level III normally distributed functions
4. Level III various functions
5. Level III dependency
6. Level III dependency and shocks of drivers

The different tools are used to estimate the civil construction works of a new container terminal in Moin, Costa Rica. The project and the civil works are introduced in chapter 5. An overview of the different works is provided, as well as the uncertainty on the design and the underlying assumptions. This uncertainty causes exceptional events that may happen.

Using the estimation tools two different estimates of the Moin project will be provided; the class 5 and class 4 estimate. The civil engineering department provides these estimates within the process to establish a new container terminal. As the model is developed to help the civil engineers estimating civil construction costs, these classes are chosen.

This chapter describes the final step of the development of the new model; the estimation results. The results are reported and analyzed. The different tools and approaches are verified with the APMT requirements to select the best model. The APMT requirements are described in chapter 4.

6.1 Class 5 estimate

The class 5 estimate is the first estimate developing a new container terminal. The estimate is an order of magnitude cost estimate. Little project information is known; based on a brief market survey there is a general idea of the throughput capacity and the dimensions of the design vessel. The little information makes these specifications prone to changes. Changes within these specifications generally have a large impact on the design and construction costs. Thus, structural design and the resulting costs are heavily prone to uncertainty due to the specification changes.

Apart from variability of the requirements the blurred insight on local conditions introduces additional uncertainty. Lacking information on e.g. soil conditions and local prices of construction materials drives additional variability of the construction costs. The uncertainty makes it pretty difficult to come to an accurate forecast of the civil scope of works. Expert experience and judgment determine the structural design to base the costs. Often assumptions on the terminal requirements and structural design are underlying the cost estimate. Additional information in a later stage may prove incorrectness of the assumptions. This introduces uncertainty on the structural design and also uncertainty on the estimate.

6.1.1 *Deterministic estimate*

Within the Moin introduction in Chapter 5, the scope of works was specified. In Table 5 the scope elements are given again. The identified scope is estimated to set the raw estimate. CRAF 1 (or development allowance) and CRAF 2 (or construction contingency) are added to the estimate to account for lacking information and uncertainty. The premiums are set by expert judgment and the project stage, although specific project information is not used. As the contingency and estimated variability depend on the project stage, the approach is often called "black box" method.

The applied contingency is two-fold. First, CRAF 1 accounts for lacking formation introducing additional costs to the scope of works. It is known by experience additional costs will come to the estimate.

The CRAF 2 contingency is incorporated as a safety factor to protect the company against its risk exposure. Risk is introduced by variability of the scope of works and the occurrence of exceptional events. It is unknown whether the additional premium will be spent or not.

Scope Elements	QTY	Price per unit	Total
General	1	50,000,000	50,000,000
Breakwater	2,500	30,000	75,000,000
Wharf	1,500	60,000	90,000,000
Dredging	11,000,000	8	88,000,000
Reclamation/Revetment	6,000,000	5	30,000,000
Container yard	750,000	80	60,000,000
Mechanical	1	70,000,000	70,000,000
Buildings & gates	21,000	750	15,750,000
Raw estimate			478,750,000
Development allowance	15%		78,812,500
Construction contingency	15%		78,812,000
Total			622,375,000

Table 5; Class 5 deterministic estimate construction costs

Table 5 is pasted from a Microsoft Excel sheet. The green fields allow for the input; blue fields show the outcomes. The white fields show additional information to support the estimate. Accordingly the deterministic tool, the civil CAPEX should be set at USD 622,375,000. APM Terminals' internal documents(2) point out the estimate is expected to present the expected value of the civil works (i.e. P50).

The internal documents also show the actual costs are expected within a range of +/- 40% of the total costs. This means the lower limit of the interval is USD 373,425,000; the upper limit is set at USD 871,325,000. The confidence interval is assumed to hold with an 80% probability. Any distribution function can be assigned to the estimate. It is assumed expected value is 622.4mn and the 80% confidence boundaries are USD 373,4mn and USD 871.3mn.

6.1.2 *Normal distribution estimate*

The deterministic estimate does not explicitly incorporate estimate uncertainty. As a first approach towards a probabilistic estimate the semi probabilistic Level II method is used. Uncertainty is incorporated by taking the mean and standard deviation of all (to be) estimated elements into account. Normality is assumed, but that does not mean the elements are normally distributed. Of any function the mean and standard deviation can be determined and inserted into the normal distribution.

Of all the different scope items both quantity and unit rate can vary (normally distributed). The product of both is assumed to be normally distributed as well. The justification is shown in Appendix 4. The experts have to estimate the mean and standard deviation. The standard deviation is estimated as a percentage of the mean. This is called the variability. The identified scope is presented in Table 6 below;

	Quantities			Price			Statistical characteristics		
Normal events	Mean	Stdev	Variability	Mean	Stdev	Variability	Mean section	Std section	Variability
General	1	0	0%	50,000,000	7,500,000	15%	50,000,000	7,500,000	15%
Breakwater	2,500	125	5%	30,000	4,500	15%	75,000,000	11,871,875	16%
Wharf	1,500	75	5%	60,000	7,500	13%	90,000,000	12,129,670	13%
Dredging	11,000,000	1,650,000	15%	8	1	15%	88,000,000	18,772,331	21%
Reclamation/Revetment	6,000,000	900,000	15%	5	1	15%	30,000,000	6,399,658	21%
Container yard	750,000	37,500	5%	80	12	15%	60,000,000	9,497,500	16%
Mechanical	1	0	0%	70,000,000	8,750,000	13%	70,000,000	8,750,000	13%
Buildings & gates	21,000	2,100	10%	750	75	10%	15,750,000	2,232,948	14%
Unforeseen	1	0	0%	35,000,000	10,500,000	30%	35,000,000	10,500,000	30%
Subtotal							513,750,000	31,932,388	6%

Table 6; Class 5 Normal estimate construction costs

To the different scope items exceptional events are added. The different events were already introduced in Chapter 5. The probabilities of occurrence, consequence and variability on the consequence have to be estimated. The result of this estimate is shown below;

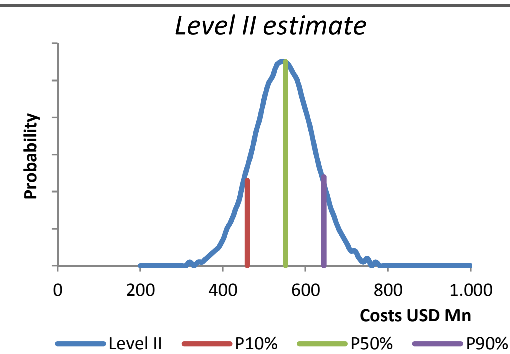
Scenarios	Prob.	Price	Variability	Mean	Stdev
Contaminated soil	1.0%	40,000,000	5%	400,000	3,984,972
Changing wave conditions	1.0%	15,000,000	10%	150,000	1,500,000
Soft soil	15.0%	100,000,000	10%	15,000,000	35,916,570
Hard soil	15.0%	150,000,000	10%	22,500,000	53,874,855
				0	0
				0	0
				0	0
				0	0
Subtotal				38,050,000	64,889,367

Table 7; Exceptional events normally dsitributed

A few remarks to the input data and first results are necessary. n additional line item has been added to the scope of works; Unforeseen. This item is similar to the application of CRAF 1; it is known additional costs will come to the estimate. The reason for application is to align the expected costs to the deterministic estimate only including CRAF 1. CRAF 2 is defined as a safety factor. The expected costs of the probabilistic approach should therefore not meet the deterministic estimate including CRAF 1 and 2. This also implies the first assumptions of APM Terminals on the output results will not be met. The assumptions defined the deterministic estimate the expected costs of the probabilistic approach. Still it is interesting to position the deterministic input within the probabilistic approach and verify the APM Terminals assumptions on the variability of the estimated costs.

This probabilistic output also shows a few interesting other phenomena. The variability of the total scope of works is estimated 6%. The variability of the events is 170%. This leads to the conclusion the exceptional events heavily contribute to variability of the total cost density function. The contribution of the mean is the other way around; the scope mainly contributes to the expected costs. The total distribution function is shown in Figure 36.

This 80% confidence interval bounded by the P10 and P90 deviate 17% from the mean. The APM Terminals assumed a +/- 40% deviation from the mean. Clearly, the output does not match the initial assumptions. It is also interesting to position the deterministic estimate within the probability density function. The deterministic estimate presents the P84. If the deterministic estimate including CRAF 1 and 2 is the budget, the costs do not overrun by 16%.



Mean	551.800.000
Standard deviation	72.320.864
Variability	13%
P10	459.117.083
P90	644.482.917

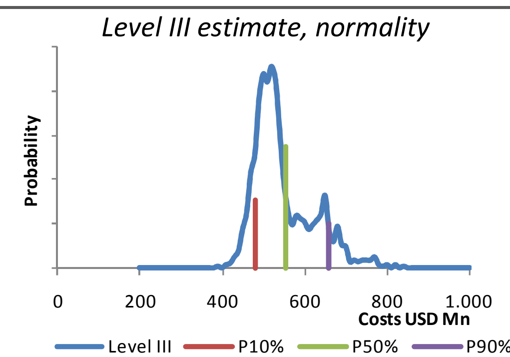
Figure 36; Level II estimate

6.1.3 Level III Standard Normal distribution

The Level II approach is useful as a first approximation. The assumption of normality is justified as the behavior of many independent variables converges towards the normal distribution function. This is proved by the Central Limit Theorem.

But this simplification introduces errors modeling the events. The discontinuity causes this error. As a first approximation, the error can be neglected. At the other hand, one is interested in an accurate solution. Therefore the same input is used, but the discontinuity of the events is taken into account to achieve a more accurate result. The shape of the line differs as an exact solution cannot be obtained; simulation is used. The project density function is shown in Figure 37.

The incorporation of the discontinuity shows a asymmetric distribution function skewed towards the right. It also changes the shape of the confidence



Mean	552.092.306
Standard deviation	72.135.425
Variability	13%
P10	479.681.318
P90	658.884.483

Figure 37; Level III estimate, normality interval from +/- 17% to -13%/+20%. But the assumptions of +/-40% are not met. Within this density function the deterministic estimate represents the P83; a minor difference with the level II approach.

6.1.4 Level III Various distribution functions

The first two probabilistic models did meet the assumptions on the mean and confidence interval. The deterministic estimate does not equal the mean of the project density function. As the CRAF 2 contingency is a safety factor, the probabilistic mean can simply not equal the deterministic estimate. Above that, the confidence interval does not meet initial assumptions of +/-40%. A tool that allows for various distribution functions may present better results. There is additional flexibility as the expert can estimate the upper, lower limit and expected value of the different scope elements. The expert can also assign a distribution to each element. In this thesis six different distribution functions are chosen. These functions were shown in chapter 3. In Figure 38 each density function (PDF) is assigned a number. The different numbers and belonging distribution functions are;

0. Fixed estimate
1. Standard normal distribution
2. Lognormal distribution
3. Student-t distribution
4. Triangular distribution
5. Uniform distribution

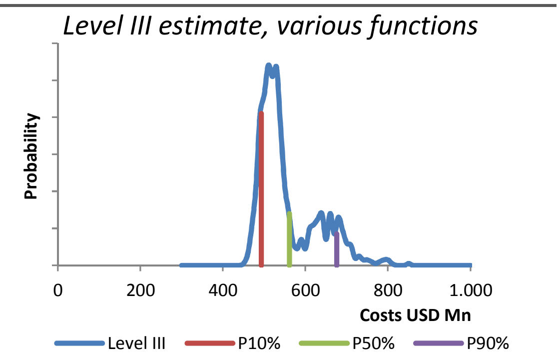
The scope elements, different distribution functions and confidence interval limits are shown in Figure 38. Again, the unforeseen item is used to align the probabilistic estimate to the deterministic estimate (only including CRAF 1). Below the scope items, the exceptional events are quoted.

Tender Year	Quantities				Price per quantity				
Scope elements	Low	Expected	High	PDF	Low	Expected	High	PDF	Mean section
General		1		0	35,000,000	50,000,000	70,000,000	2	50,000,000
Breakwater	2,400	2,500	2,600	1	25,000	30,000	42,000	4	75,000,000
Wharf	1,400	1,500	1,600	1	55,000	60,000	75,000	4	90,000,000
Dredging	9,000,000	11,000,000	12,500,000	5	6	8	10	4	88,000,000
Reclamation/Revetment	4,500,000	6,000,000	7,500,000	5	5	5	8	4	30,000,000
Container yard	700,000	750,000	800,000	1	60	80	90	4	60,000,000
Mechanical, electrical & power supply		1		0	50,000,000	70,000,000	90,000,000	1	70,000,000
Buildings & gates	16,000	21,000	24,000	5	500	750	825	4	15,750,000
Unforeseen		1		0	20,000,000	35,000,000	50,000,000	1	35,000,000

Scenarios	Probability	Consequence				Mean section
		Low	Expected	High	PDF	
Contaminated soil	1.0%	30,000,000	40,000,000	70,000,000	2	400,000
Changing wave conditions	1.0%	10,000,000	15,000,000	20,000,000	5	150,000
Soft soil	15.0%	90,000,000	100,000,000	130,000,000	4	15,000,000
Hard soil	15.0%	135,000,000	150,000,000	200,000,000	4	22,500,000
						0
						0
						0
						0
Final COST						**551,800,000**

Figure 38; Monte Carlo input

The simulation is run in Matlab. The function files with the scripts of the simulations are shown in Annex 3. The output is transferred into Microsoft Excel to visualize the results. The probability density function using the Matlab tool is shown in Figure 39. Again, the confidence interval does not meet the +/-40% constraint as it is -12%/+20%. It is also shown the input mean does not comply with the simulated mean. This difference is introduced by the difference between the most likely and the expected value of several functions and was discussed extensively in Chapter 3. As the deterministic estimate does not reflect a P50, it is still interesting to determine its position with the project density function. For this particular function, the deterministic estimate is the P77.



Mean	561.039.946
Standard deviation	71.977.919
Variability	13%
P10	493.239.903
P90	675.679.813

Figure 39; Level III estimate, various functions

Now we have simulated two different estimates (apart from the Level I and Level II method). The first estimate assumed normal distribution and took discontinuity of the events into account. The other simulation took use of various distribution functions. Between both estimates a minor difference is present. The mean of both functions differ; the estimated mean is USD 551.9mn; the simulated mean USD 561.0mn. This is caused by the use of asymmetric functions. The standard deviation as well as the P10 and the P90 of both functions (correcting for the app. 10mn difference) are quite close. The difference between both functions is visualized in Figure 40.

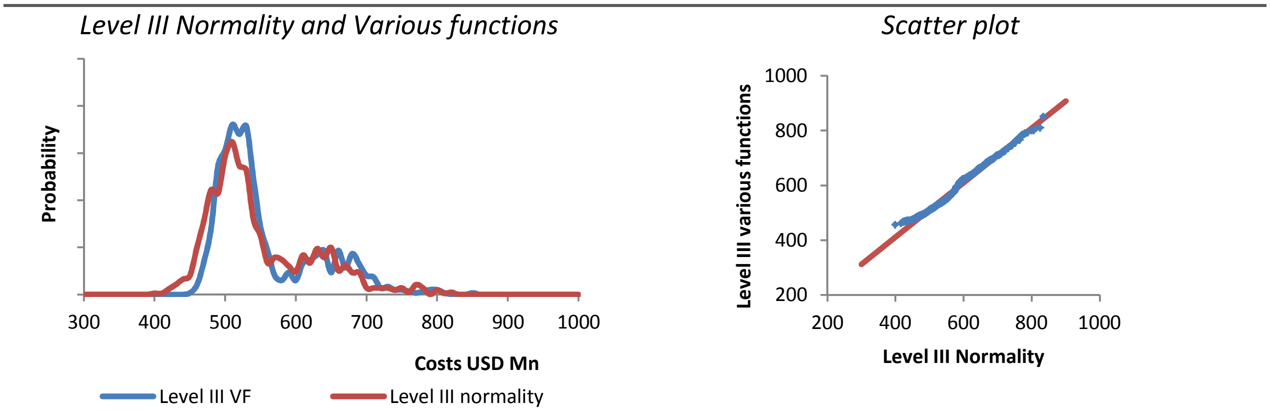


Figure 40; Difference between Monte Carlo and normal distribution

The left graph shows the combined distribution functions. A minor difference between both functions is shown. The right graph shows both simulated series simulated from low to high. The scatter plot also shows a straight line drawn through the scatter plot. The data points almost follow the linear function. Additional regression analysis is in accordance with the graphics; there is minor difference between the shapes of both functions. Analysis shows a R^2 value of 98% indicating a 98% similar behavior.

6.1.5 Level III dependency

The fifth model decomposes the price of the various elements to different commodities. Within this class 5 estimate five different drivers are identified. The continuous drivers behave accordingly the standard normal distribution and the behavior is interrelated (measured by correlation). The time series is monthly data derived from the US Bureau of Labor Statistics except for the oil time series. The time

series of oil is described by the Gasoil day data derived from Reuters'. (The oil time series is converted to monthly data.) All series start June 2007 and end June 2010.
The five different identified drivers are;

1. Steel price; approximated with the index series of iron, steel mills and ferroalloy steel
2. Oil; assumed to be represented by Reuters' Gasoil spot price (converted to monthly data)
3. Cement prices; presented with Portland cement prices
4. Contractor prices; reflected with producer price index data of the new industrial building construction.
5. Labor market; assumed to behave accordingly the average weekly earnings of production and nonsupervisory employees within construction industry.

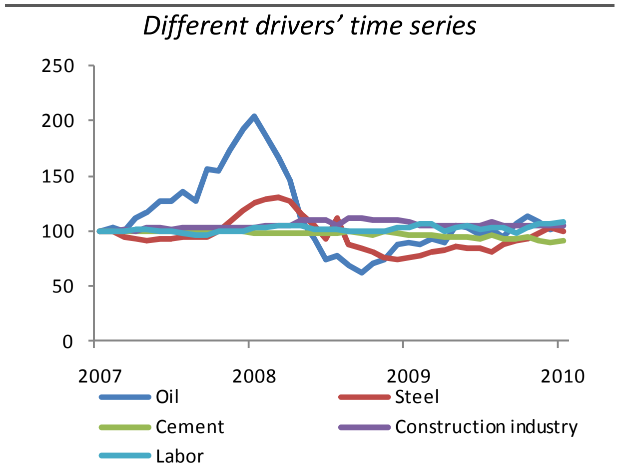


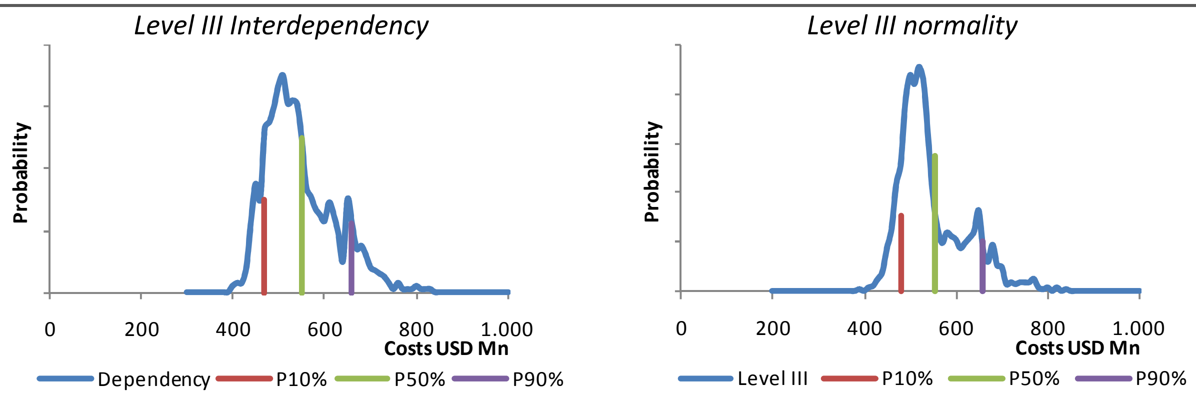
Figure 41; Decomposition input

The different time series are shown in Figure 41. All time series start at 100 to present comparable behavior. The volatility of oil and, to a lesser extent, steel is clearly visible.
Figure 42 below shows the estimated scope. This estimated scope has not changed from previous estimates. Next to the scope, the price is decomposed to the five different drivers. For example, oil accounts for USD 4,500(15% out of USD 30,000) per unit breakwater. As the variability of oil is known and the variability of the price per unit (for each line item) the variability of each element can be determined.
Without determining the R^2, one may already have an idea of the outcome. Figure 42 clearly shows the error term of the price (measured by "other") accounts for approximately 50%. As this is only the contribution to the mean, it does not necessarily mean the variance is large enough. But it indicates the estimated contribution is pretty small. This may lead to a larger standard deviation of the error term than the standard deviation of the identified part. Therefore the tool is, at forehand, not expected to have a (significant) impact on the variability of the project density function.

Scope Elements	QTY	Price per unit	Price decompsition					
			Driver 1 [Oil]	Driver 2 [Steel]	Driver 3 [Cement]	Driver 4 [Contractor]	Driver 5 [labour]	Other
General	1	50,000,000	0%	0%	0%	8%	40%	53%
Breakwater	2,500	30,000	15%	0%	10%	10%	10%	55%
Wharf	1,500	60,000	10%	5%	30%	10%	20%	25%
Dredging	11,000,000	8	35%	0%	0%	15%	5%	45%
Reclamation/Revetment	6,000,000	5	35%	0%	0%	15%	8%	43%
Container yard	750,000	80	10%	0%	5%	15%	25%	45%
Mechanical	1	70,000,000	5%	25%	0%	10%	30%	30%
Buildings & gates	21,000	750	15%	15%	10%	10%	30%	20%
Unforeseen	1	35,000,000	0%	0%	0%	0%	0%	100%
	0	0						100%
	0	0						100%
	0	0						100%
	0	0						100%
	0	0						100%
	0	0						100%
Total		513,750,000.00						

Figure 42; Input decomposition

The exceptional events are added to the scope. As the scope is decomposed assuming normality, the exceptional events are assumed normally distributed as well. Thus, the event input of the Level III normal tool is used. Another advantage is that we can easily compare this approach to the Level III normal distribution approach. Figure 43 therefore shows the Level III dependency and the Level III normal distribution estimates.



Mean	550,809,646	*Mean*	552.092.306
Standard deviation	74,825,125	*Standard deviation*	72.135.425
Variability	14%	*Variability*	13%
P10	469,314,581	*P10*	479.681.318
P90	660,321,429	*P90*	658.884.483

Figure 43; Decomposed cost estimates

The two estimates above show the earlier presented feeling. The difference between both functions is negligible and introduced by the simulations. A recalculation of dependency also points this out. The R^2 is 7%. This leads to the conclusion the contribution by dependency is marginal.

6.1.6 Level III dependency and shocks

The historical data reflects historical supply and demand behavior. Future development may be different. To account for these deviations the drivers are allowed (with an estimated probability) to deviate from historical supply and demand. The deviation is introduced by change (i.e. jumps) of the drivers' mean. For each driver the mean is defined as its stake of the price. This contribution may jump up or down. The estimated jumps and probability of occurrence for the different drivers are shown below. The variability of the drivers is now composed of the shocks and the historical behavior.

	Perc. change of discrete drivers				
	Driver 1 [Oil]	Driver 2 [Steel]	Driver 3 [Cement]	Driver 4 [Contractor]	Driver 5 [labour]
down	-50%	-20%	-10%	-50%	-10%
equal	0%	0%	0%	0%	0%
up	50%	20%	10%	50%	10%

	Probability of change discrete drivers				
	Driver 1 [Oil]	Driver 2 [Steel]	Driver 3 [Cement]	Driver 4 [Contractor]	Driver 5 [labour]
down	5%	5%	5%	10%	8%
equal	90%	90%	90%	80%	85%
up	5%	5%	5%	10%	8%

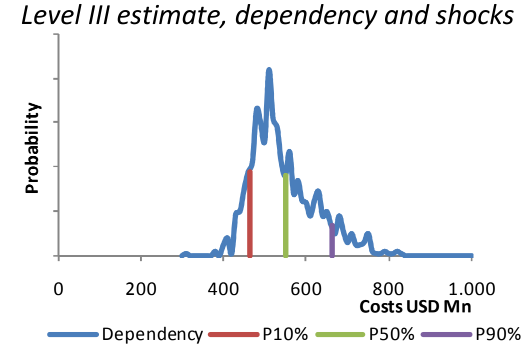
Figure 44; Supply and demand changes drivers

The shocks introduce additional variability. The total estimated variability cannot change, as well as the earlier estimated dependent variability. The introduction of these shocks should therefore increase the explained variance and the R^2. But it still is doubtful whether it would result in a significant impact as a small contribution of the drivers was estimated.

Figure 45 shows the estimation results. These results show no difference with the earlier presented estimate. The shocks have increased the R^2 to 8%. This difference will not influence the variability of the total project density function.

Both attempts incorporating dependency did not have a significant influence on variability. This also means the assumption of a +/-40% confidence interval are not met by this approach. The additional effort did not have an influence on this class 5 estimate.

Level III estimate, dependency and shocks



Mean	550.331.447
Standard deviation	79.310.227
Variability	14%
P10	465.176.859
P90	661.492.337

Figure 45; Level III estimate, dependency and shocks

6.1.7 Remarks to the estimation output

The preceding sections showed the estimation results of a class 5 estimate using the six different tools. These results were verified with APM Terminals' assumptions on the outcomes. According to these assumptions the probabilistic mean should equal the deterministic estimate and the confidence interval is bounded +/-40% off the mean. The estimation results did not meet these assumptions due to several reasons.

The difference between the deterministic estimate and the probabilistic mean is present by the use of contingency. Two contingencies, CRAF 1 (or development allowance) and CRAF 2 (or construction contingency) are applied. The definition of CRAF 2 is interpreted as a safety factor. That makes it likely to conclude the deterministic estimate will never equal the probabilistic mean.

Unfortunately the assumptions are not backed by a survey of the construction costs. It is only expert judgment and no proof of the assumption is present. That has led to doubt on the assumptions within APM Terminals itself. Some experts expected the costs to present between P70 and P75; the estimate holds with 70% to 75% probability. The results showed the deterministic estimate was approximately the 80% upper limit of the costs. Thus the P80 seems pretty high comparing with the renewed assumption. But variability is important calculating the upper confidence interval and the probabilistic estimate did not meet the assumed variability. It is therefore worthwhile to further analyze variability.

APM Terminals has pointed out the confidence interval is based on industry outlines. The class 5 estimate is heavily prone to variability. The design conditions and assumptions are likely to change. It may be possible the model does not incorporate all present variability. Some part of the variability cannot be identified because of limited time or information. This class 5 estimate is estimated on only 1 structural design, but multiple options are present and not incorporated. This makes it likely to conclude not all variability is incorporated.

This issue cannot be solved by incorporating dependency. This case study showed the contribution of dependency is too low. Above that, the level II results showed the majority of the variability is introduced by the events. The analysis of the variability, presented in Figure 46, shows the variability of the events is mainly introduced by the jump(i.e. $(1-p)*\mu_p$) as σ_p is only a partition of μ_p. Having multiple options, several clocks are present. The variability is introduced by the discontinuity of the different options as the sigma's only have a marginal influence. This leads to the conclusion the variability of the clock has a minor influence. Dependency only influences the variability of the clock. Based on this analysis, the conclusion is drawn dependency is of marginal influence and one should focus on the different events and scenarios while estimating variability.

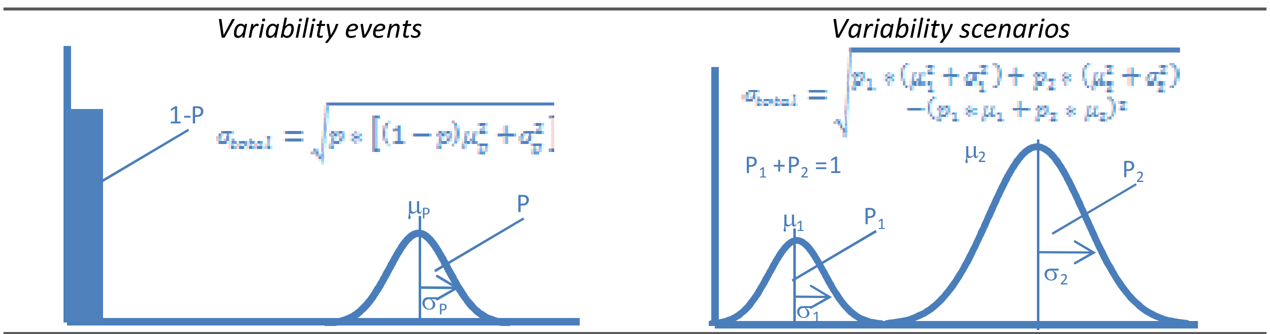


Figure 46; Decomposition input

6.1.8 Conclusion

This section has shown the output results of the different estimation tools. The output of the various probabilistic models was already compared to the deterministic estimate. It is shown the probabilistic output does not meet the probabilistic assumptions of the current provided estimate; the deterministic estimate is not accordingly the expected value and the confidence interval is considerably smaller.

Two reasons cause the deviations between the assumptions and the probabilistic output.

1. The assumptions may be incorrect. Accordingly its definition, CRAF 2 is a safety factor to protect the company against its risk exposure. This implies the CRAF 2 cannot contribute to the expected value but increases the probability of cost control. Above that the assumptions

are based on expert opinion and not proven with an inquiry. Experts have stated the civil CAPEX estimate presents a P70 or P75.

2. The estimate does not incorporate all present variability. The structural design is likely to change. Often the estimate does only account for 1 design instead of multiple designs. The different designs and possible events mainly cause the variability. As only 1 design is used, not all variability is incorporated. It was also clearly shown dependency has a marginal influence on the variability.

In the last paragraph of this chapter the best model will be selected. This will be achieved by aligning the different tools to the various requirements. Based on this verification the most suitable model will be selected.

6.2 Class 4 estimate

The class 4 estimate is an estimate used to verify the costs as part of an in depth analysis to develop a new terminal. The engineer visits the site to determine the local conditions, clarify the structural design and meet possible contractors. Based on this first research the engineer comes up with an estimate. Whereas an engineer did spend limited time to a class 5 estimate, a class 4 estimate allows additional time.

6.2.1 Deterministic estimate

This class 4 deterministic estimate is similar to the class 5 estimate. A similar design is assumed. But, the estimate includes additional scope elements complementing the main line items. The summary of this estimate is presented with the costs of main line items.

Scope Elements	Total
General	**57,370,043**
Breakwater	**75,715,144**
Wharf	**88,941,885**
Dredging	**90,846,487**
Reclamation/Revetment	**33,203,040**
Container yard	**57,216,598**
Mechanical	**45,232,630**
Buildings & gates	**39,862,150**
Raw estimate	**488,387,977**
Development allowance - 10%	48,838,798
Construction contingency - 10%	48,838,798
Total	**586,065,573**

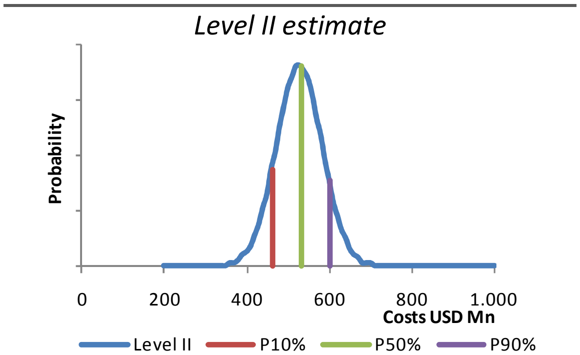
The detailed presentation of the numbers already reveals the estimate is determined with additional detail. The increasing amount of information has led to an improvement of the raw estimate. Additionally, the applied contingency decreases as the estimate is prone to less variability. This leads to a considerable lower estimate; USD 586.1mn vs. USD 622.4mn.

The APMT assumptions (2) estimate the Class 4 confidence interval at +/- 20%. The confidence interval should be bounded by USD 468.9mn (P10) and USD 703.3mn (P90).

6.2.2 Level II estimate

Again the level II method is used as a first attempt to estimate construction costs probabilistically. Of the different elements the variability was estimated and different exceptional events were reviewed. These exceptional events were subjected to in depth analysis. The consequence of each event was estimated as a scenario; a cost density function of each scenario was constructed. The experts preferred this procedure as it was close to their approach to estimate consequences of events. The unforeseen item was used to align the probabilistic estimate to the deterministic estimate (only including CRAF 1). This choice will result in a difference between the probabilistic mean and the deterministic estimate(including CRAF 1 and 2). Therefore the confidence interval will be verified with the deterministic assumptions (i.e. +/- 20% of the mean). Above that, the deterministic estimate will be positioned within the probabilistic estimate.

Figure 47 shows the estimation results of the class 4 input. The estimate shows less variability that justifies the use of smaller contingency factors. At first sight the confidence interval bounded by the P10 and P90 is too small; +/- 13% in contrast to the assumed +/-20%. But the interval is based on the 80% confidence interval. If the (commonly used) 95% confidence interval is taken the confidence interval is bounded +/-20%. Now we get somewhere. Based on a 95% confidence interval it can be concluded the probabilistic estimate presents realistic results.



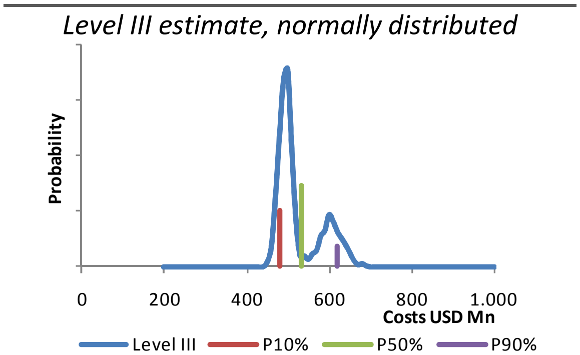
Mean	532,719,035
Standard deviation	55,235,133
Variability	10%
P10	461,932,364
P90	603,505,706

Figure 47; Level II estimate

The next step is to position the deterministic interval within the probabilistic estimate. This level II estimate shows the deterministic estimate holds the P83. Assuming the expert opinion of P70-P75 to be true, the probabilistic estimate does not provide a realistic solution. But this estimate only incorporates 4 possible events. That leads to errors in the tail. Simulation may present better outcomes.

6.2.3 Level III standard normal estimate

The preceding section showed accurate results of the confidence interval. Based on a 95% confidence interval variability exactly reflected expert opinion. The probabilistic estimated the deterministic estimate as the P83. That is too high with respect to the P70-P75 assumption. But, the level II analysis is sensitive to error distributions in the tail having estimated few events. Using simulation methods an more realistic and accurate estimate will be obtained.



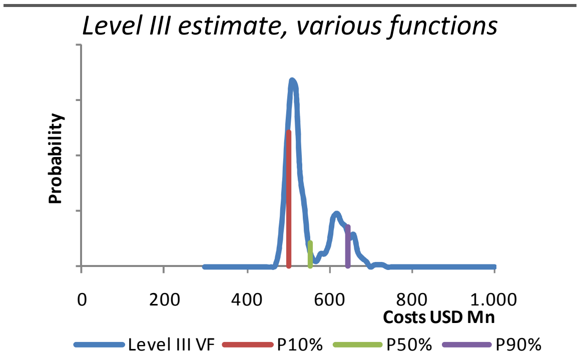
Mean	531.392.309
Standard deviation	54.677.294
Variability	10%
P10	482.188.882
P90	620.707.458

Figure 48; Level III estimate, normality

Figure 48 shows the simulated outcomes assuming all elements and normally distributed, but accounting for the discontinuity of the events. The variability and mean have remained equal (a small deviation is present due to the simulation). Using this approach, the deterministic estimate now represents the 77% upper limit. This outcome is almost reflecting the expert opinion that stated the deterministic estimate is the P70-P75. The outcome is a little on the upper limit. Possibly, the other estimation tools using various functions or dependency show an even better result.

6.2.4 Level III estimate various functions

This level III estimate used various functions to come to the estimate. The experts had the flexibility of assigning six different functions. The functions are shown in paragraph 3.4. The estimate is shown in Figure 49. The mean of this estimate has increased due to the asymmetry of some estimated functions. Although the mean has changed, the standard deviation, variability and confidence interval are clearly similar to the Level III normally distributed estimate. Above that, analysis has shown the shape of both functions complement for 99.5%. This is due to the use of the standard deviation. Of all elements the expert estimated the upper and lower boundaries of the various distribution functions. All the different functions were fitted to it. At the other hand, the Central Limit Theorem proves itself as the sum of differently distributed line converges toward the standard normal distribution.



Mean	555.006.712
Standard deviation	57.417.790
Variability	10%
P10	502.721.505
P90	646.443.479

Figure 49; Level III estimate, various functions

The deterministic estimate is again positioned within the estimate. The deterministic estimate is the P71 of this estimate. This is exactly reflecting the P70-P75 assumption. But, one should notice this outcome is driven by the increase of the mean.

6.2.5 Level III; dependency

Next to the probabilistic tools, the dependency was again incorporated. This (class 4) decomposed input differed from the class 5 estimate. Another expert estimated 6 different drivers. To the five earlier defined drivers, equipment was added. All the different drivers and the figure with the time series are provided below;

1. Steel price; approximated with the index series of iron, steel mills and ferroalloy steel
2. Oil; assumed to be represented by Reuters' Gasoil spot price (converted to monthly data)
3. Cement prices; presented with Portland cement prices
4. Contractor prices; reflected with producer price index data of the new industrial building construction.
5. Labor market; assumed to behave accordingly the average weekly earnings of

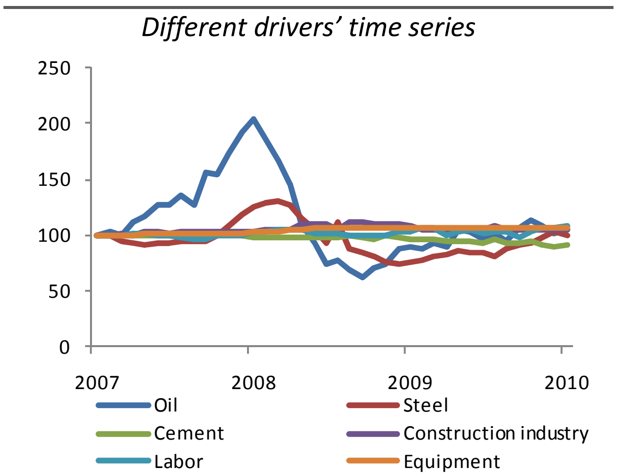


Figure 50; Decomposition input

production and nonsupervisory employees within construction industry.

6. Equipment; described by the index of construction machinery and equipment prices

The expert estimated the contribution of the different line items differently from the class 5 estimate. The expert based his judgment on a consultant estimate. A summary of the different estimated elements is shown below;

Scope Elements	QTY	Price per unit	Price decompsition						
			Driver 1 [Oil]	Driver 2 [Steel]	Driver 3 [Cement]	Driver 4 [Contractor]	Driver 5 [labour]	Driver 6 [equipment]	Other
Summary									
General	1	57,370,043	1%	0%	0%	24%	48%	3%	23%
Breakwater	2,500	30,286	4%	0%	10%	17%	21%	10%	38%
Wharf	1,500	59,295	14%	8%	7%	17%	21%	12%	22%
Dredging	10,790,066	8	12%	0%	0%	17%	49%	21%	0%
Reclamation/Revetment	5,913,533	6	12%	0%	0%	17%	35%	27%	8%
Container yard	740,624	77	1%	1%	15%	17%	50%	2%	14%
Mechanical, electrical & p	1	45,232,630	0%	9%	4%	17%	21%	1%	48%
Buildings & gates	21,000	1,898	1%	27%	5%	17%	38%	2%	11%
Unspecified	1	10,000,000	0%	0%	0%	0%	0%	0%	100%
Total		488,387,977.19							

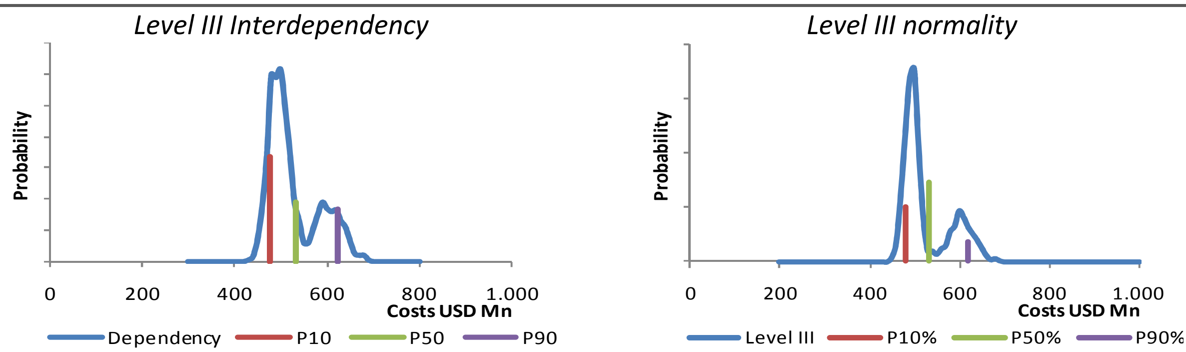
Figure 51; Input decomposition

The input clearly shows a difference with the class 5 estimate. The class 5 estimate showed a contribution of the estimates around 50%. In this example the contribution is mostly around 70%. This approximates the threshold value to have a large explained variance and may imply the different drivers have a significant impact on the variability. This only holds if variability is equally distributed.

Above that, this estimate is a single expert estimate. The class 5 estimate differed significantly. This makes it likely to state that an inquiry on the contribution of various drivers is worth additional research. An analysis of different estimates among different experts may point this out.

To measure the influence of dependency, the variability of each element equals the variability of the Level III normal distributed estimation tool. The total variability may increase by the influence of dependency. This means the variance of each element is composed of the drivers' variance and an error term. This error term is defined as "other".

Figure 43 below shows the estimation results of both the dependent estimate and the (independent) Level III normally distributed estimate. The impact of dependency is negligible. The standard deviation did not change. The error term is still large. Above that, it can be concluded the error term is more widely distributed than the identified parts. This is not entirely surprising as analysis shows an R^2 value of 30%. The R^2 measure is based on the variance of the price. But this R^2 measure increases the total variability of the price by 15%. As the price is multiplied by quantity and different events are introduced the impact on the total standard deviation damps and has become negligible.



Mean	531.797.500	*Mean*	531,392,309
Standard deviation	55.538.665	*Standard deviation*	54,677,294
Variability	10%	*Variability*	10%
P10	476.932.422	*P10*	482,188,882
P90	622.552.745	*P90*	620,707,458

Figure 52; Decomposed cost estimates

An additional note should be provided. The lower limit P10 has changed 6Mn comparing both estimates. This cannot be explained by the standard deviation and mean of both estimates. Those values are similar. But we observed an increase of the standard deviation of the price. We also know an estimate of the costs is composed of five scenarios; one scenario for an event occurring plus a scenario if no event occurs. Each scenarios has its own project density function. If no event occurs the independent cost density function (just the clock itself) varies 3%. Due to interdependency this has increased to 4%. This is the reason for the changing P10. It also leads to the conclusion the total variability (of the entire estimate) is mainly caused by the shocks. The variability of each scenario has minor influence including dependency. The majority of the variability is caused by the shocks. This is the same conclusion drawn in section 6.1.7.

6.2.6 Level III dependency and shocks

Variability of historical data incorporates historical behavior. Future behavior may be different. To account for this future behavior shocks of the drivers are introduced. The possible up and downward movement (in percentages) is shown in Figure 44.

	Perc. change of discrete drivers					
	Driver 1 [Oil]	**Driver 2 [Steel]**	**Driver 3 [Cement]**	**Driver 4 [Contractor]**	**Driver 5 [labour]**	**Driver 6 [Equipment]**
down	**-50%**	**-20%**	**-10%**	**-50%**	**-10%**	**-10%**
equal	**0%**	**0%**	**0%**	**0%**	**0%**	**0%**
up	**50%**	**20%**	**10%**	**50%**	**10%**	**10%**

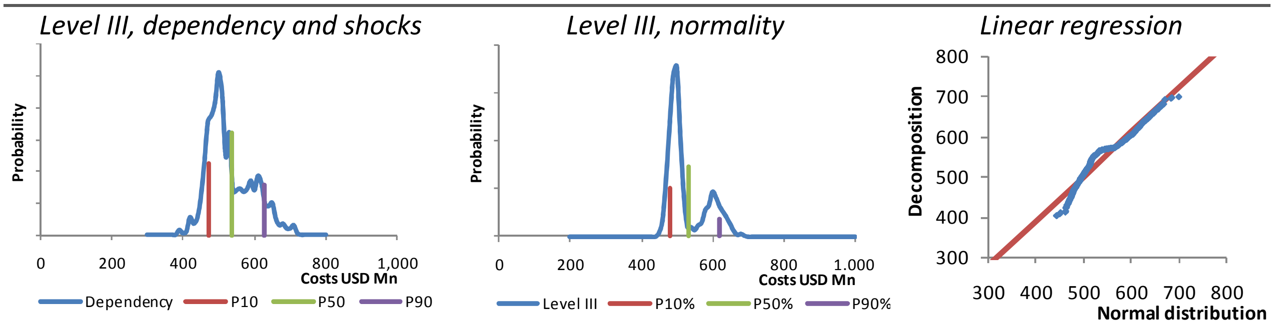
	Probability of change discrete drivers					
	Driver 1 [Oil]	**Driver 2 [Steel]**	**Driver 3 [Cement]**	**Driver 4 [Contractor]**	**Driver 5 [labour]**	**Driver 6 [Equipment]**
down	**5%**	**5%**	**5%**	**10%**	**8%**	**3%**
equal	**90%**	**90%**	**90%**	**80%**	**85%**	**95%**
up	**5%**	**5%**	**5%**	**10%**	**8%**	**3%**

Figure 53; Supply and demand changes drivers

In Figure 54 both the estimate including dependency and shock and normally distributed estimate are presented. Next to both graphs a scatter plot is provided. The output shows a significant change of the shape of the probability density function using the decomposition tool. This is due to the 15% increase of the (total) standard deviation. Analysis has also shown out the R^2 of the price is 49%. The total variance of the price is 49% explained by dependency. 25% originates from dependency and 24% results from the shocks. This leads to 15% increase of total variability. Multiplication of price and quantity and the different scenarios damp the direct influence of dependency. In section 6.1.7 it was pointed out jumps mainly drive variance. The standard deviation of the elements plays a minor role. In fact, the variability of the price has increased by 25%; total variability has increased by 15%.

The graph clearly shows that the standard normal estimate is composed of two bumps; the left bump presents the probability density function without events and the other if an event occurs. Due to the increasing variability the separation between the two bumps vanishes. The reason is the increasing variance of the clocks; the section in between vanishes.

The right figure clearly shows the two density functions are deviating in shape. The blue dots start to deviate from the linear regression line. It is calculated the R^2 value is still at 95%. This is the result of the large cost interval (from 400mn to 700mn) and the small probabilities. The area under the graph should always sum to 1. This will always result in a small difference between the two density functions and a large R^2 value.



Mean	535,972,016	*Mean*	531,392,309
STdev	63,000,327	*STdev*	54,677,294
Variability	12%	*Variability*	10%
P10	469,772,520	*P10*	482,188,882
P90	628,578,967	*P90*	620,707,458

Figure 54; Decomposition vs. normal distributed cost estimate

Furthermore this probabilistic approach shows the deterministic estimate is the P75 of the project density function. The expected value of the decomposed tool is in line with the deterministic output including CRAF 1. The variability bounding a 95% confidence interval is accordingly this estimate 24%.

From these observations it is concluded that the probabilistic approach well presents the variability using the 95% confidence interval. On top of that, the civil CAPEX estimate is likely to present a P75; the estimate holds with 75% probability.

6.3 Comments to the results

Before selecting the best estimation tool for APM Terminals, any general remarks on the output results can still be made. The two probability density functions of both estimation classes are quite similar. As the initial expected costs are based on the similar design, the similarity cannot be not surprising.

But variability of both classes is also quite similar and especially for the class 4 estimate is quite narrow. The next section will treat whether or not it is possible to estimate variability. Here, a few remarks are noted on the variability itself. The output clearly showed the behavior of variability as sketched in Chapter three. There a triangle was constructed composed of three different layers. The width of the layers indicates the drivers' contribution to the cost variability. Above that, the layers are placed in chronological order. Local conditions and requirements are important in the first stage of developing a new terminal; later on the market conditions and construction materials become important. This figure is shown again in Figure 55.

Analysis of variability

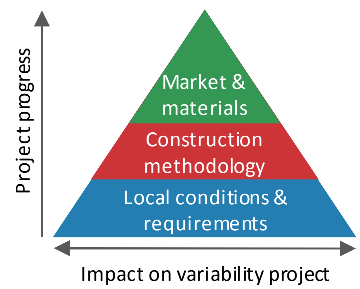


Figure 55; Variability

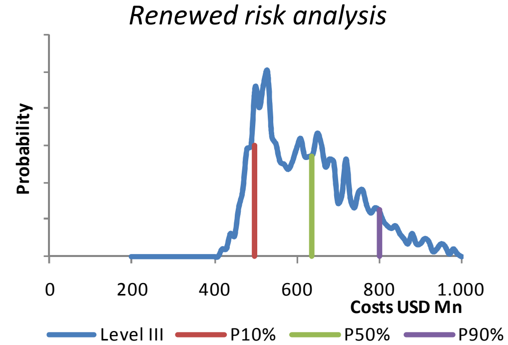
Exceptional events mainly introduce variability as proven by research. It is not the spread on the individual elements but the occurrence of events that cause variability! The estimate results reflect this analysis. Initially, the exceptional events drive variability. The events should be classified to the first two layers. The last layer creates interdependency and has a minority influence. It only increased variability by 20%.

Based on historical observations 1 to 2 events occur every construction projects. Within this estimate only 0.32 events are expected to happen. Identifying a few other events would be more in line with academic literature. It will also widen the project density function and slightly increase expected costs.

An example is shown below to illustrate this analysis. Four different scenarios are added to the earlier presented class 4 estimate;

1. The soil under the quay slides off due to insufficient compaction; 20% probability, 200Mn costs
2. A construction worker loses his life during construction ; 10% probability 25Mn costs
3. Uneven settlements cause additional costs to restore the pavement; 20% probability, 100Mn costs
4. Political interference causes delay and additional costs; 15% probability, 100Mn costs

Renewed risk analysis



Mean	632,312,816
Standard deviation	122,004,834
Variability	20%
P10	498,163,151
P90	803,067,455

Figure 56; Renewed project density function

Now 1.02 events are expected to happen. Figure 56 shows the new project density function using the Level III normally distributed tool. The costs and mainly variability increases significantly. The expected costs increase by approximately 15% and variability increases by almost 50%. This illustrates the need to analyze the risks and possible exceptional events carefully. Interdependency can function in a later stage to fine-tune the risk analysis.

6.4 Selecting the models

Knowing the estimate outcomes of the different tools, it is possible to select the most suitable estimation tool. First the different tools within the class 5 estimation tool are evaluated and scored. The tools are scored from 1 to 5; 1 is the worst, 5 is the best score. The scoring of the class 5 tools is presented in Table 8;

Class 5 estimate	**Level I**	**Level II**	**Level III normality**	**Level III VF**	**Level III dependency**	**Level III dep/shocks**
Input information						
Required information	4	2	2	2	1	1
Engineering judgment	4	3	3	3	2	2
Output figures						
Expected value	3	3	3	3	3	3
Confidence interval	3	2	2	2	2	2
Practical requirements						
Use of statistics	3	3	3	2	3	3
Model size	4	4	4	4	4	4
Use of software	4	4	4	4	4	4
Total	25	21	21	20	19	19

Table 8; scores of the tools providing a class 5 estimate

The class 5 estimate is (in general) provided using little information. If an expert provides an estimate within 4 hours, very little information is known. At the other hand, if the expert spends 4 days to it, he will probably have an understanding of the structural design and an outlay of the site.

Spending four hours to the estimate, it is hardly possible to determine the structural design, let alone to estimate variability and any exceptional events. The input information is as such best aligned to the deterministic estimate. Based on a few assumptions an order of magnitude estimate can be provided quickly.

It is difficult to use a probabilistic approach. Uncertainty on all assumptions is present. This makes it difficult to determine a structural design and leads to an inaccurate estimate of variability. A realistic presentation of the variability can be provided if multiple designs are elaborated. Taking the available time into account, it is impossible to use probabilistics.

The lacking information on the structural design and variability makes the use of dependency not suitable to use. The expert is hardly able to determine a design and variability, let alone to state a meaningful estimate on the various drivers and their contribution.

If an engineer provides a class 5 estimate within 4 hours the Level I deterministic estimate provides the most accurate result. This estimate does not show the variability. But the uncertainty cannot be shown due to the limited time versus the required time to present the variability.

6.4.1 Selecting class 4 tool

Spending 4 days to the class 5 estimate the engineer should have a rough idea on the structural design. Possibly, the expert can determine the possible risk drivers using his experience and a rough desk research. This would make it possible to use the standard normal tool and present the estimate probabilistically.

Providing the class 4 estimate the civil engineer has additional time available, to his benefit. As additional information is present and less prone to changes, a more accurate forecast can be provided. This has led to the evaluation of the various tool presented in Table 9;

Class 4 estimate	**Level I**	**Level II**	**Level III normality**	**Level III VF**	**Level III dependency**	**Level III dep/shocks**
Input information						
Required information	3	4	4	4	3	3
Engineering judgment	2	4	4	2	4	5
Output figures						
Expected value	2	4	4	2	4	4
Confidence interval	2	2	4	3	4	5
Practical requirements						
Use of statistics	1	3	4	4	4	4
Model size	4	4	4	4	2	2
Use of software	4	4	4	2	4	4
Total	18	25	28	21	25	27

Table 9 scores of the tools providing a class 4 estimate

Using the Level III normally distributed tool the engineer has the ability to state his uncertainty on the various line elements and incorporate exceptional events. It is easy to show uncertainty; a probability density function and statistical characteristics can be derived from the output. The model also fulfills the practical requirements; the model is based on MS Excel, the output can easily be transferred into the APMT financial model and remains below 6Mb file size.

The Level I deterministic tool is in accordance with present estimation methodology; the expert cannot incorporate nor present the uncertainty involved with the estimate. Not all information can be incorporated and the uncertainty cannot be presented.

The Level II method does not account for the discontinuity of events. This may introduce incorrect tails and estimates.

The Level III various functions methods provides a similar output to the Level III normally distributed estimation tool. However, the difference between the expected value and most likely value may not reflect expert's expectations and thereby introducing estimation errors. Additionally, this tool cannot be run within MS Excel. That makes it difficult to align the tool to the financial model and APMT daily practices.

The Level III dependency and shocks tool is powerful to use as it can provide an accurate forecast. Only incorporating dependency (i.e. without shocks) did not show significant difference from the Level III normally distributed tool. The tool using dependency and shocks requires quite some information. Probably, the model can therefore be used after spending a few weeks to the estimate. The time span of the class 4 estimate is between 4 days and 4 weeks. The Level III dependency and shocks tool can be used when considerable is spent developing the estimate. Before coming to a correct and accurate forecast one should do additional research on the correct time series. Preferably the model should also be downsized as it now occupies approximately 31 MB of hard drive space.

6.5 Conclusions

This chapter has presented the results of the different estimation tools and selected the best tool for APM Terminals to estimate civil costs. The estimation results of a class 5 estimate showed the difficulty to incorporate all variability. The various design assumptions and conditions are prone to changes. This uncertainty is visualized by the incorporation of the events. Above that, the uncertainty by events drives variability. But it is difficult to monitor all variability and the incorporation will take considerable amount of time. Several designs, exceptional events and variability of the scope of works should be determined.

Based on the different requirements the model has to meet, the Level I deterministic estimate has shown to be the best suitable model. As little time is spent to come to the estimate, the probabilistic models cannot map all variability. A probabilistic approach leads to an incorrect estimate of variability. The deterministic model is therefore best suitable to provide a class 5 estimate. If additional time is spent to the class 5 estimate (i.e. 4 days) it may become possible to use the Level III normally distributed method.

As the time span to provide the class 4 estimate is between 4 days and 4 weeks a two-fold recommendation is given. Spending less than 2 weeks to the estimate, it is recommended to use the Level III normality distributed tool. Thereafter enough information may be present to incorporate dependency and shocks. That requires a lot of information. The detailed information makes it possible to accurately estimate future construction costs of new terminals.

Apart from the selection of the tools any other observations are monitored. The simulations indicated the assumptions on the deterministic (Level I) civil CAPEX estimate (i.e. P50) and confidence interval may not be true. Based on the observations and reviewing the definitions it is concluded the estimate including CRAF 1 represents the expected costs. The estimate including CRAF 1 and 2 is likely to represent a P75; the estimate is assumed to hold with a 75% probability. Above that, the variability was well approximated with the probabilistic approach assuming a 95% confidence interval.

In general it is also shown that shocks mainly determine variability. Shocks originate from different sources; exceptional events or shocks in the different commodities. Above that, it is doubtful whether the assessment of the shocks is sufficient. As one expects the occurrence of 1 to 2 exceptional events the sum of all probabilities (of the events) should equal 1 to 2. All in all, this clarifies the need for engineers to focus on the shocks to map variability correctly.

7 Presenting and using the estimate

Within the preceding chapters different models to estimate the civil construction costs have been developed and selected. This model is delivered separately from this report. A few words will be spent to the model. The model is developed in MS Excel and for each of the classes the model has been developed. The first page gives an introduction to the model. Thereafter the estimation input can be provided. On the last sheet the outcome of the estimate is shown accompanied with a P50 (expected costs), P75 (budgeted costs) and the expected costs if budget overrun takes place. This is summated in a graph.

Now, one is able to distribute the civil CAPEX estimate among the different departments. The civil CAPEX estimate is used to optimize the business case developing a new terminal.

Understanding civil constructions requires a profound understanding of engineering. Deriving the risks of the constructions is therefore not easy. People with an economic (or other than engineering) background may encounter many difficulties. Above that, the AP Moller Maersk group policy is to keep risk simple. Thus the complicated civil works and resulting risks should be reduced to simple basics in the description and review of the business case.

Above that, the civil CAPEX estimate is the first step towards a probabilistic approach. In order to secure a successful implementation, the estimate is aligned to present approach determining project profitability; the deterministic way. The deterministic approach can only take one number into account. Distilling various numbers from the probabilistic analysis would not affect the deterministic process and takes the probabilistic insights into account.

As a result of these remarks a PowerPoint template has been designed to present the civil works and its variability driving factors. The template can be part of the board proposal showing both the results and some of the underlying effort of the estimation process. The sheet is shown in Figure 57 on the next page.

Moin Civil scope of works (output sheet clarifying civil works and risk)

Civil Scope of Works

The Moin green field project is to reclaim a terminal in the bay. The dredged soil from the access channel will be used to reclaim the area.

The terminal area will approximately be a 500 by 1500 meters new island. At the front a 1500 meters quay structure (deck on piles) will be constructed.

Directly at the back of the quay a 450 meters wide yard will be built. The yard will be paved with blocks

A breakwater with concrete armour units protect the new terminal area from incoming waves

Scope specification

	QTY	Price	Total
General	1	5,000,000	5,000,000
Breakwater	2,500	30,000	75,000,000
Wharf	1,500	60,000	90,000,000
Dredging	11,000,000	8	88,000,000
Reclamation/Revetments	6,000,000	5	30,000,000
Container Yard	750,000	80	60,000,000
Mechanical, electrical and power supply	1	70,000,000	70,000,000
Buildings & Gates	21,000	750	15,750,000
Unspecified & exceptional events			38,050,000
Total			471,800,000

Risk driving variables

Several scenarios are incorporated within the estimate to determine the estimate

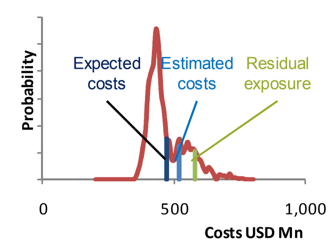
Local technical conditions

- Discovery of hard soil drives additional 150 Mn costs
- Useless dredging material increases costs with 100 Mn
- Presence of contaminated soil leads to 40 Mn cleaning costs
- Larger wave forces from assumptions result in additional15 Mn to improve breakwater construction

Commodity price uncertainty

- The influence of changing commodity prices is not incorporated

Cost outlook



Financials (USD Mn)

Expected costs	471.8
Estimated construction costs	519.3
Residual risk exposure[1]	581.4

Note (1); Residual risk exposure represent the expected costs as cost overrun takes place

APM TERMINALS

3

Figure 57; Presenting civil CAPEX

At the top left of the figure the structural design is described with an introduction to the scope of works. The estimated costs of the civil construction works including the estimated exceptional events and unspecified scope elements is estimated at the top right. The total of the estimated scope shows the expected costs.

The exceptional events mainly drive variability and risk. A detailed description of most important events and the consequence of the events (by means of the expected value) is presented at the bottom left. An insight in the main risk drivers makes it possible to reduce variability and project risk.

At the bottom right the project density function using the probabilistic approach tool is presented. Three numbers are presented;

1. Expected costs; this is the P50 and equals the bottom number of the table specifying the scope of works.
2. Estimated construction costs; this is the civil CAPEX estimate P75. If the estimated construction costs are presented to the board it will also include project contingency or management reserve.
3. Residual risk exposure; presenting the expected costs in case of cost overrun.

This format provides an insight in the civil CAPEX estimate. The civil works are explained and the cost estimate is motivated. Additionally a clear presentation of the risk driving factors is provided. The estimate is made clear with the graphical presentation and the different numbers are clearly presented within the estimate.

Using this layout provides the different end users a good understanding of the civil works and the estimate. Using the residual risk exposure the business case profitability can be back tested. This provides APM Terminals additional value to take a profitable investment decision.

The impact of risk mitigating measures could also be shown easily. A measure will increase cost, but reduce variability. The variability is an important driver of the P75. A measure will set a new estimate (i.e. P75). If the sum of the measure and the new budget are lower than the first budget it is profitable. Otherwise one should accept risk or develop other measures.

8 Concluding remarks & recommendations

This first part of the Master's thesis proposes an approach to improve the estimation process of civil construction costs within APM Terminals. Within the development of new terminals, these costs are an important driver to the forecasted profitability. Projects require major upfront investments in civil infrastructure before revenues start.

At present the costs are estimated in a purely deterministic matter. A fixed point estimate is provided. Uncertainty and variability are accounted for by applying various contingency factors. These factors are added to account for variability in pricing, quantities and exceptional events.

This estimation process has its limitations. Two different projects with a similar CAPEX estimate are valued equally. As the uncertainty differs, the probability and consequence of cost overrun differ. Now APM Terminals neglects this information that influences the investment decision.

Above that, engineers apply contingency without insight on the project variability. This makes the applied contingency prone to inconsistencies. Above that, insight in the project variability can directly show which measures need to be taken to reduce project variability.

To overcome the issues this study analyses six different estimation tools. These tools are;

1. Level I deterministic tool; this is accordingly the current APM Terminals estimation process
2. Level II semi-probabilistic approach; of all the different elements and events the mean and standard deviation are accounted for. These are fitted to the standard normal distribution. the estimated scope (i.e. unit rates and quantities) are assumed normally distributed as well as its product.
3. Level III normally distributed tool; this simulation tool is used to incorporate the discontinuity of exceptional events and model tails more accurately
4. Level III various functions method; all elements are assigned a different distribution function.
5. Level III dependency tool; the price is assumed to be driven by various price drivers. The contribution is estimated to determine correlation.
6. Level III dependency and shocks; possible future behavior is not accounted for. Shocks of the drivers are incorporated to present this behavior.

These tools are used to estimate the costs of a new terminal in Moin, Costa Rica. Two estimates were provided; a class 5 and class 4 estimate. The output results were aligned with APM Terminals' assumptions on the estimate and with APM Terminals' requirements.

The class 5 estimate is a rough estimate to roughly verify project feasibility. Within 4 hours to 4 days an estimate is provided. This short time, especially for the 4 hours' estimate, makes it difficult to determine a structural design and have an idea of the main risks. Above that the design requirements are heavily prone to changes driving additionally variability. Being aware of this unknown variability, it has led to a poor performance of the probabilistic models. Therefore it has been concluded the Level I deterministic method would be best suitable to APM Terminals' estimation process. If additional time is spent (i.e. several days) it may be possible to start using the Level III normally distributed tool.

The class 4 estimate is a more detailed estimate based on additional research and information. As the design requirements are determined and the site is visited, a rough structural design and a general idea of the risks can be determined. The engineers now have an idea of the project variability.

The Level I deterministic estimate does not incorporate all information and is therefore not recommended for future use in the more detailed classes. Based on the evaluation criteria, the Level III normally distributed tool is recommended. Spending additional time on the class 4 estimate (i.e. several weeks) the Level III dependency and shocks tool may be of additional advantage. A future estimate can be provided to estimate the price development of the different drivers more accurate. Other tools have not shown additional value. The trade-off is based on choice of software, output performance and required information. The output does not differ from the Level III normally distributed tool or the tools require specialized software.

The observations have also led to the conclusion the APM Terminals' assumptions probably do not reflect reality. Simulations have shown the deterministic estimate including CRAF 1 represents the expected costs of the probabilistic approach. The deterministic estimate including CRAF 1 and 2 is shown to present the P75; the estimate is assumed to hold with a 75% probability.

It was also shown shocks mainly drive variability. To correctly map variability one should carefully analyze possible occurrence of exceptional events in design, during operations or price driving factors.

The insights resulting from the simulations and the current APM Terminals and AP Moller Maersk approach towards risk management are combined to successfully use the probabilistic approach. Three different numbers and the project density function are provided to the end users. This provides the decision makers information to take a solid investment decision and secure future profits.

8.1 Recommendations

The probabilistic approach is a step ahead to APM Terminals determining its future possible profits and losses. Further improvement can be achieved in the following ways;

- The decomposition tool can be further improved. A considerable difference was found between the classes 5 and 4 estimated contributions of the drivers to the unit rates. A survey to determine the contribution of the drivers to other estimates improves the application of the tool.

  Within this thesis assumptions are made on the applied time series. Possible other time series are available. These series might present the behavior more accurate.

  Using the tool it may also be possible to take the flow of time into account. The behavior of the different drivers can be forecasted. Based on the factor time is taken into account providing a more accurate forecast.
- A research on the estimates, realized costs and exceptional events. Such an inquiry provides a benchmark of the expert's estimates. It will also show the presence and behavior of shocks.

  This would improve future estimates as the effect of scope changes and exceptional events can be measured. Additionally the assumptions on the provided estimate can be verified. Is the estimate represented with a P75 and is variability plus or minus 20%?

  Additionally the differences between the experts' estimates can be presented. The variability may be driven by project characteristics, the estimation class or both. It is interesting to measure the estimated variability among different experts and different classes for various projects.

Part B – Rotterdam Study

The impact of shocks on optimal investments of uncertain costs

Acknowledgment

This final part covers the last three months of work to finish the Master's thesis. This last part has a strong academic focus and contrasts with the first part that has a strong focus on daily practice. This last part was very challenging by the opportunity to cross various academic boundaries. The model itself has an econometric origin, it also relates closely to applied mathematics and civil engineering, but also touches upon information technology.

My sincere gratitude goes to three people that helped me crossing these various boundaries. Michel van der Wel helped me going through the complicated theory and summarize all the results clearly and succinctly. Prof. Kees Oosterlee took a few hours of his Friday afternoon to explain me all the difficulties on numerical analysis and come to the solutions using numerical mathematics. Finally, my twin brother Aad van Buuren helped by converting the Matlab scripts to C and the expansion of his internal memory to a hexa-core processor. The conversion and expansion increased the calculation speed of 1 graph from 3 days to 1 hour. The entire result is what you find next. I hope you will enjoy this second part as much as I enjoyed to develop it.

Coos van Buuren
11th of January 2011

Introduction

This second part of the Master's thesis studies the valuation of projects and optimal stopping rules using option pricing theory. The pricing of financial options originates from Black and Scholes (1973) and Merton (1973). This thesis studies options of real projects that are therefore called Real Options. This study uses academic research by Pindyck (1991, 1993) to analyze investment decisions and define optimal stopping rules of large construction projects subjected to cost uncertainty. Moreover, I add shocks as academic research showed most construction projects face unexpected events (Vrijling and Boschloo (2001)) and Part A showed shocks have a major impact on the estimation of construction costs. This study analyses the impact of shocks on the option value and the decision to invest.

Scientists commonly understand that option theory better explains investment behaviour of real projects than traditional investment models. These traditional models define investments to take place if the sum of discounted cash flows is positive. But this approach implicitly assumes investment decisions are a "now or never" proposition. Option theory accounts for the value to "wait and see". Agents tend to "wait and see" because of irreversibility of investments and uncertainty on cash flows. Irreversibility covers sunk costs that cannot be recovered with some disinvestment. Uncertainty covers the changing value of cash flows. The irreversibility and uncertainty causes postponement, delay or temporary stopping of investments. Then investors wait for new information to reduce uncertainty and improve profitability. Ultimately investors may decide to stop and abandon the project to protect the company from excessive losses.

The fundamentals of Real Options originate from Brennan and Schwartz (1985) and McDonald and Siegel (1986). These papers propose the application of option theory to real projects. Much literature has been published thereafter. The influential book by Dixit and Pindyck (1994) provides the theory to study the general question of investment under uncertainty.

This thesis mainly relates to papers by Grenadier and Malenko (2010) and Miao and Wang (2007). Both papers include shocks and discuss the option to learn from stochastic behaviour and especially shocks. Often there is not only uncertainty on the future shocks, but also on the permanence of past shocks. The addition of Bayesian uncertainty introduces another option "to learn" (next to the option "to wait and see") to distil permanent from temporary shocks. Like most studies of investment Grenadier and Malenko (2010) and Miao and Wang (2007) assume the payoff from the investment that is uncertain. This thesis studies investments under cost uncertainty. The costs of construction projects are particularly more uncertain than the payoff. This different uncertainty (on the costs instead of the project payoff) also changes the nature of shocks. In the studies of Grenadier and Malenko (2010) and Miao and Wang (2007) investors observe shocks before investments (ex ante). Under cost uncertainty shocks occur ex post (i.e. after the investment decision has been taken). Then, investors question the profitability to continue investments and the ability to learn has minor importance. This study analyzes investment decisions under cost uncertainty and shocks may occur while investments take place.

This study restricts to the classical Real Options and does not consider competitive bid processes and other frictions of investments under uncertainty. Nevertheless this thesis shares the importance of competition and frictions. Lambrecht and Perraudin (2003) study the competition between two firms for a single investment opportunity when information about investment costs is private. Because of this, as time goes by, while the competitor has not invested yet, each firm updates its belief about the competitor's investment costs upward. Smit and Trigeoris (2006) illustrate the use of real option valuation and game theory principles to analyze investment opportunities that involve strategic decisions under uncertainty. Smit and Trigeoris (2006) showed that R&D investments in oligopolistic and innovative industries improve the competitive position and the ability to better capture future growth opportunities in the Industry. Sundaresan and Wang (2007) study the effect of capital structure on investments under uncertainty. Sundaresan and Wang (2007) showed that the (re)negotiation of loan capacity and facility influences the firm's investment and financing decisions, that ultimately affects firm value. A stronger bargaining power of equity holders lowers debt capacity, reduces firm value, and discourages growth option exercising. Moreover, the quantitative effects of strategic renegotiation on the firm value may be large.

Other papers studied the relationship of investments under uncertainty with aspects that are beyond profitability and competition. Bernardo and Chowdhry (2002) relate the lifecycle of firms to investments under uncertainty and results of these investments. They argue that firms relate the learning of investments to the commitment of resources by undertaking real investments and observing the outcomes. This implies that the valuation of potential investments depends on the cash flows and the information they gain by investing. Bernardo and Chowdhry (2002) relate this to the corporate strategy, future investment decisions and lifecycle of companies. Their approach predict that young firms specialize when young, then experiment in a different line of business, and then either expand into large multisegment firms or focus and scale up their specialized business. Grenadier and Wang (2007) show time-inconsistent preferences among investors. Agents are impatient about choices in the short term, but are patient when choosing between long term options. Grenadier and Wang (2007) merge the real options approach that emphasizes the benefits of waiting to invest in an uncertain environment, and the literature on hyperbolic preferences in which decision makers face the difficult problem of making optimal choices in a time-inconsistent framework.

This study defines optimal investment rules of large construction projects in four parts. Chapter 1 explains the emerging interest for academic research on option theory. The best explanation of the increasing preference for option valuation is to value an imaginary project using conventional theory. This shows the difficulties and the need for a different approach. Chapter 2 uses the knowledge to construct a model and value projects using option theory. It discusses the model behaviour and solution characteristics extensively. The approach defines points of optimal investment. Chapter 3 adds shocks to the developed model. Most construction projects face unexpected events (or shocks). These shocks not only increase cost but also have a major impact on project variability. Above that, the shocks do not drive a combined up and downward behaviour of project costs, but more of a skewed distribution. The change of stochastic behaviour of the model results in new points of optimal investments. Chapter 4 of this research describes a practical application. This last chapter examines the decision to invest in a new container terminal.

1 Introduction to option theory

The introduction briefly discussed the emerging interest to explain investment behaviour with option theory. Historically the investment decision is based on the net present value (NPV) using discounted cash flow (DCF) modelling. This model implies that investments take place if revenues are at least as large as its costs. An example explains the use of DCF the best.

Consider an imaginary opportunity to develop a new container terminal. The initial costs are I_0=30 and after 1 year the investor receives the terminal with certain value V_1=50. Assuming the investor demands a 10% return on investments, the NPV is

$$\text{NPV} = -30 + \frac{50}{1.1} = 15.$$

The sum of discounted revenues and costs is positive indicating investments should take place.

The investor may also postpone the investment 1 year. The situation changes during this year. The opportunity still requires initial investment I_1=30, but the payoff in year 2 has changed. The value of the container terminal changes with equal probability to 25 or 75. Figure 58 shows both situations.

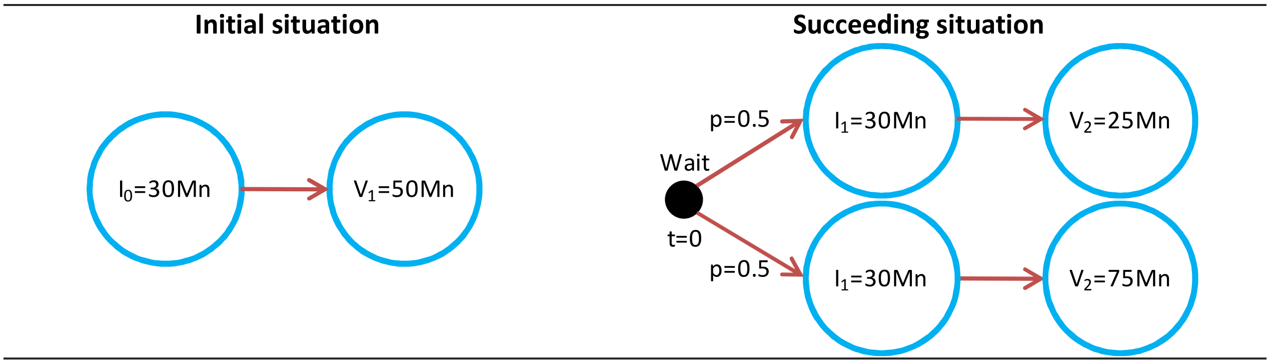


Figure 58; Investment opportunity with option to wait. Initial investments remain 30, but the (later) received asset may change from 50 to 25 or 75 with equal probability.

According to traditional investment theory, the investor immediately spends his money. The expected cash flows for both situations is equal. As discounting of cash flows reduces the value of the postponed cash flows, the initial situation reflects a greater value;

$$-30 + \frac{50}{1.1} > -\frac{30}{1.1} + \frac{50}{1.1^2}.$$

But this approach does not reflect the situation correctly because waiting has a value. The investor only invests if the value increases to 75. The NPV should account for the investor's decision to investment or not. The NPV of each situation is;

$$\text{NPV} = \begin{cases} -\frac{30}{1.1} + \frac{75}{1.1^2} & \text{if investments take place} \\ 0 & \text{if no investments take place} \end{cases}.$$

As each situation has equal probability to occur, the NPV is

$$\text{NPV} = {}^{1}\!/_{2} * \left(-\frac{30}{1.1} + \frac{75}{1.1^2}\right) + {}^{1}\!/_{2} * 0 = 17.$$

The NPV increases from 15 to 17 if the investor postpones the investment. This proves the value of waiting[4]. Closing a costless contract to secure the value of the asset does not help the investor. Then the investor agrees to sell container services at a fixed price. This contract offsets any fluctuations and fixes the value of the asset at its expected value (i.e. 50). Securing a higher value requires a payment to the counterparty that compensates the expected loss. As a result, the contract fixes the NPV at 15 and does not add any value.

There is an analogy between this example and financial options. Financial options give the investor the right to take a position on the financial market e.g. to buy a share of a listed company at a given price. In the example the investor has the opportunity to postpone the investment. In a later stage the investor can withdraw the opportunity or acquire the container terminal (with value 75 at costs 30). This is similar to a financial option as the investor has the right acquire the terminal at a given price. This real option has no date that ends the right. That makes it a perpetual option.

1.1 Valuation of options

The preceding section showed flexibility to choose for immediate or delayed investments has a value. Investors have the (valuable) opportunity to wait. The value of this flexibility is 2 in this example. That is the difference between the NPVs of the consecutive situations. There is also a formal proof to value flexibility aside from the quick valuation (i.e. taking the difference between both NPV's). The formal approach starts by constructing a riskless portfolio. A portfolio is some combination of assets. A riskless portfolio implies a fixed value as risk covers the sensitivity for a value change. The portfolio is composed of the investment opportunity and another asset. In fact, the investor buys the investment opportunity (long position) financed with another asset (short position). This asset replicates the value of the investment opportunity. For this example it is the value of container services; the product of cargo flow, cargo captivity and handling fee. If the investor would construct an oil rig, the spanned (i.e. replicating) asset is oil. If spanning does not hold (i.e. there is no asset that replicated all possible values of the short position) dynamic programming should be used (Pindyck (1991)). Dynamic programming is a computational method that computes the asset value and the option payoff at the same time. For more information on dynamic programming see Bellman (1957). Pindyck

[4] The approach assumes risk neutral investors who are indifferent for risk. Behavioral economics studies economic consumption and individual preferences such as Tversky and Kahnemann (1979) studied.

(1991) describes the application to real options. This research assumes spanning to hold because cargo flow, cargo captivity and handling fee determine the terminal value and can be estimated.

Now return to the portfolio and the example of Figure 58. The portfolio is composed of the investment opportunity and spanned asset. The value of the investment opportunity and the spanned asset change over time. But, the value of the portfolio is fixed by its risk-free property. The initial value of the investment opportunity (F_0) values what the investor is willing to pay today to have the option to invest in the container terminal. In fact, F_0 is the option value that this analysis determines. The value of the investment opportunity in the succeeding situation (F_1) depends on the price change and is

$$\text{Value opportunity } F_1 = \begin{cases} -30 + \frac{75}{1.1} & \text{price goes up i.e. investing} \\ 0 & \text{price goes down i.e. no investing} \end{cases}.$$

The value of the financing asset (P_0) is initially 50. That equals the initial value of the container terminal. For the succeeding situation the price P_1 drops to 25 or increases to 75. The investor can buy multiple assets *n* to finance the portfolio.

The value of the investment opportunity financed with *n* assets makes the value of the portfolio in the succeeding situation

$$\text{Value portfolio } \varphi_1 = \begin{cases} -30 + \frac{75}{1.1} - 75n & \text{price goes up} \\ -25n & \text{price goes down} \end{cases}.$$

Using the risk-free property of the portfolio determines its value in the succeeding situation. The risk-free property implies that the value of the portfolio is independent of price changes. Thus $-30 + \frac{75}{1.1} - 75n = -25n$. This makes the value of the portfolio at the succeeding situation ϕ_1= -19 with n=0.76 borrowed assets.

Initially the investor borrows 38 to finance the portfolio (because $n*P_0$=0,76*50=38). To finance the portfolio the investor pays 10% interest rate. The size of the interest rate has a rational explanation. The counterparty takes a long position on the asset (i.e. invests in the asset) and therefore demands at least 10% return. Thus, the financing costs are 3.8. The return on the risk free portfolio (which is the risk free rate r= 10%) should pay out the interest. Therefore it should hold that;

$$\Delta\varphi - rnP_0 = r\varphi_0,$$

$$\varphi_1 - \varphi_0 - rnP_0 = r\varphi_0 \text{ and } \varphi_0 = F_0 - nP_0.$$

Solving the equation shows the price (i.e. F_0) the investor is willing to pay to have the option to invest. For this example the option to invest F_0 is F_0=17. This equals the value of the option to invest. Above that, it is exactly the NPV if the investor waits.

1.2 Optimal investments

Knowing the value of the option, consider the example again. The investor has the opportunity to invest I_0=30 to receive the payoff V_1=50. Above that, today's investment (of I_0=30) also includes the exercise of the option with value F_0. This makes the full opportunity cost of capital composed of the initial investment and the option to wait. Thus, the full cost of investing is 30+17=47. As the payoff is 50/1.1=45, the investment is unprofitable and the investor should wait. The loss by investing is 2 which formally proves the value of waiting. Moreover, this approach defines the optimal investment rule. Investments should take place if the full opportunity cost of capital (i.e. initial costs and option value) does not exceed the payoff.

This analogy makes it also possible to determine the optimal point of investing under uncertain costs. Consider another example. The initial costs are 20 to obtain an asset with certain value 50. But the investment period may extend with 50% probability and require additional investments 60. Figure 59 shows the situation.

Situation of uncertain costs

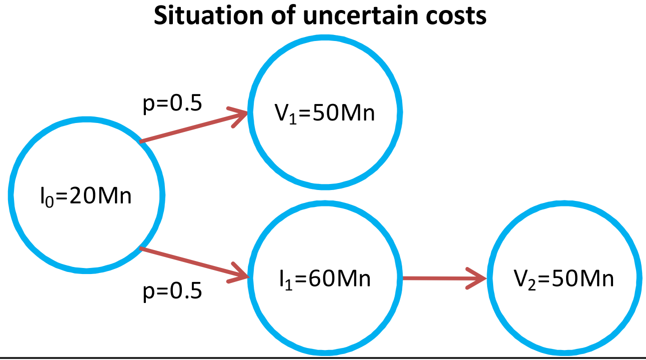


Figure 59; Investment opportunity with option to stop. Investments may require additional payment 60.

The NPV of this situation is;

$$\text{NPV} = \begin{cases} -20 + \dfrac{50}{1.1} \\ -20 - \dfrac{60}{1.1} + \dfrac{50}{1.1^2} \end{cases} = -3.88.$$

Conventional investment theory explains that the investor withdraws the opportunity because the NPV is negative. But, this approach neglects the opportunity to abandon the project. After initial investments the investor can abandon the project if costs increase. Then, the NPV of the initial situation is;

$$\text{NPV} = \begin{cases} -20 + \dfrac{50}{1.1} = 2.73. \\ -20 + 0 \end{cases}$$

This proves that an initial investment is profitable. The full opportunity cost of capital is 20+2.73. That equals the expected payoff and investments should take place. If costs turn out higher after 1 period of investing, the investor should stop the project[5]. A rationale explanation for this approach is that expected costs do not reflect the situation correctly. The investor faces less than expected costs because of the opportunity to abandon.

In reality the investor has a longer investment period. During investments, both costs and payoff change. For some combination of costs, option value and payoff, the full costs of investing have become too large. Investing is no longer profitable and the investor should stop the project. Investing may become profitable again by waiting for new information that increases the value or lowers costs. If investing is no longer profitable, the investor should terminate his investment.

[5] This approach assumes no costs to abandon. If costs to abandon are larger than 6, initial investment is not profitable.

2 Constructing the Pindyck (1993) model

Construction projects also face uncertainty that make investors postpone or (temporary) stop investments. In most studies of investment under uncertainty it is the payoffs from the investment that are uncertain. But for most construction projects it is the costs that are more uncertain than the payoff. Investors largely know the value prior to construction but are uncertain of the costs. Other examples of investments under cost uncertainty relate to large petrochemical complexes or R&D projects developing new medicines. Moreover, much of the costs to acquire the new assets are irreversible. The expenditures are sunk costs that cannot be recovered as the investment, ex post, has shown to be a bad one. As the costs are prone to uncertainty investors may wait for new information before spending.

The uncertainty of these costs is two-fold. First there is technical uncertainty which relates to the physical difficulty of completing the project. If the prices of the construction inputs are known, how much input will be used? The technical uncertainty is only known after completion. Second there is input uncertainty. This relates to uncertainty of changing input costs. Change results from economic cycles and is more uncertain the farther one looks into the future. These costs change regardless from whether or not the firm is investing. This is opposite to technical uncertainty which only changes if the firm invests. Projects that take considerable time to complete are prone to input uncertainty.

This chapter uses a continuous model that incorporates both types of uncertainty. The use of a continuous approach is straightforward. In the examples of Chapter 1 the costs and payoff change discrete. A more realistic approach is to increase the step size and prolong the investment period. But the calculations become rather unclear as the possible outcomes grow every step n by 2^n. Therefore I use a continuous approach that Pindyck (1993) developed. A first analysis studies the continuous solution for a certain investment situation, thereafter uncertainty is introduced.

2.1 Developing the Pindyck (1993) model

Consider a project with random cost to completion K and (fixed) value V. The use of cost to completion implies investors initially face costs K and during construction the costs decline to 0. Upon completion (i.e. $K = 0$), the firm receives the asset whose value V is known with certainty. During the time to completion, the firm's rate of investment is k. If no uncertainty would be present, the value of the option is easy to determine. The project is completed at time $T = {}^{K}/_{k}$ and the value of the option is therefore;

$$F(K) = max\left[Ve^{\frac{-rK}{k}} - \int_0^{K/k} ke^{-rt}dt, 0\right].$$

Figure 60 shows the value of the option to invest for random costs K. The option value has a slightly curved shape as the investment and received asset are discounted at the risk free rate $r = 0.05$. Without discounting, the option value would have a straight line.

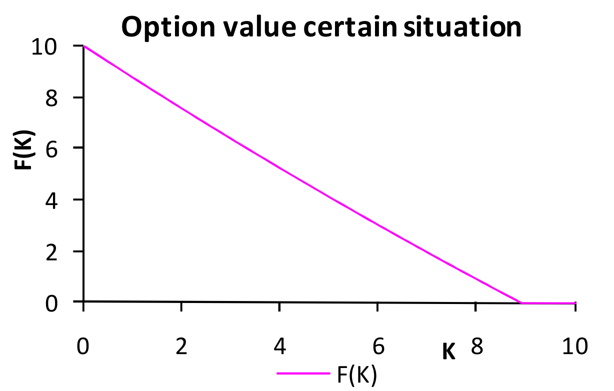


Figure 60; Option value without uncertainty for varying K, V=10, k=2 and r=0.05

At some point the option no longer has a value. That point reflects the critical costs K^* and investing is no longer profitable. Upon K^* the initial costs plus the option value do not exceed the value of the received asset (i.e. $K + F(K) \leq V$) and investing should take place. For costs larger than K^* the firm should withdraw or abandon the investment. The critical costs K^* is

$$K^* = \left(\frac{k}{r}\right) log\left(1 + \frac{r * V}{k}\right).$$

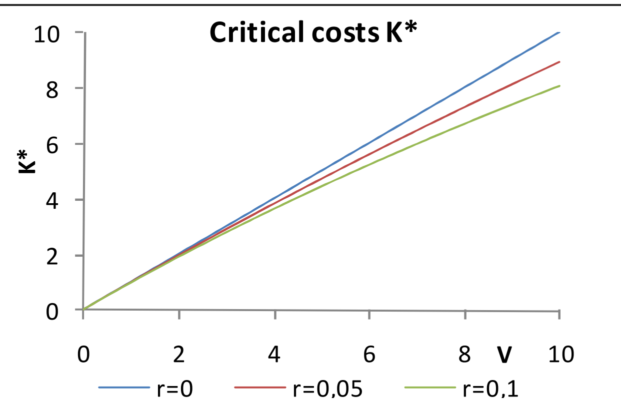


Figure 61; Critical costs K* plotted against V for different values of r with k=2.

Figure 61 shows the critical costs as a function of V for different values of r with constant investment rate $k = 2$. For $r = 0$, the critical costs have value V. But for $r > 0$ critical costs decline for the increasing discount rate. The continuously discounted investments between $t = 0$ and $t = T$ and the discounted asset received at $t = T$ lower the critical costs K^*.

Having obtained a solution for the situation under certain conditions, it is now time to study the option value for uncertain circumstances. Let's consider again the cost to completion K. A noise term is added to the cost to completion that makes its behaviour stochastic. The incremental change dK of the cost to completion (given by K) is

$$dK = -Idt + g(I, K)dz.$$

The investor faces initial costs K, like the certain situation. Upon completion costs are 0 and the investor receives the asset with value V. The expected change of costs every time step is $-I * dt$. Note that for the certain situation dK only changes by the expected investment without any noise term. The term $g(I, K)dz$ introduces uncertainty. dz is the increment of a Wiener process that introduces the stochastic behaviour. The Wiener process makes the error term for each step dK normally distributed, independent and stationary[6]. The function $g(I, K)$ determines the magnitude of error and is dependent on I and K. Moreover, the function $g(I, K)$ introduces technical and input uncertainty to the process. Technical uncertainty is only present if investments take place and both $g(0, K) > 0$ and $g_I > 0$ should hold. For any cost K the function dK behaves stochastically and therefore $g(0, K) > 0$. But technical uncertainty is only present if investments take place. This implies that the first derivative of g to I should (i.e. g_I) be larger than 0. If costs change independent from I, it is called input uncertainty. Then only $g(0, K) > 0$ should hold.

The stochastic behaviour makes the time to complete the project $\check{T}$ stochastic as well. The combination of randomly fluctuating costs and a constant investment rate I lengthen or shorten the construction period. If dK is positive, the costs to completion accumulate and lengthen the investor's investment period. The investor continuously invests at the maximum rate k between $t = 0$ and $t = \check{T}$ and receives the asset with (discounted) value V after completion. This makes the option value;

$$F(K) = maxE_0\left[Ve^{-\mu\check{T}} - \int_0^{\check{T}} I(t)e^{-\mu t}dt, 0\right].$$

Here μ is an appropriate risk adjusted discount rate and depends on the investor's risk aversion. In order to determine optimal investment and stopping rules, the stochastic diffusion process requires any additional structure;

i. $F(K;\ V, k)$ should be homogeneous of degree one.

ii. $Fk < 0$, this implies that the value of our option declines as costs increase.

[6] For more information on Wiener processes and Brownian motion consider Baxter and Rennie (1997).

iii. The variance of dK is bounded and approaches zero if K goes to 0.

iv. The expected cost to completion is given by $K = E_0 \int_0^{\tilde{T}} I(t)dt$

Based on these conditions the stochastic component is;

$$g(I, K) = \beta K(IK)^{\alpha} \text{ for } 0 \leq \alpha \leq {}^{1}\!/_{2}.$$

This study only considers two extreme values of α; $\alpha = {}^{1}\!/_{2}$ and $\alpha = 0$. Both correspond to the two identified types of uncertainty. If only technical uncertainty is present, $\alpha = {}^{1}\!/_{2}$ should hold. K only changes if investments are taking place and the variance of ${}^{dK}\!/_{K}$ changes linearly with ${}^{I}\!/_{K}$. If $\alpha = 0$ the costs change irrespective of what the firm does and the variance of ${}^{dK}\!/_{K}$ is constant and independent of I. These two situations are joined in one equation;

$$dK = -Idt + \beta\sqrt{IK}dz + \gamma K dw.$$

Here, β and γ are coefficients that are determined later and both dz and dw are increments of uncorrelated Wiener processes. This entire model combines uncertainty over the required input to complete the project, uncertainty over the cost of that effort and uncertainty over the time to complete the project. The stochastic process dK combines the first two types of uncertainty. The uncertainty over the time to complete is a consequence of uncertain costs and the constant investment rate. If dK is positive, the costs to completion accumulate and lengthen the investor's investment period. A natural consequence is increasing costs and declining profit.

Figure 62 shows the behaviour of the model. Initially the investor starts investing at expected costs K until completion where $K = 0$. The investment rate I and discount rate r determine the time span from $K = K_0$ to $K = 0$. During construction (between $K = K_0$ and $K = 0$) the investor invests continuously at rate I. The horizontal line in the upper right graph shows the invested costs for each individual project. The line is slightly curved by the discounting of the investment rate. Upon completion the investor stops and receives the (discounted) asset. The lower left graph of Figure 62 summarizes the simulations as it shows the payoff from the invested costs and received asset over time. A greater time span to completion drives increasing costs and lowers profitability. Note that the option value is similar to a financial put option. A put option gives the investor the right to sell a share at a certain value. Here, its interpretation is that the investor issues a contract worth costs K to obtain an asset with value V.

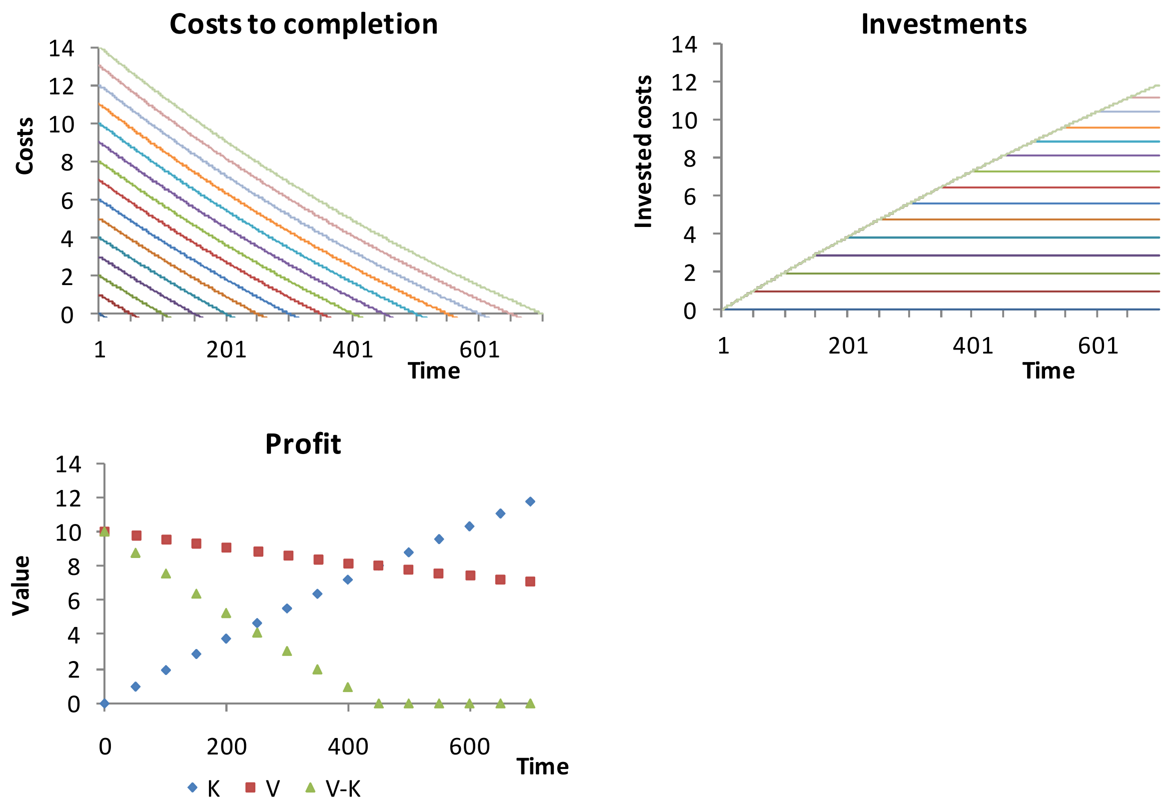


Figure 62; Cost to completion, invested costs and project payoff of projects for K_0=0 to K_0=14, I=2, r=0.05, V=10 and β=0.

Figure 62 shows the impact of varying costs on the option value under certain circumstances. The diffusion term of the stochastic differential equation (i.e. $\beta\sqrt{IK}dz + \gamma K dw$) has a similar impact on the option value. The term shortens or prolongs the construction costs that drives invested costs and profits. Every time step dK the costs change fixed by the drift term and stochastically by the normally distributed error term. Here, $\beta\sqrt{IK}$ and γK are the magnitude of the error term. The left graph of Figure 63 shows the stochastic behavior. The change from K_t to K_{t-I} reflects expected change. The steps from K_{t-I} to K_{t+dt} show the stochastic behavior (normally distributed with magnitude $\beta\sqrt{IK}$ and γK). This shows the change of costs is composed of three steps. The right graph of Figure 63 shows the cost paths under certain and uncertain circumstances. The thick line marks certain conditions. The other irregular lines show cost development under uncertainty. Uncertainty progresses or delays construction. The results of Figure 62 imply that delay enlarges invested costs and reduces profit.

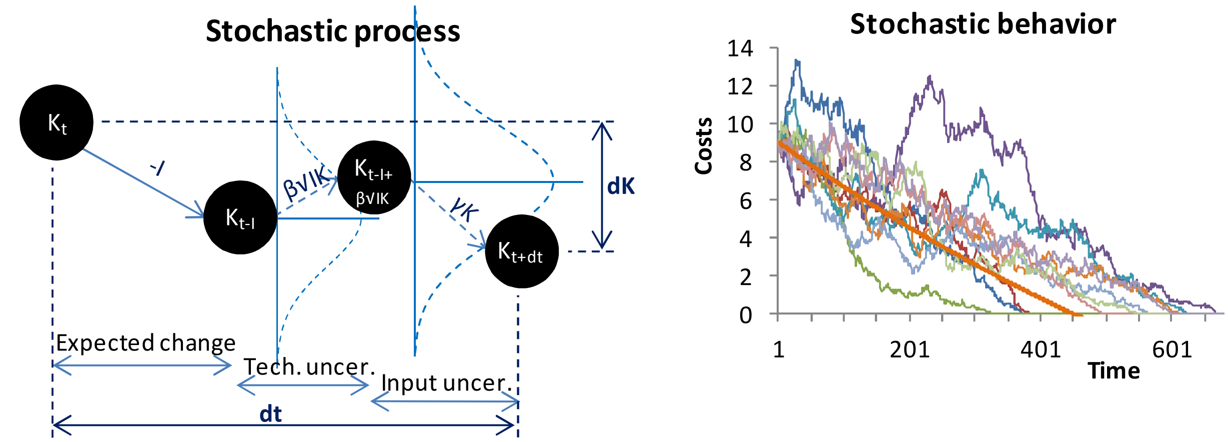


Figure 63; Impact of stochastic behavior. The right graph shows simulations of 25 projects for initial costs K=9, β=0.63 and r=0.05 and one simulation for β=0.

All uncertain simulations in the right graph of Figure 63 start at initial costs $K = 9$ and pay off $F(K) = \left[Ve^{-\mu\check{T}} - \int_0^{\check{T}} I(t)e^{-\mu t}dt, 0\right]$. When more simulations are used (all starting at $K = 9$) the average of $F(K)$ converges to a value. For other initial costs K the option also converges to some value F dependent on K. Thus the option value approaches some function $F(K)$ for given parameters r, β and γ. In fact, the exact solution is known; it is a differential equation. One obtains the differential equation by constructing a portfolio that consists of the investment opportunity (long position) and an asset that spans the value of the costs (short position). Section 1.1 already explained the approach to obtain the option value of a discrete example. This approach still holds, but continuous time obliges some changes. The asset x that spans the value of the portfolio is denoted by a Wiener process (like dK) and fully correlated with dw[7]. This implies

$$dX = \alpha_x x dt + \sigma_x x dw.$$

The appendix shows the derivation of the differential equation. Here, the closed form solution of the option F dependent on initial costs K is given;

$$\frac{1}{2}\beta^2 IKF_{kk} + \frac{1}{2}\gamma^2 K^2 F_{kk} - IF_k - \varphi\gamma KF_k - I = rF.$$

The differential equation introduces φ that relates $F(K)$ with the market. This implies $\varphi = \frac{(r_x - r)}{\sigma_x}$. By the CAPM, the return of the asset x is given by $r_x = r + \theta\rho_{xm}\sigma_x$ where θ is the market price of risk and ρ_{xm} is the correlation of x with the market portfolio (i.e. β of the CAPM). Thus $\varphi = \theta\rho_{xm}$. This research does not study correlation with the market. Pindyck (1993) showed the correlation has a negligible difference on the critical costs for small φ.

[7] Full correlation with dw is required as the asset replicates all values of the investment opportunity.

The first derivative of F to I shows that the investment rate I maximizes $F(K)$ for;

$$I = \begin{cases} k & \text{for } \frac{1}{2}\beta^2 K^* F_{kk} - F_k - 1 \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

This condition already reveals a boundary condition to determine the critical costs K^*, that are denoted with a (*). One should invest at rate k to achieve an optimal value for $F(K)$ as long as $\frac{1}{2}\beta^2 K^* F_{kk}(K^*) - F_k(K^*) - 1 \geq 0$ holds. More structure is required to solve the differential equation and find a solution for $F(K)$. For the optimal stopping rule $F(K)$ should meet;

$$F(0) = V,$$

$$\lim_{K\to\infty} F(K) = 0 \text{ and,}$$

$$\frac{1}{2}\beta^2 K_S^* F_{kk}(K^*) - F_k(K^*) - 1 = 0.$$

The other (i.e. first) two conditions have a straightforward meaning. For K is zero there are no costs, the net pay-off is the value V of the asset and the option should have value V. The second condition implies that if K is very large, the probability is very small the project becomes profitable over some finite time. The last equation is derived from the condition for I. For some critical value K^* the investor should stop investing. The condition is also known as a smooth pasting condition that $F(K^*)$ be continuous at K^*. If $F(V)$ were not continuous and smooth at the critical exercise point K^*, one could do better by exercising a different point.

2.2 Solution characteristics of Pindyck (1993)

Combining the differential equation and the different boundary conditions shows a solution for the identified types of uncertainty. The solutions of both types of uncertainty are analyzed separately. First technical uncertainty (i.e. $\beta > 0$ and $\gamma = 0$) is examined and thereafter input uncertainty (i.e. $\beta = 0$ and $\gamma > 0$). Each solution is obtained by simulation and numerical analysis. Similarity of both answers proves the solutions are correct and enables to use the most accurate solution. The scripts to obtain a simulated and numerical solution are added in the appendix.

With only technical uncertainty present the differential equation reduces to

$$\frac{1}{2}\beta^2 IKF_{kk} - IF_k - I = rF.$$

An analytical solution is obtained for $\beta = 0$. For $\beta > 0$ the differential equation does not have an analytical solution. Numerical analysis or simulations approximate the solution. This approach requires to specify initial values. Most values can be picked randomly, but β requires some additional analysis. This Part A uses, like Part B, variability (i.e. σ/μ) to describe technical uncertainty. The variance of the cost to completion for only technical uncertainty is

$$var(K) = \left(\frac{\beta^2}{2-\beta^2}\right)K^2.$$

This study analyses variability of 25% and 50%. This implies that one standard deviation of the project is 25% or 50% of the expected costs. That makes values of β are $\beta = 0.343$ and $\beta = 0.63$ by solving $\frac{\sigma}{\mu} = \sqrt{\frac{\beta^2}{2-\beta^2}}$. Figure 64 below shows values for $F(K)$ as a function of K with $V = 10, I = 2, r = 0.05$ and selected values of β.

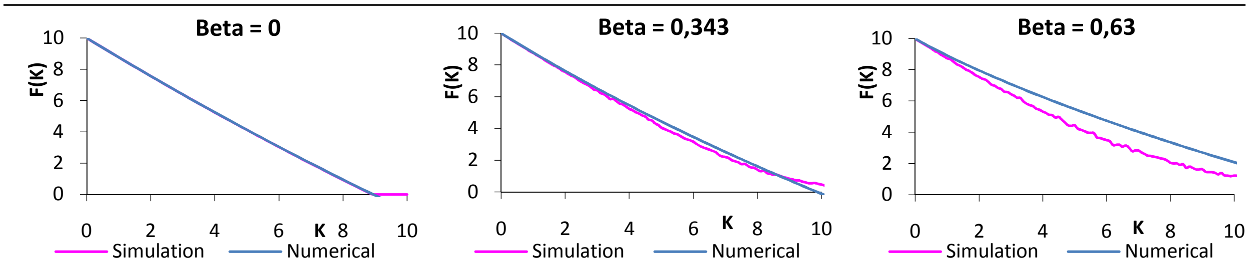


Figure 64; Option value $F(K)$ numerical and simulation results for different values of β. Other values are V=10, I=2 and r=0.05.

The first left graph shows the option value for no uncertainty. Both the numerical and the simulation solutions are identical in the absence of uncertainty. The discount rate r somewhat curves the graph. Initially the option value is $F(0) = V$ as no costs leave payoff V. For increasing values of K the investment horizon prolongs and the option is eventually no longer profitable.

Having incorporated uncertainty, the value of the option has a similar path as the option value for $\beta = 0$. Initially, the option has the same value as within the certain situation. It always holds that for no costs the payoff is the asset. But for increasing cost the option longer retains its value. A few projects remain profitable because uncertainty is in favour of these projects. The uncertainty shortens the construction time and thereby increases profits. For growing uncertainty more projects are profitable that drives increasing profit. This proves volatility increases the option value.

The figures also show the difficulty to determine the exact solution of the option value under uncertainty. The numerical and simulated solutions show a considerable difference. It is especially difficult to model the option's tail correctly. The numerical method has difficulties to approximate the curvature of the option value. Changing the step size or using a different script may improve the approximation. Here a first order Euler approximation was used with an accuracy of order O(h). The simulated solution starts to move up and down for increasing costs. Less and less simulations have a profitable payoff for increasing costs. And precisely the profitable projects determine the option value. The line is initially straight but starts to move for larger costs. To obtain a more accurate solution many simulations should be run for every start value K and the time step should be small. The comparison points out that the simulation results show a better approximation of the option value. Therefore the numerical solution will not be considered further.

2.2.1 Solutions under technical uncertainty

Knowing the option value allows to determine the optimal points of investing. Section 2.2 defined investors should stop investing if $\frac{1}{2}\beta^2 K^* F_{kk}(K^*) - F_k(K^*) - 1 \geq 0$ no longer holds. Figure 65 shows the option value for the different types of uncertainty. The right graph is a close-up of the left graph. The outcomes shows that a larger uncertainty drives a larger option value. The application of the stopping rule determines the critical costs to stop investing (i.e. K^{*}[8]). The figure shows that increasing technical uncertainty enlarges the critical costs. Technically the optimal stopping rule shows that a larger β enlarges critical costs because F_k is positive and F_{kk} is negative. A rational explanation is that technical uncertainty is only discovered after spending (in contrast to input uncertainty that investors know in advance). The investor may always abandon if additional payments are required. Investing under technical uncertainty may therefore continue longer.

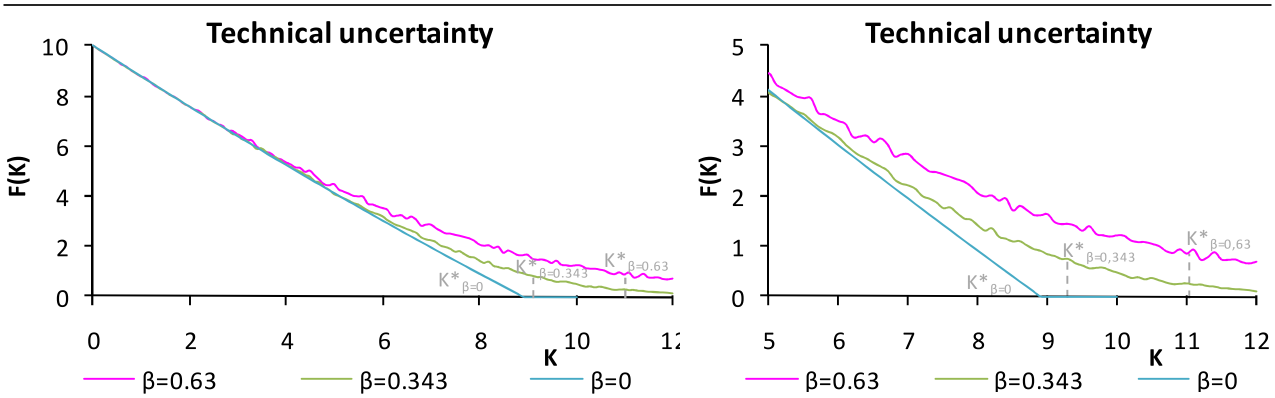


Figure 65; Option value F(K) under technical uncertainty for different values of β. Other values are V=10, I=2, r=0.05. Critical costs to stop enlarge for increasing technical uncertainty

Additional uncertainty increases critical costs progressively. The magnified uncertainty from $\beta = 0.343$ to $\beta = 0.63$ almost triples the difference with the critical costs for $\beta = 0$. Upon critical costs the option still has a value (except for the certain situation).

[8] Note that these critical costs approximate of the exact costs. The simulation models make it difficult to determine an exact solution. The appendix shows that especially the second order derivative is subjected to noise.

2.2.2 *Solutions under input uncertainty*

Having obtained a solution to investing under technical uncertainty, the research now examines optimal investment rules for input uncertainty. At first, the analysis reviews the differential equation that gives the option value. Thereafter this study examines optimal investment rules.

The differential equation stating the option value reduces for only input uncertainty to

$$\frac{1}{2}\gamma^2 K^2 F_{kk} - IF_k - \varphi\gamma K F_k - I = rF.$$

The earlier defined boundary conditions still hold for input uncertainty. But the last condition of the stopping rule changes from $\frac{1}{2}\beta^2 K^* F_{kk}(K^*) - F_k(K^*) - I = 0$ to $F_k(K^*) = -1$ because β is now 0.

Figure 66 shows the simulation results for the different types of input uncertainty. In general, input uncertainty lowers critical costs. This behavior contrasts with critical costs to stop investing under technical uncertainty. Under technical uncertainty critical costs enlarge for increasing uncertainty. This contrasting behavior relates to the nature of the different types of uncertainty. Technical uncertainty only reveals after investments have taken place. Input costs behave independent from investments and are known beforehand. Optimal investing may continue longer for technical uncertainty than for input uncertainty. This proves critical costs to stop differentiate between the different types of uncertainty.

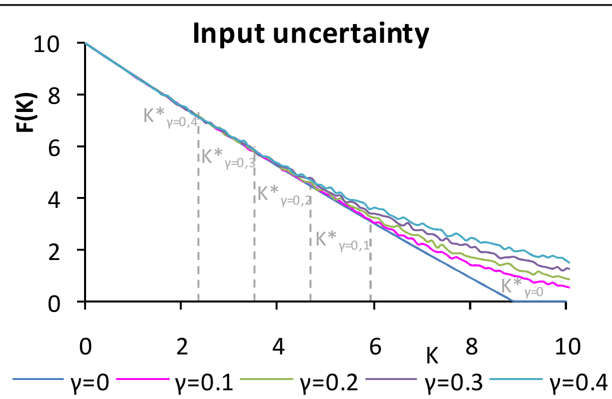


Figure 66; Option value under input uncertainty varying from γ=0 to γ=0.4. Critical costs to invest and abandon investments are equal for input uncertainty. The costs to stop behave contrasting from technical uncertainty.

The declining critical costs for increasing input uncertainty also have a rational explanation. A (logical) consequence of increasing uncertainty is a demand for larger profit. A greater profit can only be achieved through decreasing costs if the value is assumed constant. This is exactly what we observe; declining critical costs for increasing uncertainty.

Above that, the inclusion of additional input uncertainty has a different impact on critical costs from technical uncertainty. The impact of additional 0.1 input uncertainty decreases critical costs less and less. This is opposite for technical uncertainty that enlarges critical costs progressively. This proves input uncertainty shows the contrasting behavior from technical uncertainty. But where increasing input uncertainty has a declining impact on critical costs, the stand-alone impact on critical costs is relatively large. The critical costs for $\gamma = 0$ almost triple critical costs for $\gamma = 0.3$. Considering technical uncertainty, the critical costs for $\beta = 0.3$ only increase by 10% from the critical costs for $\beta = 0$.

2.3 Conclusions

The analysis of this chapter allows drawing two conclusions. First, input and technical uncertainty have a contrasting behaviour for optimal stopping costs. Technical uncertainty enlarges critical costs to stop. Investors face technical uncertainty only after investments and always have the opportunity to terminate investments. Moreover, the inclusion of additional technical uncertainty progressively enlarges costs. On the other hand, input uncertainty reduces critical costs. Input uncertainty is known beforehand. Additional uncertainty has a decreasing impact on critical costs to stop.

Second, critical costs are particularly sensitive for input uncertainty while determining the optimal point to stop. Technical uncertainty has a negligible effect on critical costs with respect to the strong effect of input uncertainty.

3 Inclusion of shocks

This third part adds shocks to the model and reconsiders the optimal point of investment. Investors face unforeseen cost increases at most construction projects. Every individual shock is a rare 'black swan' or 'white elephant', but shocks as such happen at most projects and largely contribute to costs of infrastructural projects (Vrijling and Boschloo (2001))[9]. Examples of shocks relate to (fatal) accidents, deficient designs or miscalculations. Shocks are of great importance as they contribute to both expected costs and cost variance. In fact, shocks mainly drive project variability instead of the noise of individual line items (Part A). Moreover, the probability density function takes a skewed form by the incorporation of shocks. This effect relates to the discontinuity and the unidirectional behaviour of shocks. Shocks only cause cost increase instead of up- and downward movement. All in all shocks improve the model incorporating a more realistic cost development.

This third part has the same structure as the second part. First it (re)considers the model. Thereafter an analysis studies and discusses the results for the different types of uncertainty. This provides a solid foundation to review an investment in chapter 4.

3.1 Reconsidering the model

The occurrence of shocks is related to technical uncertainty. It enlarges the (physical) difficulty to complete the project. Only after completion investors know how much input was required to obtain the asset. Based on experience investors may expect shocks to happen, but this is not certain at all. By including shocks technical uncertainty is now composed of shocks and continuous behaviour.

The inclusion of shocks transforms the stochastic differential equation to;

$$dK = -Idt + \beta\sqrt{IK}dz + \gamma K dw + \kappa J.$$

Here, J is a Poisson distributed arrival process and κ is the value of the occurring shock. The operation of the model does not change. Upon $t = 0$, the investor faces expected costs K_0 and after completion $t = \check{T}$ (which is stochastic) the investor receives the asset with discounted value V. The option value (or payoff) of every simulated project is;

$$F(K) = max E_0 \left[V e^{-\mu \check{T}} - \int_0^{\check{T}} I(t) e^{-\mu t} dt, 0 \right].$$

[9] The terminology 'black swans' or 'white elephants' expresses the small probability of occurring. Despite all preparation the events still occur. For more information, Kahneman and Tversky (1974) studied biases of expert judgment under uncertainty.

This approach combines technical uncertainty, input uncertainty and shocks of project costs. Figure 67 further clarifies the stochastic behaviour. The change of K from K_t to $K_{t-I+\beta\sqrt{IK}+\gamma K}$ equals stochastic behaviour in Figure 63. It shows the expected change, technical and input uncertainty. The last change (i.e. from $K_{t-I+\beta\sqrt{IK}+\gamma K}$ to K_{t+dt}) shows the shock term. For some probability a shock with magnitude κ occurs. Otherwise there is no change of costs.

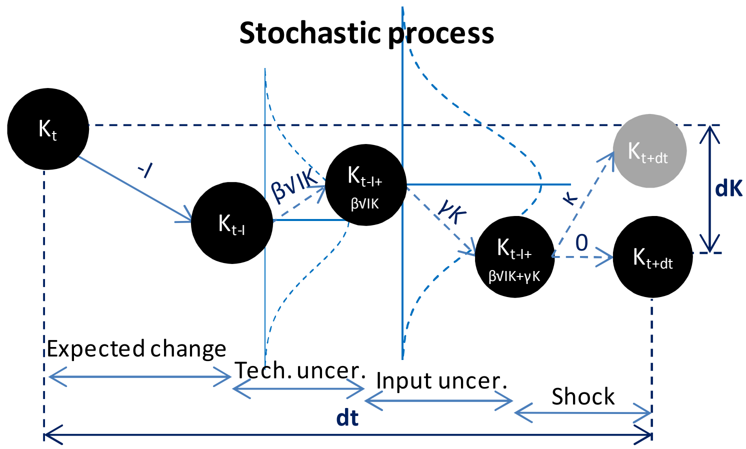


Figure 67; Stochastic process including shocks

The inclusion of shocks increases the costs to completion K and enlarges uncertainty (it is not sure shocks occur). Above that, shocks change the shape of the probability density function (PDF). Technical and input uncertainty introduce symmetrically up- and downward movements. Shocks only enlarge costs to completion. There is no downward movement.

All uncertainty explained in Figure 67 introduces variability on the time to completion, the firm's invested costs and profitability. Figure 68 further clarifies the impact of shocks on the entire model. The occurrence of shocks lengthens construction time as shown in the left upper graph. The project is completed the fastest if no shocks occur. Then, the invested costs are the lowest and the project is most profitable. Also remark that the shapes of the graph showing investments and profit did not change. Invested costs cut off at a different point, but the shape is indifferent from Figure 62. The lower graph shows the option has the shape of a put option.

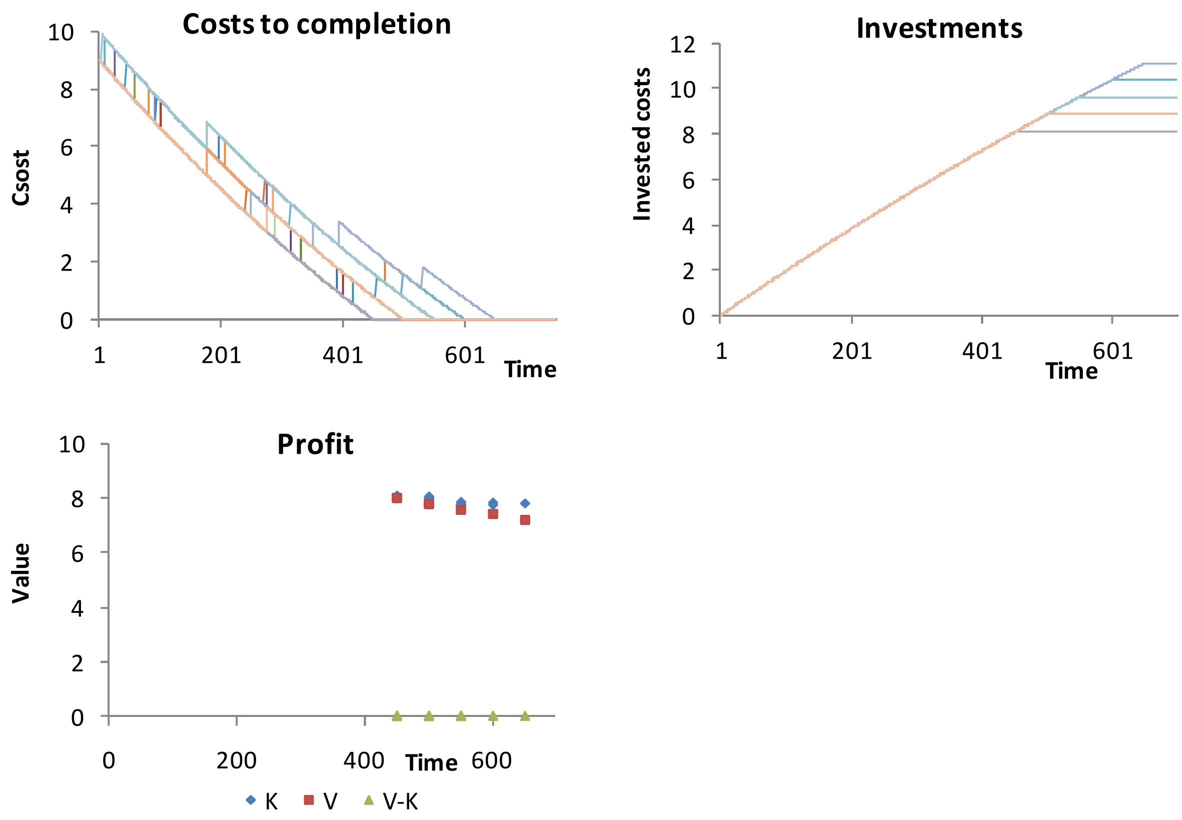


Figure 68; Inclusion of shocks and model response for K_0=9, I=2, β=0, γ=0, κ=0.1*K_0 and λ=K_0/I.

Chapter 2 derived an exact solution of the option that was given with a differential equation. The development of (again) a closed form solution is difficult by the inclusion of discontinuities (like shocks). The development requires the use of different mathematics. Recent studies by the Delft University of Technology describe a method to solve the stochastic differential equation under shocks[10]. This study restricts to investigate optimal investments under cost uncertainty. The (required) profound understanding of mathematics limits to study the simulation results and determine optimal investments under cost uncertainty.

This study uses the previously defined boundary conditions to determine $F(K)$;

$$F(0) = V,$$

$$\lim_{K\to\infty} F(K) = 0 \text{ and,}$$

$$\frac{1}{2}\beta^2 K_S^* F_{kk}(K^*) - F_k(K^*) - 1 = 0.$$

[10] Fang (2010) defines the COS method to solve the stochastic differential equation and come to a closed form solution. This method is based on the idea to replace the probability density function, appearing in the risk-neutral valuation formula, by its Fourier-cosine series expansion.

These conditions give structure to the solution and define the optimal point of investing. The first two conditions define that for no costs the payoff is the asset value V (thus $F(0) = K$) and for large costs it is not likely the net payoff has a value (thus $\lim_{K\to\infty} F(K) = 0$). The last condition defines the critical costs if $\frac{1}{2}\beta^2 K_S^* F_{kk}(K_S^*) - F_k(K_S^*) - 1 \geq 0$ no longer holds. Shocks presumably have no influence on the terms of the exact solution that includes the investment rate I. The first derivative to I fades all terms out that do not include the investment rate. This feeling drives the assumption that conditions do not change. Additional research is recommended to determine the exact solution and review the accuracy of the assumptions.

3.2 Solution characteristics under shocks

The shock process requires any additional structure in order to value the option that includes shocks. First consider the arrival process. Studies show investors should expect 1 to 2 shocks during construction. Increasing complexity of projects enlarges probability of shocks to occur during construction. This research assumes shocks to arrive according to a Poisson distribution. That makes the expected amount of shocks during construction *n*

$$n = \lambda T_0.$$

Here λ is the arrival rate of the shocks and $T0$ is the initially expected time to completion. In fact, T_0 is $T_0 = \frac{K_0}{I}$ which relates λ and n by $\lambda = \frac{n*I}{K_0}$. For now, this study assumes $n = 1$. Next, consider the magnitude of the shock. Vrijling and Boschloo (2001) showed the magnitude of the shock is 10% of the initial construction costs. Therefore I assume $\kappa = 0.1 * K_0$.

Shocks do not only enlarge expected costs but also introduce additional uncertainty. Here, only the number of shocks during construction introduces uncertainty[11]. By definition of the Poisson process, the shocks' variance equals the expected value. Multiplying this variance by the scalar $0.1 * K_0$, makes the total variance of the system;

$$Var(shocks) = (0.1K_0)^2 \lambda T_0.$$

Figure 67 showed the impact of shocks on the cost to completion. Shocks enlarge cost to completion, increase the variance and change the density function of the firm's investments. Figure 69 shows the impact of shocks on the firm's investments. Both figures relate to the upper right graphs of Figure 62 and Figure 68 by showing the probability density function (PDF) of the invested costs. Figure 69 proves shocks increase expected costs, change variability and shape of the invested costs. The

[11] This study does not analyze variability of the magnitude of shocks. The variability of the magnitude has a negligible impact on project variability (Vrijling (2003)).

inclusion of shocks increases expected costs to completion. The increased costs to completion lengthen the time to invest and thereby enlarge invested costs. Figure 69 also shows the impact of the discount rate on construction costs. The PDF changes by the influence of discounting. Larger costs relate directly to a longer construction and investment period. That magnifies the impact of discounting and reduces larger costs increasingly.

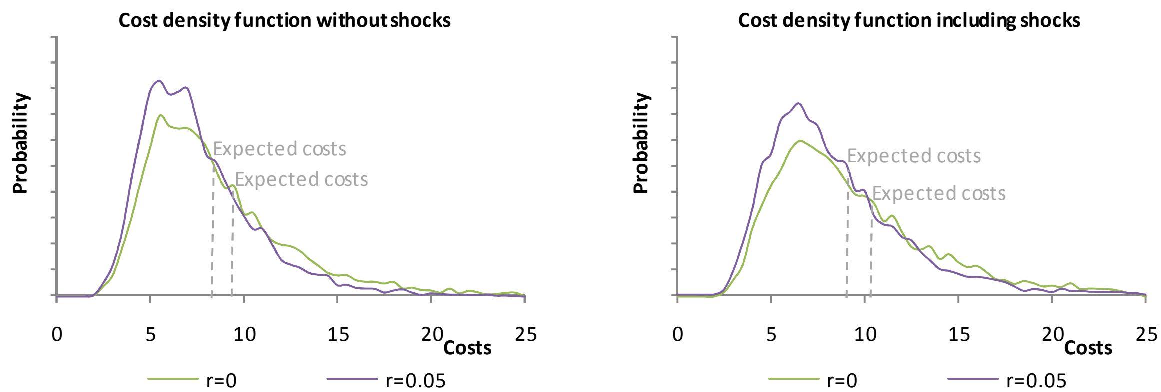


Figure 69; Impact of shocks on the cost density function for K_0=9, I=2, β=0.63, γ=0, κ=0.1*K_0 and λ=K_0/I.

3.2.1 *Solutions of technical uncertainty*

Figure 70 shows the impact of shocks on the option value for technical uncertainty. The graphs show the option value with and without shocks to clarify the effect of shocks. The lower right graph gives an overview for the option value and critical costs under technical uncertainty and shocks. In general, shocks devalue the option. Enlarged expected costs decline profit and the option value as well. For increasing costs the option value converges towards the initial situation. Simulations of profitable projects determine the option value. Shocks prolong the time to completion and increase invested costs. For large costs this implies that a project is only profitable if no shocks occur. As a result the option value converges to the option value without shocks. The situation for $\beta = 0$ shows the option no longer has a value above initial costs $K_0 = 9$ like the option value without shocks. Figure 68 shows the model behavior for $K_0 = 9$ that affirms the analysis.

Shocks reduce the critical costs of optimal investments. Shocks enlarge costs and cost uncertainty that reduces the option value. The absolute cost increase reduces expected payoff by the magnitude of the shock. But the additional uncertainty declines critical costs even further. Investors demand a higher pay off on more uncertain projects. Critical costs should decline to if the received asset has fixed value to ensure a larger pay off. The inaccuracy of the simulations makes it impossible to state the exact cost decline of critical costs. But the decline is the magnitude of the shock plus a premium for additional uncertainty. The analysis in Chapter 4 shows the investment criterion (i.e. $\frac{1}{2}\beta^2 K_S^* F_{kk}(K_S^*) - F_k(K_S^*) - 1 \geq 0$) is heavily prone to uncertainty.

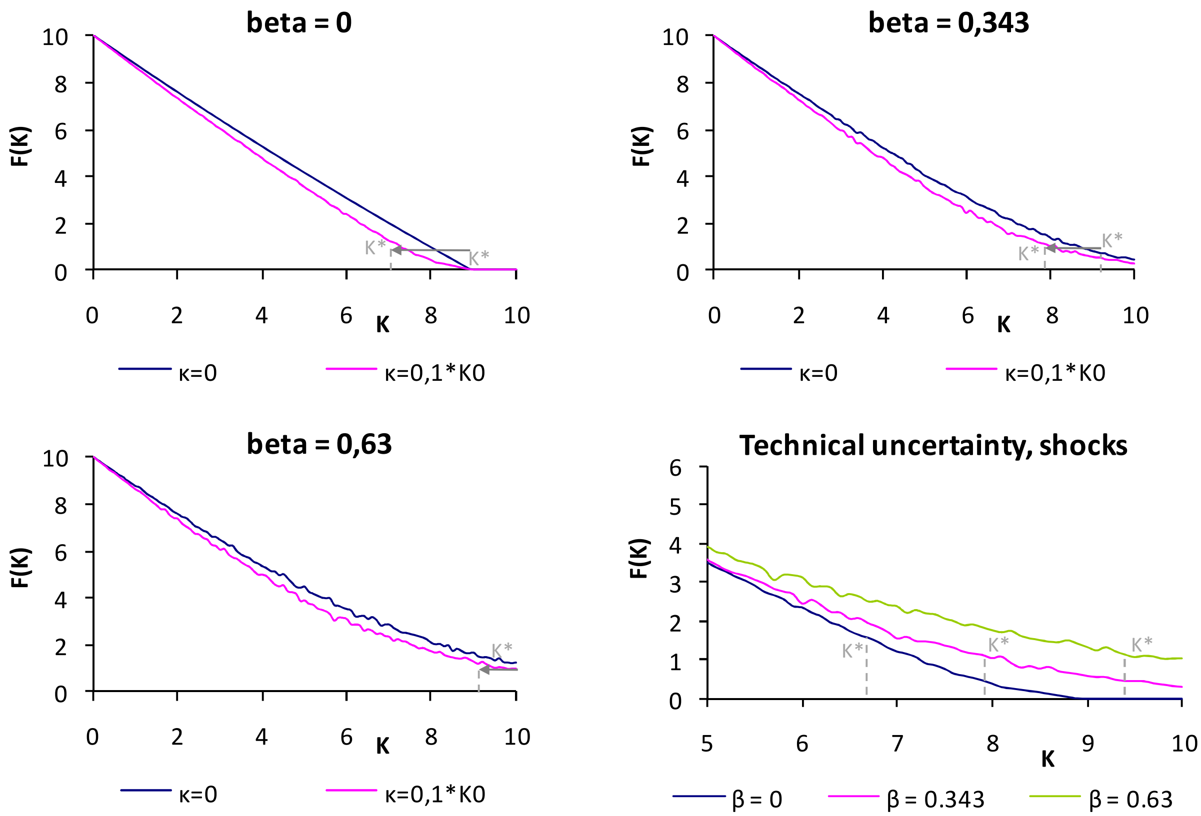


Figure 70; Option value F(K) for technical uncertainty and shocks. Other values are V=10, I=2 and r=0.05

3.2.2 Solutions of input uncertainty

Figure 71 shows shocks also decline the option value for input uncertainty like for technical uncertainty. The lower right graph gives an overview of the option value and critical costs under input uncertainty and shocks. The increase of expected costs declines the option value. Critical costs also decline because of additional costs and uncertainty makes investing more expensive. Investors only invest in profitable projects and demand a compensation for additional risk exposure. Additional profit is gained by declining costs as the asset value is fixed. Above that increasing uncertainty enlarges the full opportunity cost of capital and lower critical costs reflect a large(r) option value.

Figure 71 also shows the large impact of shocks on the option value. For all critical costs under input uncertainty, the critical costs reduce by approximately 1/3. Moreover the reduction of critical costs by the shocks is quite similar for all types of uncertainty. Note that the critical costs under input uncertainty reduce for increasing uncertainty.

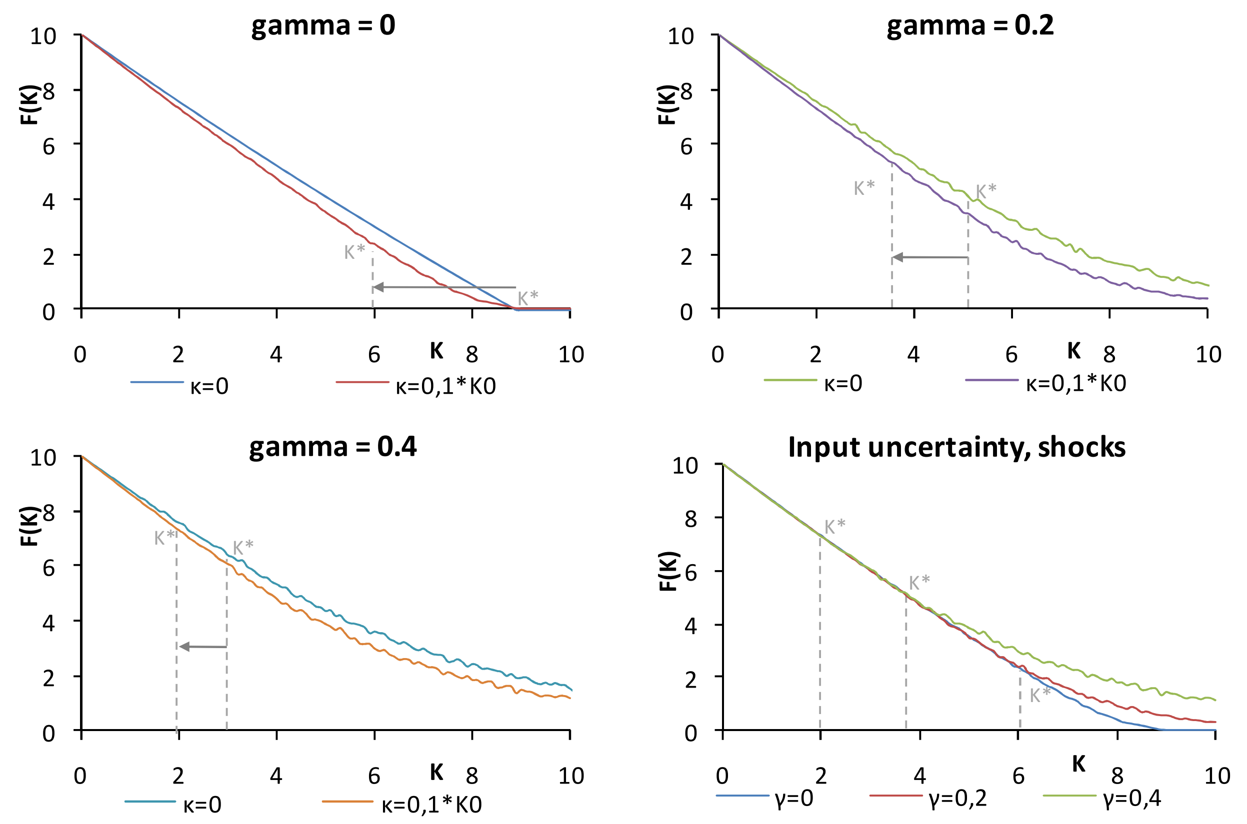


Figure 71; Option value F(K) for input uncertainty and shocks. Other values are V=10, I=2 and r=0.05

Figure 72 shows the sensitivity of the option value for different arrival rates λ and shock magnitudes κ. For increasing κ, the option value declines because κ magnifies cost to completion. That increases investments and lowers payoff. The result is a lower option value. An increasing λ has a similar impact. A larger arrival rate increases the occurrence of shocks during construction. More shocks lower the option value. Shocks prolong construction time, enlarge investments and lower profitability. For all values of λ and κ, options converges to the option value with no shocks. Simulations that obtain a profitable payoff determine the option value. For large initial costs projects only remain profitable if no shocks occur during construction. But the increasing λ and κ lower all critical costs.

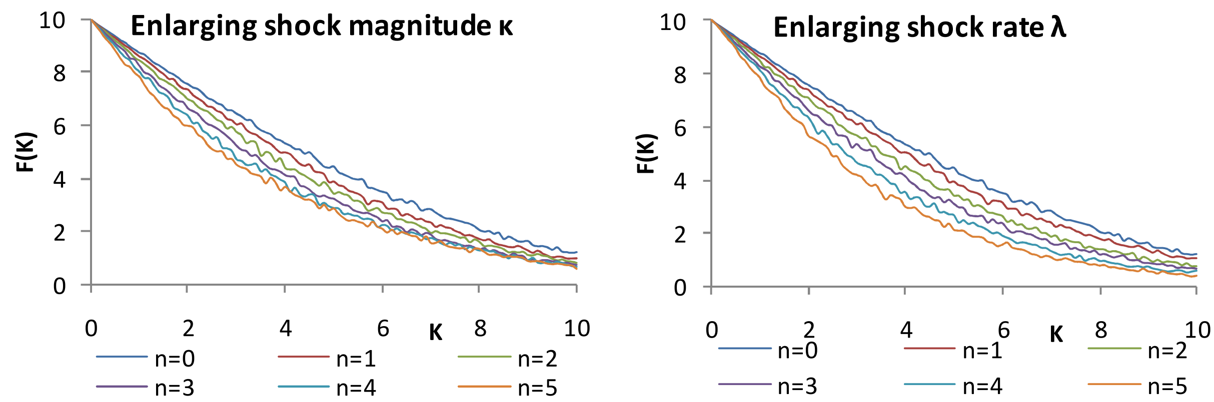


Figure 72; The sensitivity of the option value for changing arrival rate λ and shock magnitude κ. Here, n is $\kappa=0.1nK_0$ respectively the expected number of shocks. Other values are V=10, I=2 and r=0.05

3.3 Conclusion

This chapter showed the impact of shocks on optimal investments under cost uncertainty. The occurrence of shocks lengthens construction time and drives additional larger costs. Shocks introduce additional cost uncertainty because the amount of shocks that occur is unknown beforehand. A first consequence of larger costs and cost uncertainty is a lower option value. But for larger initially expected costs the option including shocks converges to its initial value. Then, projects without shocks have a profitable payoff. Shocks also reduce critical costs for both technical and input uncertainty. Investors demand additional premium to compensate additional risk. To pay this premium costs should decline if the value of the received asset is fixed.

4 Investments in Moin

This last section applies the theory of Chapter 1 to 3 of Part B. As illustration of workings, I study an investment opportunity in Costa Rica, Moin. There, APM Terminals identified the opportunity to construct a new container terminal. Chapter 5 of part A describes the construction works. The subsequent chapters of part A estimate the costs of these construction works. This chapter reviews the investment opportunity and identifies an optimal point to invest. A verification of the optimal point to invest with the initially invested costs verifies the reason to invest.

The chapter starts with a short introduction to the project and makes a number of assumptions in section 4.1. The second section (i.e. section 4.2) calibrates the different parameters. Section 4.3 shows the results and verifies the profitability of the investment opportunity. Section 4.4 puts the results in a broader perspective to further review the profitability of the opportunity. Note that the information on the business case is confidential and may not be distributed in any sense without the permission of APM Terminals.

4.1 Introducing the investment opportunity and model assumptions

APM Terminals identified the opportunity to establish a new container terminal in Moin, Costa Rica. The Costa Rican economy has a strong focus on agriculture and in particular export of soft fruits like bananas and pineapples. The current infrastructure is inferior and limits both export and economical growth. A new container terminal, focusing on the export of these products, drives economic growth. The infrastructure is so inferior that APM Terminals proposed the authorities to create a completely new terminal. APM Terminals has a clear idea of the terminal value but the construction costs are prone to uncertainty.

APM Terminals developed a terminal design and planned to construct the terminal in three different phases. This implies the existence of three different options to invest. APM Terminals has the option to postpone the construction of each part of the terminal. But in reality legal documents oblige APM Terminals to construct the entire terminal. Postponing the latter investments requires APM Terminals to pay a penalty. Therefore this study assumes the first construction phase of the terminal is prone to uncertainty. It is assumed APM Terminals budgeted the two latter phases correctly and these phases are considered fixed. Moreover, information on the additional revenues by the latter investments is not apparent. This makes it impossible to estimate the value of the latter options.

This study also assumes that the stochastic behaviour of construction time does not influence the terminal value. Reality is more obstinate. APM Terminals receives a concession to construct and operate the terminal 33 years after approval from the local authorities. A delay of operations postpone revenues and narrows operation time to regain the invested cost. Here, the terminal value is assumed constant regardless from the end of construction.

4.2 Calibrating the parameters

The review of the investment opportunity requires calibration of the different parameters. First consider the value of the investment opportunity. The sum of the discounted free cash flows determines the asset value. The discount rate is defined easily. The listed mother company, the A.P. Moller Maersk group, set the discount rate for this project at 11.4%. The free cash flows require additional analysis. First, the valuation only considers the free cash flows from operations on (i.e. after construction). The free cash flows equal operational revenues minus tax and investments. In the valuation model, investments include the initial investments and the renewal of operational and civil infrastructure. Initial investments are the payments to obtain the terminal. Mechanical wear requires renewal of pavement and cranes to continue operations. Leaving the initial payment of the first phase out, the terminal has a value of approximately USD 730 million[12].

Second, consider the costs to construct the terminal. The costs of the first phase are approximately USD 410 million. These costs include all civil and operational infrastructures. A verification of optimal point of investing with the estimated budget indicates whether investing is profitable or not. Moreover, the analysis also uses the estimated costs to determine the investment rate. The construction time is estimated at 4 four years. This means that APM Terminals expects to operate the terminal after 1000 days (i.e. working five days a week during 4 years). This implies that the rate of investment I is USD 102.5Mn per year. The time step dt that replicates the progress of the simulated project is 1/250, i.e. one day of construction. Thus the change of costs dK also give the change of costs per day.

The next step is to define uncertainty. The cost analysis in part A of this thesis shows that the standard deviation of the costs is approximately 20% of the mean. This uncertainty relates to technical uncertainty which makes it likely to assume $\beta = 0.277$. The cost analysis of part A also identified shocks. The expected value of the shocks is approximately 15% of the estimated costs. A comparison with the results of Vrijling and Boschloo (2001) shows the value is rather high. But, underground construction introduces additional risk. The considerable amount of dredging and the

[12] If initial payments of all phases are left out, the terminal has a value of USD 800 million. But the construction of all elements at once doubles initial costs. This shows that the postponed construction works significantly increase the terminal value.

related reclamation works enlarge the shock magnitude. This justifies $\kappa = 0.15 * K_0$. Because of little amount of identified shocks I assume $n = 1$[13]. Furthermore, Pindyck showed that $\gamma = 0.2$ for input uncertainty is reasonable and the risk free rate is assumed at $r = 5\%$.

4.3 To invest or not to invest?

Figure 73 shows the value of the option and the optimal investment rule for random costs K.

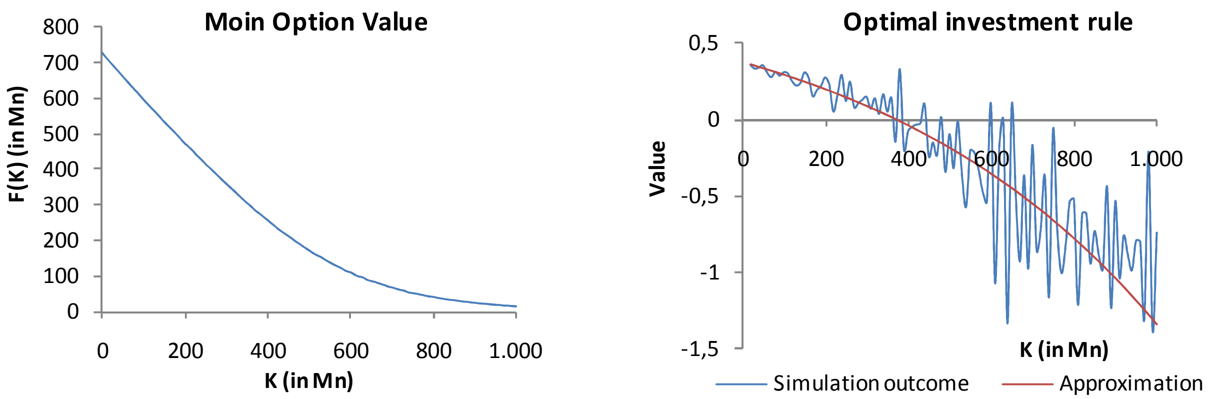


Figure 73 Option value of the investment opportunity in Moin, Costa Rica. Parameter values that value the option are V=730Mn, I=102.5Mn, γ=0.2, β=0.277, κ=0.15*K_0,and r=0.05. The curve $0.9 - 0.525 * e^{-K/690Mn}$ approximates the investment criterion $\frac{1}{2}\beta^2 K F_{kk}(K) - F_k(K) - 1$ and determines critical costs are roughly K*=372Mn.

The left graph of Figure 73 shows the option value. The option of Moin has a similar shape like the example options in the preceding chapters. The discount rate curves the option value. For increasing costs, the option value converges towards 0. Volatility drives the option value for large initial costs because uncertainty advances construction for some projects. This behavior makes some projects profitable that particularly drives the option value for large costs.

The right graph of Figure 73 shows the investment criterion that defines optimal investments;

$$\text{Investment criterion} = \frac{1}{2}\beta^2 K F_{kk}(K) - F_k(K) - 1.$$

The figure shows considerable noise on the condition. The graph is particularly subjected to noise for costs exceeding 600 million. The noise makes it difficult to determine the exact point to stop (which is eventually the objective of this analysis). This research estimates an exponential function that approximates the boundary condition[14] to determine the critical costs K^*. The approximation indicates when the boundary condition is no longer positive that defines optimal investing. The general form of the function is;

---

[13] At t=0 total technical uncertainty is 21.7% of the expected costs. $Var(K) = \alpha^2 E(K)^2$ implies $\left(\frac{\beta^2}{2-\beta^2}\right) K^2 +$ *0,15K2n=α2K+0.15K2* with α=0.217.

[14] The approximation is some exponential function as the solution of the option is likely of an exponential form.

$$\text{Approximation} = A - B * e^{-K*C}.$$

The figure shows an approximation for initial values of $A = 0.9, B = 0.525$ and $C = {}^{1}/_{690Mn}$. The squared error term is used to calibrate the approximation. This solution approximates the curve well, especially for costs smaller than 600Mn. The χ^2-statistic shows the real solution is significant different from the approximation. The approximation shows the critical costs are 372Mn. This indicates that it is not optimal to invest. The initially estimated costs are 410Mn. Based on this analysis APM Terminals should revise its investment opportunity. The initial costs are, given the cost uncertainty and the payoff, too large. This outcome is remarkable because the NPV indicates a highly profitable project. At the other hand, this project is prone to considerable uncertainty. The first chapter shows the difficulty of the DCF-model to incorporate uncertainty. The DCF-model uses the discount rate to capture uncertainty. The option model is closer to reality as it values uncertainty. The example proves that the DCF-model may forecast profits, but a poor incorporation of uncertainty may lead to unprofitable investments.

4.4 Remarks to the outcome

The analysis requires some remarks given the outcome not to invest. One may point out that the approximation of the optimal investment is quite rough. The graph is subjected to enough noise to question correctness of the critical costs. The approximation may also cross the axis more close to 410Mn. This section sets out the various arguments and puts the model assumptions in a broader perspective. The conclusion shows whether the critical costs should be larger or lower.

First, the different parameters are subjected to uncertainty. A sensitivity analysis may provide additional value. But the values of the payoff, the costs and the uncertainty are carefully chosen. Adjustment is more justified if new information becomes available. Then, the information changes the asset value, costs and uncertainty. As a result, the information also changes the point of optimal investment.

Second, the Moin investment is composed of three parts. The analysis assumed the last two parts fixed. Separating the investment likely has a minor influence on the critical costs. The costs and additional payoff from the other two parts have a minor influence on the asset value.

Thus, these first two assumptions only have a minor influence on the critical costs. Consider therefore the more general assumptions of the model. In particular, consider the incorporation of uncertainty within the model. Some uncertainty is neglected within the model. The model assumed a certain payoff but experts within APM Terminals asserted that transport of cargo can be volatile. This implies that the free cash flows that determine the asset value are prone uncertainty. But the uncertainty on the free cash flows is accounted for in valuation. The discount rate is a risk-adjusted discount rate that incorporates uncertainty on the cash flows. This withdraws the arguments because the model accounts for uncertainty on the payoff.

The model does not incorporate the conditions of the concession agreement. The agreement bounds APM Terminals to a fixed time frame to construct and operate the terminal. Delayed construction costs shorten time to operate the model. This lowers both the payoff and the option value. The consequence is further declining critical costs. This last analysis strengthens the decision to postpone the investment and wait for new information.

A further improvement of the model is to incorporate friction costs and a competitive market. Friction costs make halfway stopping costly instead of costless. The start of projects also involve legal obligations. Terminating the project breaks the obligations which the investor (i.e. APM Terminals) has to compensate. The omission of friction costs would reduce critical costs even further because it increases the aversion to invest.

This approach only considered the investor’s interest from a profitable point of view. A competitive market may investors tempt to take other decisions. Investors may decide to invest to retain its market share although investments are not profitable at first sight. Then, the impact of not investing is more unprofitable than investing. This analogy enlarges critical costs but it has no impact on this business opportunity because APM Terminals was the only competitive bidder.

4.5 Conclusion

This last chapter argues to postpone an investment opportunity in Moin. The results of the simulation show that the critical costs are larger than the estimated construction costs. These results are based on the analysis of the first two sections. The first section introduces the business opportunity and simplifies the business case. The second part calibrates the different parameters. These input parameters determine the option value and the point of optimal investment. The last section further discusses the reason to wait and invest later. The analysis strengthens the outcome as the analysis omitted any risk. The model excludes uncertainty on the value as APM Terminals has a concession to construct and operate the terminal for 33 years. Delay of construction works has a considerable impact on the terminal value. Above that, the model excludes friction cost. There is no costless option to abandon. Legal contracts make it more difficult to suddenly terminate the project.

5 Conclusion and suggestions for further research

This chapter discusses the conclusions and recommendations of the second part of this Master's thesis. It is a brief section that summarizes the most important findings. The various chapters elaborate the results. Recommendations for further research are distilled from the conclusions.

5.1 Conclusion

This second part of the Master's thesis extends the optimization of investments under uncertainty with the inclusion of shocks. Each expense towards the completion of a project is a single investment opportunity that investors can postpone. The outcome of the investment is the uncertain progress. Pindyck (1993) splits this uncertainty to technical and input uncertainty. This research adds shocks to the stochastic behavior. Academic research showed most construction projects face unforeseen events during construction (Boschloo and Vrijling (2001)).

Shocks are a special form of technical uncertainty. It only reveals after investments and solely increases the difficulty of physical completion. The inclusion of shocks enlarge invested costs. These enlarged costs reduce (expected) payoff. The amount of shocks that occur during construction is uncertain and introduces additional uncertainty on the investments.

This study shows that the inclusion of shocks has a substantial impact on the investment decisions and increases the aversion to invest. The enlarged aversion lowers critical costs, that define optimal investment. Lower (critical) costs enable a larger profit as the payoff after construction is fixed. This analogy is related to conventional economic theory: Investors demand a higher premium to compensate additional costs and uncertainty.

5.2 Recommendations for future research

This study uses the traditional Real Option to determine optimal investments under uncertainty. The traditional approach neglects many frictions that are apparent in reality. In general, these frictions are important aspects for future research. Here, I identify three interesting topics to better define optimal investments under uncertain costs.

First, this thesis did not study the analytical solution of the model. It requires a profound understanding of mathematics that goes beyond this thesis. Nevertheless the outcome may provide interesting results. Therefore this thesis recommends further research to study the differential equation that describes the exact solution. Fang (2010) developed an efficient and reliable method to determine the option value.

Second, this study neglects the costs to abandon the investment. Agents cannot costless stop the commitment of resources and abandon the investment. Legal documents, signed with the commitment of investing, oblige investors to compensate (sub-)contractors. The learning by investing may justify the termination of projects, but that is certainly not costless. Most investors construct a Special Purpose Vehicle (SPV) that facilitates construction. Terminating the project will therefore imply the liquidation of the SPV. The study of Surandesan and Wang (2007) relates to this topic as they combined investment under uncertainty, financing capacity and costs and (possible) liquidation.

Third, investors attempt to reduce cost uncertainty by undertaking additional engineering studies. The investment problem is more complicated because one has three choices instead of two; start construction now, undertake an engineering study now, and then begin only if the study indicates costs are likely to be low, or abandon the project completely. The application of Bayesian learning during engineering studies may prove the value of these studies. Bayesian learning incorporates the changing uncertainty by the engineering studies. It is even more interesting to account for the incomplete information by the studies and the costs of these studies. Therefore future research is recommended to study the investments under uncertainty using Bayesian learning from engineering studies while the information of the studies is incomplete and costly. Miao and Wang (2007) have already touched upon this topic studying investments of entrepreneurs under incomplete markets.

Literature

Literature related to Part A;

1. **S.A. Ross, R.W. Westerfield, J. Jaffe.** *Corporate finance.* s.l. : McGraw Hill, 2005.
2. **Weert, van.** *Improving CAPEX estimate reporting - Internal memo APMT.* 2010.
3. **Vrijling, J.K.** *PAO Course probabilistic design.* s.l. : TU Delft, 2009.
4. *Is "estimate uncertainty" an oxymoron?* **Dysert, L.R.** s.l. : AACE International Transactions, 2006.
5. *Evaluation of uncertainties in cost estimation.* **Boschloo, Van Gelder, Vrijling.** Delft : IABSE Conference Report on Safety, Risk and Reliability - Trends in Engineering, 2001.
6. **AACE.** *Cost estimation classification system.* s.l. : AACE, 2003.
7. **Knight.** *Risk, uncertainty and profit.* New York : Cosimo, 2005.
8. **Hubbard.** *How to Measure Anything: Finding the Value of Intangibles in Business.* s.l. : John Wiley & Sons, 2007.
9. **Heij, C., et al.** *Econometric methods with apllications in business and economics.* s.l. : Oxford University Press, 2004.
10. **Grinold, Kahn.** *Active portfolio management.* s.l. : McGraw-Hill, 1999.
11. **Horn and Johnson.** *Matrix analysis.* s.l. : Cambridge University Press, 1985.
12. **Gelder, P.H.A.J.M. van.** *Risk based design of civil structures.*
13. **Vrijling.** *Westerschelde survey.*
14. **McNeil, Frey, Embrechts.** *Quantitative Risk Management.* s.l. : Princeton, 2005.
15. *Understanding and managing risk in large engineering projects.* **R. Miller, D. Lessard.** 2001, International journal of project management.
16. *Building real options into physical systems with stochastic mixed-integer programming.* **T. Wang, R. de Neufville.** Montreal : Real option annual conference, 2004.
17. *Judgement under uncertainty heuristics and biases.* **A. Tversky, D. Kahneman.** s.l. : Science, 1974.
18. *Expert elicitation and Bayesian analysis of construction contracts: an investigation.* **Adams, F.K.** s.l. : Construction management and science, 2006.
19. **Langdon.** *Spon's civil engineering and highway works price book.* s.l. : Taylor & Francis, 2007.
20. *Belief networks for construction performance diagnostics.* **B. McCabe, S.M. AbouRizk, R. Goebel.** s.l. : Journal computational civil engineering, 1998.
21. *Investments of uncertain costs.* **Pindyck, Robert S.** Cambridge : Journal of Financial Economics, 1993, Vol. 34.

Literature related to Part B;

22. **Bernardo, A.E., Chowdhry, B.**, 2002. *Resources, real options and corporate strategy*, Journal of Financial Economics 63, 211-234.
23. **Black, F., Scholes, M.**, 1973. *The pricing of options and corporate liabilities*, The Journal of Political Economy 81, 637-654.
24. **Brennan, M.J., Schwartz, E.**, 1985. *Evaluating natural resource investments*, Journal of Business 58, 35-57.
25. **Dixit, A.K., Pindyck, R.S.,** 1994. *Investment under Uncertainty*, Princeton University Press, Princeton.

26. **Fang, F.,** 2010. The COS method: An efficient Fourier method for pricing financial derivatives, Delft University Press, Delft.
27. **Grenadier, S.R**., **Malenko, A.,** 2010. *A Bayesian approach to real options: the case of distinguishing between temporary and permanent shocks*, Journal of Finance 65, 1949-1986.
28. **Grenadier, S.R., Wang, N.**, 2007. *Investment under uncertainty and time-inconsistent preferences*, Journal of Financial Economics 84, 2-39
29. **Lambrecht, B.M., Perraudin, W.,** 2003. *Real options and preemption under incomplete information*, Journal of Economic Dynamics and Control 27, 619-643.
30. **McDonald, R**., **Siegel, D.,** 1986. *The value of waiting to invest,* Quarterly Journal of Economics 101, 707-728.
31. **Merton R.S.**, 1973. *Theory of rational option pricing*, Bell Journal of Economics 4, 141-183.
32. **Miao J., Wang, N.,** 2007. *Investment, consumption and hedging under incomplete markets*, Journal of Financial Economics 86, 608-642.
33. **Pindyck, R.S.,** 1991. Irreversibility, Uncertainty, and Investment, Journal of Economic Literature 24, 1110-1148.
34. **Pindyck, R.S.,** 1993. *Investments of uncertain costs,* Journal of Financial Economics 34, 53-76.
35. **Smit, H.T.J., Trigeoris, L.,** 2006. *Real options and games: Competition, alliances and other applications of valuation and strategy*, Review of Financial Economics 15, 95-112.
36. **Surandesan S., Wang, N.,** 2007. *Investment under uncertainty with strategic debt service*, American Economic Review 97, 256-261.

Annexes

Annex 1: Introduction to statistics

Describing a project by its density function *p(x)* shows an expectation off the outcome. The expected value or mean is calculated with . If one would like to know the sum of two projects, the expected value is the sum of the expected values of the two different projects.

The expected value has a 50% probability of exceedance as well as 50% probability of deceedance. No information on the likelihood of the outcome to the expected value is presented. Three parameters are used to describe the likelihood and distribution of the data around the expected value.

1. Variance or standard deviation

   The variance is used to measure variability from the expected value. A high variance indicates a wide spread distribution and the outcome will deviate largely from the expected value. The standard deviation is also often used as an indicator of the spread and is the squared root of the variance.

2. Skewness

   The skewness measures the symmetry of the probability distribution. In the figure 1 the skewness is clarified. A negative skewness indicates wide distribution in the left tail and likely high values. A positive skewness indicates a large right tail.

3. Kurtosis

   The kurtosis measures the peakedness of the distribution function. A high kurtosis indicates fat tails and likely (extreme) values in the tails.

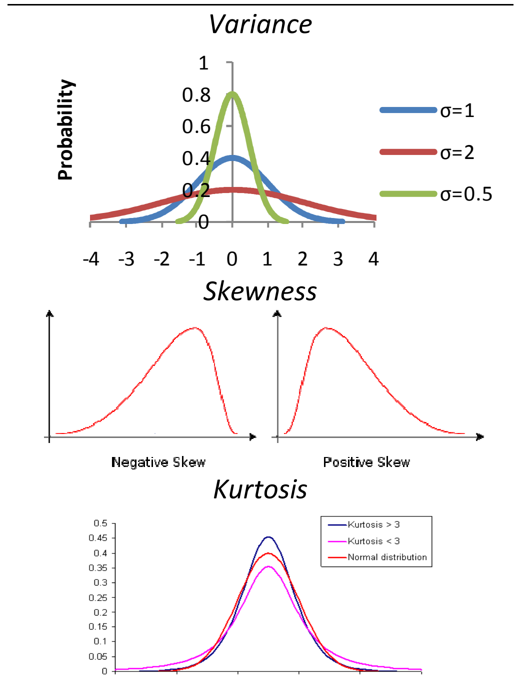


Figure 1: Variance, skewness and kurtosis explained

The skewness and kurtosis are compared to the standard normal distribution. Therefore the demeaned series z is taken to determine the skewness and kurtosis; . The demeaned series has and can therefore be compared with the standard normal distribution. The formulas to calculate the variance, skewness and kurtosis are stated below;

Variance

Skewness

Kurtosis

If one would like to know the combined spread of two density functions it is not possible just to take the sum of the variances. As the two functions are interrelated the variance of the sum becomes:

The part $cov(x, y)$ accounts for the dependency between elements and is called the covariance. The covariance measures the linear interdependency between elements. As the deviations are small, a linear dependency is justified.

The formula to calculate the covariance is:

$$cov = E[(x - E[x])(y - E[y])]$$

Often the correlation is also mentioned to measure the dependency. The correlation is given by

$$corr = \rho = \frac{cov(x, y)}{\sqrt{Var(x)Var(y)}}$$

The variance formula can now be transformed to:

$$var(x, y) = var(x) + var(y) + 2\rho\sqrt{var(x)var(y)}$$
$$var(x, y) = \sigma(x)^2 + \sigma(x)^2 + 2\rho\sigma(x)\sigma(y)$$

Annex 2: Derivation of the CAPM

This section focuses on economic theory on risk management. It is well described in Ross, Westerfield and Jaffe (1). The theory described next is the foundation for portfolio management and investment theory and therefore important theory for people interested in those topics. The basic idea is how to deal with multiple projects and how to position optimally based on the pay off and risk profile. More extensive and in depth theory on portfolio management can be found in Grinold and Kahn(10).

Economists review risk as a deviation from the expected value. Within an ordinary density function like the normal distribution the standard deviation is reviewed as the risk. If the standard deviation is zero and there is only a single expected value there is no risk at all.

If one wants to invest in two different assets, for example two shares, the probability density function of both (based on historical data) can be determined. The returns of the last half year of well traded stocks are often a good approximation as trading conditions do not change in the meantime. These two stocks are normally distributed with mean, variance and correlation as shown in table 1.

If one would buy one share of both stocks, the combined mean and variance are:

Mean $10 + 9 = 19$

Variance $4^2 + 3^2 - 0.2 * 3 * 4 = 22.6$

	Stock 1	Stock 2
Mean	10	9
Variance	16	9
Correlation	-0.1	-0.1

Table 1; mean and variance stocks

One can also buy multiple shares. For each combination of shares 1 and 2 a mean and variance can be calculated. Buying at maximum 100 shares in total, one can determine the mean and variance of the portfolio. The formulas used are shown below, next to the figure showing the portfolio mean and variance:

Mean portfolio $mean = n * \mu_x + (100 - n) * \mu_y$

Var portfolio $var(x, y) = n * \sigma(x)^2 + (100 - n) * \sigma(x)^2 + 2 * \rho * \sqrt{n} * \sigma(x) * \sqrt{100 - n} * \sigma(y)$

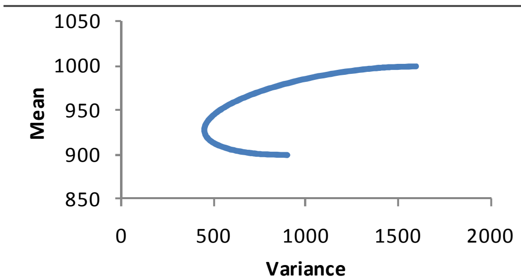


Figure 2: Portfolio's mean and variance

Figure 2 shows a curved line with the investment opportunities. In reality one can invest in more than two shares and a similar line showing the mean and variance of the portfolio can be constructed.

In addition there are also risk free assets having no variance and a fixed mean. Examples are putting your money on a bank account or buying a government bond. In this case the risk free asset has mean 875.

Now one can invest in shares to maximize profit, but as the variance is not zero variability is introduced in the expected profit. This variability is called risk. To limit our risk exposure and ensure a certain payoff, we can invest in the risk free asset. Now a line between the risk free asset and the portfolio can be drawn. This is shown in figure 3. This line shows the optimum positions between the mean and the variance. The maximum return is achieved with the least risk. Based on the risk appetite we can determine our point on the line and thus the risk and payoff strategy.

As the optimal investment line is the linear combination between variance and return, the formula of the line is:

$$E[r] = r_f + \frac{(r_p - r_f)}{\sigma_p^2} r$$

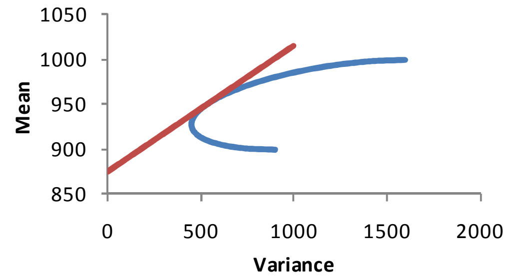


Figure 2: Portfolio's mean and variance

This above approach is the Capital Asset Pricing Model (CAPM). From a scientific perspective it is shown that risk needs to be taken to earn a risk premium. The risk premium generates a high return from which we benefit. This risk premium is twofold as the CAPM proves. The higher the risk premium, the higher the expected return. But at the same time one has to accept a higher spread.

Annex 3: Matlab coding

```
clear all
clc

%load data
Data = xlsread('100723 Monte Carlo input.xlsx', 1, 'D18:K25');
A = size(Data);
m=A(1,1);       %Number of line items
n=1000;     %Number of simulations

%Simulation input
%Simulation according to design
u = rand(m,n);
price_simulation = price_simulation(Data, m, n, u);

u = rand(m,n);
quantity_simulation = quantity_simulation(Data, m, n, u);

%Scenario analysis
scenario_data = xlsread('100723 Monte Carlo input.xlsx', 1, 'C31:G38');
A = size(scenario_data);
o = A(1,1);
u = rand(m,n);
scenario_simulation = scenario_simulation(scenario_data, o, n, u);

%%Total
for x = 1:m
    for y = 1:n
        total_simulation(x,y)                                          =
price_simulation(x,y)*quantity_simulation(x,y);
    end
end
total_simulation(m+1:m+o,1:n)=scenario_simulation(1:o,1:n);

for y = 1:n
    total(y) = sum(total_simulation(:,y));
end
```

Function file 1

```
function price_simulation = price_simulation(Data, m, n, u)

for x = 1:m
for y = 1:n

        if Data (x,8) == 0
            price_simulation(x,y)=Data(x,6);

        elseif Data (x,8) == 1
            mu = Data(x,6);
            sigma = (1/(2*1.96))*(Data(x,7)-Data(x,5));

            price_simulation(x,y)= norminv(u(x,y),mu,sigma);

        elseif  Data(x,8)== 2
            p1 = 0.025;
            p2 = 0.975;
            x1 = Data(x,5);
            x2 = Data(x,7);

            sigma = log(x2/x1)/(sqrt(2)*(erfinv(2*p2-1)-erfinv(2*p1-1)));
            mu = log(x1)-sqrt(2)*sigma*erfinv(2*p1-1);

            price_simulation(x,y)= logninv(u(x,y),mu,sigma);

        elseif  Data(x,8)== 3
            mu = Data(x,6);
            sigma = (1/(2*2.807))*(Data(x,7)-Data(x,5));

            z = tinv(u(x,y),5);

            price_simulation(x,y)= mu+z*sigma;

        elseif  Data(x,8)== 4
            low = Data(x,5);
            middle = Data(x,6);
            high = Data(x,7);

            if u(x,y)<=(middle-low)/(high-low);
               price_simulation(x,y)    =    low    +    sqrt(u(x,y)*(high-
low)*(middle-low));
            else
               price_simulation(x,y)    =   high    -    sqrt((1-u(x,y))*(high-
middle)*(high-low));
            end

        elseif  Data(x,8)== 5
            price_simulation(x,y) = unifinv(u(x,y),Data(x,5),Data(x,7));
        end
end
end
end
```

Function file 2

```
function quantity_simulation = quantity_simulation(Data, m, n, u)

for x = 1:m
for y = 1:n

    if Data (x,4) == 0
        quantity_simulation(x,y)=Data(x,2);

    elseif Data (x,4) == 1
        mu = Data(x,2);
        sigma = (1/(2*1.96))*(Data(x,3)-Data(x,1));

        quantity_simulation(x,y)= norminv(u(x,y),mu,sigma);

    elseif  Data(x,4)== 2
        p1 = 0.025;
        p2 = 0.975;
        x1 = Data(x,1);
        x2 = Data(x,3);

        sigma = log(x2/x1)/(sqrt(2)*(erfinv(2*p2-1)-erfinv(2*p1-1)));
        mu = log(x1)-sqrt(2)*sigma*erfinv(2*p1-1);

        quantity_simulation(x,y)= logninv(u(x,y),mu,sigma);

    elseif  Data(x,4)== 3
        mu = Data(x,2);
        sigma = (1/(2*2.807))*(Data(x,3)-Data(x,1));

        z = tinv(u(x,y),5);

        quantity_simulation(x,y)= mu+z*sigma;

    elseif  Data(x,4)== 4
        low = Data(x,1);
        middle = Data(x,2);
        high = Data(x,3);

        if u(x,y)<=(middle-low)/(high-low);
           quantity_simulation(x,y)  =  low  +  sqrt(u(x,y)*(high-low)*(middle-
low));
        else
           quantity_simulation(x,y)    =    high    -    sqrt((1-u(x,y))*(high-
middle)*(high-low));
        end

    elseif  Data(x,4)== 5
        quantity_simulation(x,y) = unifinv(u(x,y),Data(x,1),Data(x,3));
    end
end
end
end
```

Function file 3

```
function scenario_simulation = scenario_simulation(scenario_data, o, n, u)

for x = 1:o
    for y = 1:n
        if u(x,y) <= scenario_data(x,1)

            if scenario_data (x,5) == 0
                scenario_simulation(x,y)=scenario_data(x,3);


            elseif scenario_data (x,5) == 1
                mu = scenario_data(x,3);
                sigma          =          (1/(2*1.96))*(scenario_data(x,4)-
scenario_data(x,2));

                scenario_simulation(x,y)= norminv(rand,mu,sigma);

            elseif  scenario_data(x,5)== 2
                p1 = 0.025;
                p2 = 0.975;
                x1 = scenario_data(x,2);
                x2 = scenario_data(x,4);

                sigma   =   log(x2/x1)/(sqrt(2)*(erfinv(2*p2-1)-erfinv(2*p1-
1)));
                mu = log(x1)-sqrt(2)*sigma*erfinv(2*p1-1);

                scenario_simulation(x,y)= logninv(rand,mu,sigma);

            elseif  scenario_data(x,5)== 3
                mu = scenario_data(x,3);
                sigma          =          (1/(2*2.807))*(scenario_data(x,4)-
scenario_data(x,2));

                z = tinv(rand,5);

                scenario_simulation(x,y)= mu+z*sigma;

            elseif  scenario_data(x,5)== 4
                low = scenario_data(x,2);
                middle = scenario_data(x,3);
                high = scenario_data(x,4);
                p = rand;
                if p<=(middle-low)/(high-low);
                   scenario_simulation(x,y)    =    low    +    sqrt(p*(high-
low)*(middle-low));
                else
                   scenario_simulation(x,y)   =   high   -   sqrt((1-p)*(high-
middle)*(high-low));
                end

            elseif  scenario_data(x,5)== 5
```

```
                scenario_simulation(x,y)                                    =
unifinv(rand,scenario_data(x,2),scenario_data(x,4));

            end
        else scenario_simulation(x,y) = 0;
        end
    end

end
end
```

Annex 4: Analytical solution to standard normal methodology

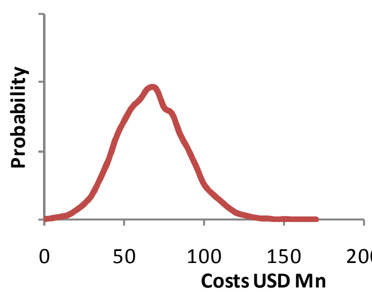
Proof mean:

$$C_i = P_i * Q_i$$
$$P_i = \mu_p + \varepsilon_p * \sigma_p$$
$$Q_i = \mu_Q + \varepsilon_Q * \sigma_Q$$
$$\varepsilon_p, \varepsilon_p \sim N(0,1)$$
$$E[C_i] = E[P_i * Q_i]$$
$$E[C_i] = E\left[\left(\mu_p + \varepsilon_p * \sigma_p\right) * \left(\mu_Q + \varepsilon_Q * \sigma_Q\right)\right]$$
$$E[C_i] = E\left[\mu_p * \mu_Q + \mu_p * \varepsilon_Q * \sigma_Q + \mu_Q * \varepsilon_p * \sigma_p + \varepsilon_Q * \sigma_Q * \varepsilon_p * \sigma_p\right]$$
$$E[C_i] = \mu_p * \mu_Q$$

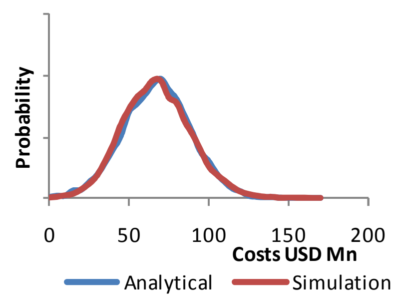
Proof variance:

$$Var(C_i) = E[(C_i - E[C_i])^2]$$
$$Var(C_i) = E[(P_i * Q_i - E[P_i * Q_i])^2]$$
$$Var(C_i) = E\left[\left(\left(\mu_p + \varepsilon_p * \sigma_p\right) * \left(\mu_Q + \varepsilon_Q * \sigma_Q\right) - E\left[\mu_p * \mu_Q\right]\right)^2\right]$$
$$Var(C_i) = E\left[\left(\left(\mu_p + \varepsilon_p * \sigma_p\right) * \left(\mu_Q + \varepsilon_Q * \sigma_Q\right) - \mu_p * \mu_Q\right)^2\right]$$
$$Var(C_i) = E\left[\left(\mu_p * \varepsilon_Q * \sigma_Q + \mu_Q * \varepsilon_p * \sigma_p + \varepsilon_Q * \sigma_Q * \varepsilon_p * \sigma_p\right)^2\right]$$
$$Var(C_i) = \mu_P^2 * \sigma_Q^2 + \mu_Q^2 * \sigma_P^2 + \sigma_Q^2 * \sigma_P^2$$

Simulated output



Comparison



		Simulation		*Analytical app.*	
Mean 1	7	*Mean*	69.6mn	*Mean*	70.0mn
Standard deviation 1	2	*Stdev*	21.3mn	*Stdev*	21.3mn
Mean 2	10,000,000	*Skew*	0.22	*Skew*	0
Standard deviation 2	1,000,000	*Kurt*	3.11	*Kurt*	3

Figure 74; Monte Carlo vs. normal distributed cost estimates

Annex 5; Methods to calculate the reliability of systems

Vrijling(3) focuses on reliability of systems. A system is defined as a coherent set of matching objects and their parts; e.g. the cash flows of a new terminal. Reliability is the rate of (no) system failure. Risk management focuses on the probability of failure and how to avoid failure. Examples of such analyses are a cost or strength analysis. The system can be described with:

$$System's\ strength = Resistence - Solicitation$$

The financial performance of an APM Terminals project is described by the system:

$$Net\ Profit = Revenues - Costs$$

Within APMT the costs are priced as a fixed value as well as the revenues. Factors are added to cover a possible cost increase or revenue decrease. If the costs are not larger than the revenues the system is reliable. As the revenues and costs change over time, the cash flows are discounted. If the sum of the discounted cash flows is positive, the system is expected to be reliable (i.e. profitable). Nevertheless there is no link between the reliability of the system and the probability of failure. I will try to establish a link.

For both revenues and costs I can derive a probability density function. Then I can calculate the probability that the costs are larger than the revenues. A direct link between probability of failure and system reliability is established.

Within this thesis I focus on costs estimation and the aim is to design a cost level C_0. From the analysis on the likelihood of the costs the probability of overrun is determined. The system called Z is therefore defined by:

$$Z = C_0 - Costs$$

Based on the elements I will be able to construct a probability density function of the costs. Thereafter a level C_0 can be defined and the reliability of the system can be determined. As part of another study it might be worthwhile to do research on the probability density function of the revenues.

Linear systems

Vrijling points out how to model the elements and starts off with linear systems. Linear systems are systems described by a linear equation; i.e. $y = a * x + b$. In general the costs of an element are non linear. The cost of an element is the product of the quantity and the price per quantity. As both the price and quantity of the elements vary, the system is non-linear. The non-linear system consisting of n elements is priced with:

$$Costs = \sum_{i=1}^{n} P_i * Q_i$$

The aim is to design a cost level based on the reliability of the system. Therefore the system now becomes:

$$Z = C_0 - \sum_{i=1}^{n} P_i * Q_i$$

First the system is simplified by reviewing the total costs of an element. This makes it possible to linearize the system to:

$$Z = C_0 - \sum_{i=1}^{n} C_i$$

As an example I have defined the following system of 7 elements. The total prices as well as the quantities are given:

Element	Price ($)	Quantity (unit)	Unit	Total
Engineering	2,000,000	1	$project^{-1}$	2,000,000
Break water	50,000	1,000	m^{-1}	50,000,000
Quay wall	40,000	1,000	m^{-3}	40,000,000
Dredging	4	3,000,000	m^{-3}	12,000,000
Container Yard	75	100,000	m^{-2}	7,500,000
Administration building	1,000	2,000	m^{-3}	2,000,000
Total				113,500,000

The linear system definition now becomes:

$$Z = C_0 - (2 + 50 + 40 + 12 + 7.5 + 2)$$
$$Z = C_0 - 113.5$$

If I define the cost budget (C_0) at $115,000,000 the system is ok; Z will be larger than 0. If the budget would be set at less than 113,500,000 the system will fail. Often a safety factor is added to the different components to allow for variability. By applying a safety factor of γ=1.2 the point estimate of C_0= $115,000,000 is insufficient. The costs are likely to overrun the budget as Z will become below 0. (Z=115-113.5*1.2<0)

There is often a link between the (expected) variability of the element and the safety factor. If the element is likely to vary, the safety factor will increase.

The above method is called a deterministic approach within Vrijling(3). Within this approach the costs are assumed fixed. A safety factor is added to every component. A (sufficient) gap should be present between the cost budget and the possible cost development to ensure reliability (i.e. profitability). However there is no direct link between the reliability and the probability of failure of the system.

To establish a link between the reliability and probability of failure I assume elements to vary; for instance the container yard of the previous example. The total costs for the yard are normally distributed with mean μ=7,500,000 and standard deviation σ=800,000. I estimate the budget at $115,000,000. Thus the system now becomes:

$$Z = C_0 - C - C_{container\ yard}$$
$$Z = 115 - 106 - N(7.5,0.8)$$

The probability of failure of the system Z is given in Figure 75. The probability of failure is given by:

$$Z = 41.5 - \mu_{yard} - \beta * \sigma_{yard}$$
$$\mathrm{P}(\mathrm{Z} < 0) = \Phi(\beta) = \Phi(-1.875) = 0.069$$

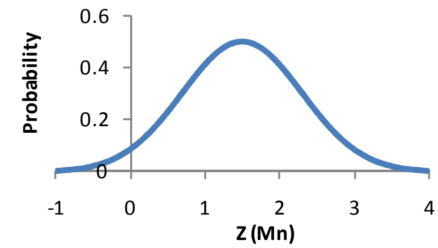


Figure 75; Failure area of a system

The calculation above shows that the probability of cost overrun is 6.9%. If APM Terminals wants to have a maximum probability of cost overrun of 50%, the budget should be set at:

$$\beta = \Phi^{-1}(0.50) = 0$$
$$C_0 = Z + C + \mu_{yard} + \beta * \sigma_{yard}$$
$$C_0 = 0 + 106 + 7.5 = 113.5$$

If two or more variables are varying, the formula remains equal under the condition that variables are normally distributed. In this example it is assumed that the administration building is normally distributed N(2,0.5) and slightly correlated (ρ=0.2) with the container yard costs. The formula and graph now become:

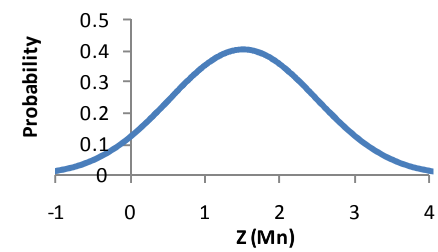


Figure 76; Failure area of a system

$$\mu_z = \mu_{yard} + \mu_{admin} = 9.5$$
$$\sigma_z = \sqrt{\sigma_{yard}^2 + \sigma_{admin}^2 + 0.2 * \sigma_{yard} * \sigma_{admin}} = 0.98$$
$$Z = 11 - \mu_z - \beta * \sigma_z$$
$$\beta = \frac{11 - \mu_z}{\sigma_z} = 1.52$$
$$\mathrm{P}(\mathrm{Z} < 0) = \Phi(\beta) = \Phi(1.52) = 0.125$$

If the maximum probability of cost overrun is 50%, the budget should be:

$$C_0 = 0 + 104 + 9.5 = 113.5$$

The probability density function in Figure 76 reflects the increase in variability. As more parameters vary in the density function of Figure 76, the width of the density function is larger than the width of the density function in Figure 75. Positive correlation between parameters will result in an even wider density function as variance increases. In contrast, negative correlation will result in decreasing variance.

Additional variables

Having more variables (all assumed to be normally distributed) the system becomes:

$$Z = C_0 - C - N(\mu, \Sigma)$$

Within the above system $C_0$is the budget, C is the sum of the fixed costs and $N(\mu, \Sigma)$ is the normal distribution of the varying elements. μ is the vector containing the expected value and the sum of the elements' mean is the total expected value.

The variability of the system is given by the matrix Σ. On the diagonal the variances are given and at the upper and lower triangle the covariances. Multiplying the matrix Σ with a vector of ones gives the variance of the system Z.

Another approach to compute σ_z, is to construct the correlation matrix P. The standard deviation is derived by multiplying a vector containing the standard deviations with the correlation matrix. Below the matrices and the operations are shown.

$$\sigma_z^2 = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \end{bmatrix} \begin{bmatrix} var(1) & cov(1,2) & \cdots & cov(1,n-1) & cov(1,n) \\ cov(1,2) & var(2) & & cov(2,n-1) & cov(2,n) \\ \vdots & & \ddots & \vdots & \\ cov(1,n-1) & cov(2,n-1) & & var(n-1) & cov(n-1,n) \\ cov(1,n) & cov(2,n) & \cdots & cov(n-1,n) & var(n) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix}$$

$$\sigma_z^2 = \begin{bmatrix} \sigma_1 & \sigma_2 & \cdots & \sigma_{n-1} & \sigma_n \end{bmatrix} \begin{bmatrix} 1 & \rho_{1,2} & \cdots & \rho_{1,n-1} & \rho_{1,n} \\ \rho_{1,2} & 1 & & \rho_{2,n-1} & \rho_{2,n} \\ \vdots & & \ddots & \vdots & \\ \rho_{1,n-1} & \rho_{2,n-1} & & 1 & \rho_{n-1,n} \\ \rho_{1,n} & \rho_{2,n} & \cdots & \rho_{n-1,n} & 1 \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_{n-1} \\ \sigma_n \end{bmatrix}$$

This system accounts for the linear relationship between the various elements within the (co)variance or correlation matrix.

The first computational method is used among financial specialists. The matrix shows the (co)variances of different assets (e.g. shares). The vector containing ones represents the position of the analyst. As multiple assets can be bought, the vector does not contain ones but numbers between plus and minus infinity.

I will use the second computation method as there is only one project to invest in. More important this method makes it easier for experts to estimate the variability. The variability and the interdependency of the elements have to be estimated separately. The covariance troubles the perception of the variability and interrelation with other elements. A dependency between -1 and 1 (or -100% and 100%) is easy to determine as well as the standard deviation of the different elements.

The above system is based on linear dependency. The change in value of the asset is described by $\partial y = \alpha + \beta * \partial x$. Actually β is equal to ρ, as it measures the linear relationship. However, various relations between elements are non-linear. In case of small deviations the linear relationship is a good approximation.

Driving variables

If the estimated costs are too high several measures can be taken. At first the company may decide to take more risk by lowering the budget. This is the simplest solution, but does not seem appropriate.

I can also compute the contribution of the individual variables to the system mean or standard deviation. Based on insight in the contribution of the different variables, the budget could be lowered as well. Measures can be thought of to reduce the mean or standard deviation. Based on the measures the budget can be lowered without changing the risk on the project. After the contribution of the parameters is identified, I recommend to start looking at the largest contributors. A change of these contributors will also have the largest impact.

The mean of the system Z is based on the sum of the different elements. The largest mean will therefore have the largest contribution to the mean of the system.
The contribution of the elements to the variance can be computed with:

$$\alpha_{dredg} = \frac{\sigma_{dredg}}{\sigma_z}, \alpha_{admin} = \frac{\sigma_{admin}}{\sigma_z}$$

The above only accounts for variables those are uncorrelated. When the elements are correlated a Cholesky decomposition should be performed. This is a special matrix operation and I will not go into detail about this operation. More information on the Cholesky decomposition can be found in Horn and Johnson (11).

Non linear systems

In the beginning the nonlinear system was shown:

$$Z = C_0 - \sum_{i=1}^{n} P_i * Q_i$$

To start off, the system was transformed to a linear system. Nevertheless, the aim is to present a realistic model. Therefore I proceed with the non-linear model.

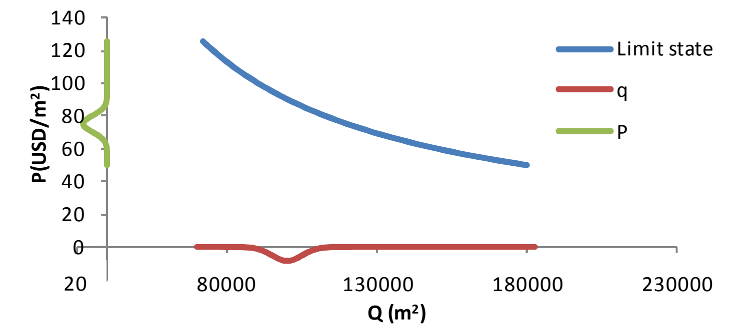


Figure 77; Limit state of a non-linear system

At first I assume that only the price and quantity of the container yard are normally distributed. Thus the system now becomes:

$$Z = C_0 - C - P_{yard} * Q_{yard}$$
$$Z = 115Mn - 106Mn - N(75{,}5) * N(100000{,}5000)$$

As the probability density function of the strength of the system is a product of two normal functions, the density function itself is not normal. The product of normal independently and identical distributed density functions will converge to the log normal distribution function based on the central limit theorem shown by Van Gelder (12).

In Figure 77 the limit state (Z=0) of the system is given for the quantity and price. On the negative side of the x- and y-axis are the probability density functions given of P and Q.
The graph clearly shows the limit state of the system. Above the line, the costs will be larger than the estimate. If the combination of P and Q result in a point below the line, the project has remained within the estimate. The probability density functions at the x- and y-axis clearly show the likelihood of P and Q and as a result the likelihood of the cost development.

Linear approximation

The system can be transformed to a linear system using the Taylor rule. Using a linearized system, the probability of failure (i.e. cost overrun) can be approximated easily. The first order Taylor polynomial to linearize the system is given by:

$$Taylor\ rule: Z \approx g(X_i) + \sum_{i=1}^{n} \frac{\partial g(X_i)}{\partial X_i}(X_i - \mu)$$

As the system is given by $Z = C_0 - C - P_{yard} * Q_{yard}$ the first order derivatives are $\frac{\partial Z}{\partial P} = Q, \frac{\partial Z}{\partial Q} = -P$. The system now becomes:

$$\mu_{z,i} = C_0 - C - P_i * Q_i + \frac{\partial Z}{\partial P}(P_i - \mu_p) + \frac{\partial Z}{\partial Q}(Q_i - \mu_q)$$

$$\sigma_{z,i} = \sqrt{\left(\frac{\partial Z}{\partial P_i} * \sigma_P\right)^2 + \left(\frac{\partial Z}{\partial Q_i} * \sigma_Q\right)^2}$$

$$\beta_{i,i+1} = \frac{\mu_{z,i}}{\sigma_{z,i}}$$

P_i and Q_i are defined by

$$P_{i+1} = \mu_P + \beta_{i+1}\alpha_{P,i}\sigma_P$$

$$Q_{i+1} = \mu_Q + \beta_{i+1}\alpha_{Q,i}\sigma_Q$$

and α_P and α_Q by

$$\alpha_{P,i} = \frac{\left(-\frac{\partial Z}{\partial P_i} * \sigma_P\right)}{\sigma_{z,i}}, \alpha_{Q,i} = \frac{\left(-\frac{\partial Z}{\partial Q_i} * \sigma_Q\right)}{\sigma_{z,i}}$$

By iterating P and Q the limit state is determined. The iteration step is noted by i. The expected value of P and Q are taken as start value. From this assumption μ_z , σ_z and β can be computed. Based on the renewed β and the contribution of P and Q (given by the alphas) the new P, Q, μ_z and σ_z can be determined.

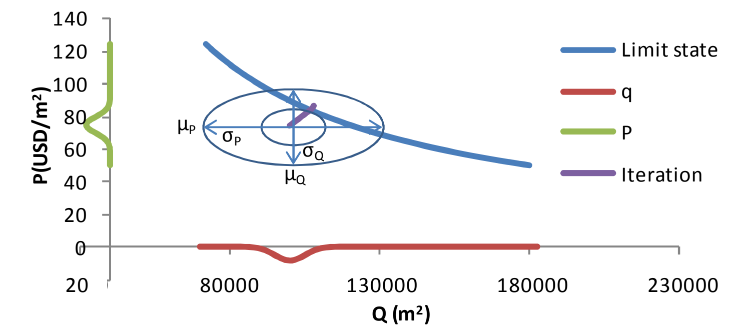
This iteration process ends with the linearized and normally distributed system:

$$Z = C_0 - C - (\mu_{z,i} + \beta_i * \sigma_{z,i})$$

The probability of failure is $\mathrm{P}(\mathrm{Z} < 0) = \Phi(\beta) = \Phi(2.86) = 0.002$. In Figure 77 the iteration steps with the μ_Z, σ_Z and β are given. If APM Terminals would allow a 50% probability cost overrun the budget should be set at $C_0 = 113{,}500{,}000$.

In Figure 78 the non linear system is also shown again. From the expected values of P and Q eclipses of equal probability can be drawn. The above computation calculates the smallest distance from the limit state P(Z<0) to the expected value. Graphically this is the eclipse with the smallest radius; numerically this is the smallest beta from the expected value of P and Q to the limit state. The iteration procedure approximates the radius of the smallest eclipse. This line is drawn as well.

The approach above is called a level II method by Vrijling.



Iteration	$\mu_{Z,i}$	$\sigma_{Z,i}$	β_i
1	1,500,000	625,000	2.40
2	1,569,120	567,649	2.76
3	1,589,950	559,665	2.84
4	1,594,418	558,077	2.86
5	1,595,332	557,757	2.86
6	1,595,514	557,694	2.86
7	1,595,549	557,681	2.86
8	1,595,555	557,679	2.86

Figure 78; linear approximation of the limit state

Additional variables

If I would assume more variables vary, the system can still be linearized according to the above Taylor rule. This would require additional computations. The system becomes with one other varying element:

$$Z = C_0 - C - P_{yard} * Q_{yard} - P_{admin} * Q_{admin}$$

$$\mu_{z,i} = C_0 - C - P_{1,i} * Q_{1,i} - P_{2,i} * Q_{2,i} + \frac{\partial Z}{\partial P_1}(P_{1,i} - \mu_{p,1}) + \frac{\partial Z}{\partial Q_1}(Q_{1,i} - \mu_{q,1}) + \frac{\partial Z}{\partial P_2}(P_{2,i} - \mu_{p,2}) + \frac{\partial Z}{\partial Q_2}(Q_{2,i} - \mu_{q,2})$$

$$\sigma_{z,i} = \sqrt{\left(\frac{\partial Z}{\partial P_{1,i}} * \sigma_{P,1}\right)^2 + \left(\frac{\partial Z}{\partial Q_{1,i}} * \sigma_{1,Q}\right)^2 + \left(\frac{\partial Z}{\partial P_{2,i}} * \sigma_{2,P}\right)^2 + \left(\frac{\partial Z}{\partial Q_{2,i}} * \sigma_{2,Q}\right)^2}$$

The above shows that it is possible to add additional varying elements. But, this is at the expense of increased complexity of the analytical solution.

Numerical approximation

The linearized Taylor approach originates from the days when little computational power was available. At present the computational power is sufficient to simulate the probability density function of a project. I use a Monte Carlo simulation to derive an approximation of the probability density function of the total dredging costs.

Characteristics density function	
Mean	$1.5*10^6$
Stdev	$0.6*10^6$
Skew	-0.12
Kurt	0.09
P(Z<0)	0.01

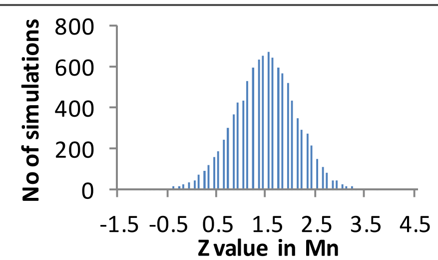


Figure 79; Numerical approximation of Z-function

Almost all computer programs have the ability to simulate a number on the interval [0,1]. By taking the inverse of the cumulative distribution you will end with the x value of the probability density function.

If I simulate the probability density functions P and Q similar to the previous example the output is shown in Figure 79. The price and quantity are simulated independently.

The probability of failure of the system is the amount of Z<0, divided over the total amount of simulations.

After 10,000 simulations in Microsoft Excel the probability of failure was P(Z<0)=0.01. There is a difference in probability of failure within the numerical and the linearized method. The distribution in the tails of distribution functions is (highly) non-linear. As a result the linearized approximation becomes less accurate.

If the different probability density functions are assigned to the elements and the output is derived numerically the method is called a level III within Vrijling. Others call the approach a full probabilistic Monte Carlo method.

Interdependency between the elements within non-linear systems

The linearized and Monte Carlo models do not take interdependency between the various elements into account. Interrelating costs of various line items within the system is very difficult based on the covariance matrix and prone to inconsistencies. Interrelating the prices or quantities with a covariance matrix and deriving one large density function for the price or quantity is not possible as the units among elements differ.

It could be possible to derive the probability density functions of the different elements. This computational method is difficult and prone to inconsistencies. The density functions of the elements can be computed. Using statistical tools like Copula models the total distribution function can be derived. This topic is out of scope of the thesis and a topic for additional research.

Exceptional events

The above approach is a good method to model the variability of elements. However things can happen causing a cost increase of certain elements. For example, during the dredging works a bomb might be found. The probability density of the system now becomes:

$$P(Z) = P(Z|no\ bomb) * \big(1 - P(bomb)\big) + P(Z|bomb) * P(bomb)$$

I assume that if a bomb is found the dredging costs per cubic meter increase with 10%; i.e. $\mu_{p,yard|bomb}$=82.5, $\sigma_{p,yard|bomb}$=5. Two different density functions $P(Z|no\ bomb)$ and $P(Z|bomb)$can be derived from this. These two density functions are presented in Figure 80.

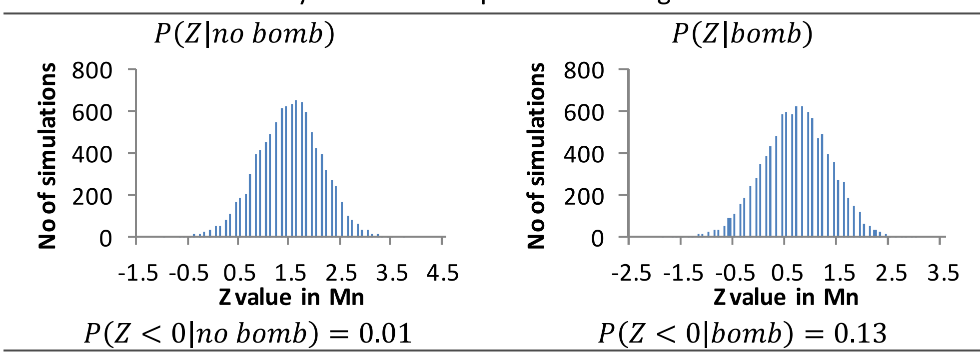


Figure 80; Failure area of a system

The output clearly shows in Figure 80 a shift in probability of system failure if a bomb is found. If I assume the probability of the discovery of a bomb is $P(bomb) = 0.01$, the total probability density function of the system now can be derived. The probability density function of the system is shown in Figure 81.

Characteristics density function	
Mean	$1.5*10^6$
Stdev	$0.6*10^6$
Skew	-0.13
Kurt	-0.02
P(Z<0)	0.01

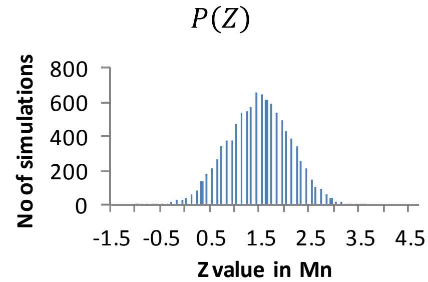


Figure 81; Numerical approximation of Z-function

As the probability finding a bomb is considerably small, the density function is similar the density function $P(Z < 0|no\ bomb)$.

Annex 6; Derivation of the option value

The portfolio consists of the investment opportunity and the asset; $\Phi = F(K) - nx$. Its instantaneous change is $d\Phi = dF - nx$. The expected growth rate of *x* is α_x and the payment of the short position is $n(r_x - \alpha_x) = n\delta x$. Holding the investment opportunity requires a payment stream *I* that makes the return on the portfolio is $dF - nx - n\delta x - Idt$.

Using Ito's lemma write dF as;

$$dF = F_k dK + \frac{1}{2} F_{kk} (dK)^2$$

Substituting the differential equation $dK = -Idt + \beta\sqrt{IK}dz + \gamma K dw$ makes

$$dF = -IF_k dt + \beta\sqrt{IK}F_k dz + \gamma K F_k dw + \frac{1}{2}\beta^2 IKF_{kk} dt + \frac{1}{2}\gamma^2 K^2 F_{kk} dt$$

Because $dx = \alpha_x x dt - \sigma_x x dw$ The return on the portfolio is now

$$-IF_k dt + \beta\sqrt{IK}F_k dz + \gamma K F_k dw + \frac{1}{2}\beta^2 IKF_{kk} dt + \frac{1}{2}\gamma^2 K^2 F_{kk} dt - n\alpha_x x dt - n\sigma_x x dw - n\delta x dt - Idt$$

Setting $n = \frac{\gamma K F_k}{\sigma_x x}$ all terms in *dw* can be eliminated. Hereby non-diversifiable risk is eliminated from the portfolio and the only risk on the portfolio is diversifiable. The expected rate of return on the portfolio must therefore be the risk free rate. Equating the return to $r(F(K) - nx)dt$ leaves the differential equation

$$\frac{1}{2}\beta^2 IKF_{kk} + \frac{1}{2}\gamma^2 K^2 F_{kk} - IF_k - \varphi\gamma K F_k - I = rF$$

Annex 7; Scripts to determine the option value

Numerical solution;

A numerical solution is the discrete approximation of a continuous formula. On some numerical grid one approximates the outcomes. Figure 82 shows the numerical approximation of some continuous function. The numerical solution has some truncation error and not all numerical scripts are stable. For more information on stability and convergence consider any textbook on numerical mathematics.

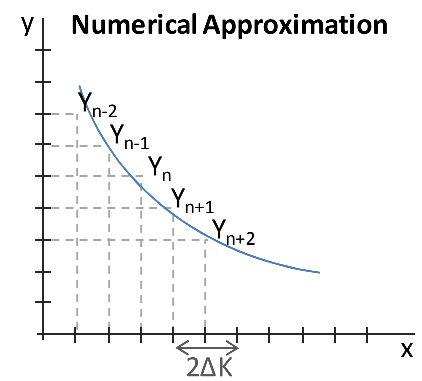


Figure 82; Concept of numerical solution

To obtain the numerical solution of the option, consider the differential equation that describes the option value;

$$\frac{1}{2}\beta^2 IKF_{kk} + \frac{1}{2}\gamma^2 K^2 F_{kk} - IF_k - \varphi\gamma K F_k - I = rF.$$

Discretize the differential equation using some numerical script;

$$F_k = \frac{F_{N+1} - F_{N-1}}{2\Delta K} \; and \; F_{kk} = \frac{F_{N+1} + 2F_N - F_{N-1}}{2\Delta K^2}.$$

Now the differential equation is;

$$\frac{1}{2}\beta^2 IK_N \frac{F_{N+1} + 2F_N - F_{N-1}}{2\Delta K^2} + \frac{1}{2}\gamma^2 K_N^2 \frac{F_{N+1} + 2F_N - F_{N-1}}{2\Delta K^2}$$
$$-I\frac{F_{N+1} - F_{N-1}}{2\Delta K} - \varphi\gamma K_N \frac{F_{N+1} - F_{N-1}}{2\Delta K} - I = rF_N.$$

Transform this differential equation to a matrix and structure it all;

$$\begin{bmatrix} -\frac{\beta^2 IK_N + \gamma^2 K_N^2}{4\Delta K^2} + \frac{I + \varphi\gamma K_N}{2\Delta K} & \frac{\beta^2 IK_N + \gamma^2 K_N^2}{2\Delta K^2} - r & \frac{\beta^2 IK_N + \gamma^2 K_N^2}{4\Delta K^2} - \frac{I + \varphi\gamma K_N}{2\Delta K} \end{bmatrix} \begin{bmatrix} F_{N-1} \\ F_N \\ F_{N+1} \end{bmatrix}$$

Because the exact solution is F(K), this matrix multiplication should equal K;

$$\begin{bmatrix} -\frac{\beta^2 IK_N + \gamma^2 K_N^2}{4\Delta K^2} + \frac{I + \varphi\gamma K_N}{2\Delta K} & \frac{\beta^2 IK_N + \gamma^2 K_N^2}{2\Delta K^2} - r & \frac{\beta^2 IK_N + \gamma^2 K_N^2}{4\Delta K^2} - \frac{I + \varphi\gamma K_N}{2\Delta K} \end{bmatrix} \begin{bmatrix} F_{N-1} \\ F_N \\ F_{N+1} \end{bmatrix} = K_N$$

Given the boundary condition F(0)=V and $\lim_{K\to\infty} F(K) = 0$, define a vector K to obtain a solution for all other values of F.

$$\begin{bmatrix} 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & -\frac{\beta^2 IK_N + \gamma^2 K_N^2}{4\Delta K^2} + \frac{I + \varphi\gamma K_N}{2\Delta K} & \frac{\beta^2 IK_N + \gamma^2 K_N^2}{2\Delta K^2} - r & \frac{\beta^2 IK_N + \gamma^2 K_N^2}{4\Delta K^2} - \frac{I + \varphi\gamma K_N}{2\Delta K} & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} F_0 = V \\ \vdots \\ F_{N-1} \\ F_N \\ F_{N+1} \\ \vdots \\ F_\infty = 0 \end{bmatrix}$$
$$= \begin{bmatrix} 0 \\ \vdots \\ K_N \\ \vdots \\ \infty \end{bmatrix}.$$

Simulation

To obtain the simulation results I used a Monte Carlo simulation. I simulated the costs to completion with:

$$dK = -Idt + \beta\sqrt{IK}\varepsilon_\beta\sqrt{dt} + \gamma K\varepsilon_\gamma\sqrt{dt} \text{ and,}$$

$$K = (K + dK)e^{-rdt}, I = Ie^{-rdt} \text{and } \varepsilon_\beta, \varepsilon_\gamma \sim N(0,1)$$

Counting the amount of (time) steps i to get the costs from $K = K$ to $K = 0$ determines the value of the option using;

$$F(K) = maxE_0\left[Ve^{-\mu\check{T}} - \int_0^{\check{T}} I(t)e^{-\mu t}dt, 0\right] \text{and } \check{T} = i.$$

The average of many (e.g. 10,000) simulations for each initial value K converges to the exact solution. It takes quite some time to determine the solution in Matlab. My Matlab version only used one processor. A conversion of the programming code from Matlab to C made it possible to use multiple processors at once.