Cost Overruns in Transport Projects
- Experiences from Sweden

ANCHALEE JENPANITSUB

Master of Science Thesis
Stockholm, Sweden 2011
Cost Overruns in Transport Projects  
- Experiences from Sweden

Master of Science Thesis

June 2011

Anchalee Jenpanitsub

Division of Transport and Location Analysis  
Department of Transport Science  
KTH Royal Institute of Technology, Stockholm
Abstract

Cost overrun of transport projects is one of the most important problems in transport planning. Apart from causing budget overruns, it also results in uncertain cost-benefit for decision making. This thesis studies cost overruns in Sweden and internationally, factors affecting cost overruns and possible improvements of cost calculations.

The literature study confirms that cost overrun problem is a global phenomenon. The average cost overruns in rail projects are always higher than in road projects. We have compared cost estimations and outcomes of 167 road and rail projects in Sweden during the period 1997-2009. This reveals that average cost overruns are 11% (SD = 24.6%) and 21% (SD = 50.5%) for road and rail projects, respectively. In Sweden, the average cost overrun in road projects is similar to other countries, while the average cost overrun in rail projects is lower than in other countries. However, the standard deviation of cost overruns in Swedish rail projects is very high. The cost overruns in road and rail projects in Sweden have been constant for the 13-year period and cost estimates have not improved over time. Furthermore, small Swedish transport projects (< 100 million SEK) have much higher percentage of cost overruns than large projects. To improve cost estimates in Sweden, the Successive Calculation method has recently been applied. We have collected data for 295 planned projects and find that the variance is significantly lower in these than in actual outcomes, and that the difference is surprisingly small between projects in different planning stages. Another method, Reference Class Forecasting, is demonstrated in two case studies - Stockholm bypass and Västlänken. The two methods are also compared in the thesis. For both case studies, the project costs by using the Reference Class Forecasting method are higher than the project costs by using the Successive Calculation based on the equal costs at 50% confident level of cost overrun.
This page is intentionally left blank.
Acknowledgements

It is pleasure to thank the many people who made this thesis possible. First and foremost, I am heartily thankful to my supervisor, Mattias Lundberg, whose generosity, encouragement, support and guidance from the initial to the final level enabled me to develop an understanding of the subject. Without his assistance and patience, this thesis would never be as good as today.

I would like to thank Roger Pyddoke and Per Näsman for their valuable comments and suggestions during thesis work.

I also would like to thank Pakpairin Uthairat, my good friend, for helping me out for with all kinds of stuffs, including trying her best to do the proof reading of this thesis.

My special thanks go to my office – the Mass Rapid Transit Authority of Thailand, for encouragement, and for giving me the opportunity of the study leave.

Lastly, and most importantly, I would like to thank my parents and my sister for their absolute confidence in me, raising me in the loving environment, educating me well in arts and science, and their unconditional endless love, support, understanding and encouragement throughout my entire life. Without them, it would have been impossible for me to achieve all the success that I have achieved throughout my life.

Anchalee Jenpanitsub

Stockholm, June 2011
## Contents

Abstract .................................................................................................................................................. 1

Acknowledgements .................................................................................................................................. 3

List of Figures ......................................................................................................................................... 7

List of Tables .......................................................................................................................................... 9

List of Acronyms .................................................................................................................................... 10

1 Introduction ......................................................................................................................................... 11

1.1 Purpose and scope of this thesis ........................................................................................................ 11

1.2 Overview of state-of-the-art .............................................................................................................. 12

1.3 Methodology ..................................................................................................................................... 16

1.3.1 Research design ............................................................................................................................ 16

1.3.2 Data .............................................................................................................................................. 17

1.3.3 Disposition ................................................................................................................................... 17

2 Cost overrun problem ........................................................................................................................ 19

2.1 Previous studies ............................................................................................................................... 22

2.2 Cost overruns in Swedish transport projects .................................................................................... 31

2.2.1 Descriptive statistics .................................................................................................................... 31

2.2.2 Differences in cost calculations .................................................................................................... 34

2.3 Key findings and discussion .............................................................................................................. 39

3 Factors contributing to cost overruns ............................................................................................... 41

3.1 Previous studies ............................................................................................................................... 41

3.2 Statistical methods ........................................................................................................................... 48

3.3 Factors of cost overruns in Swedish transport projects ................................................................. 49

3.3.1 Types of projects ........................................................................................................................... 49

3.3.2 Years of project completion .......................................................................................................... 50

3.3.3 Sizes of projects ............................................................................................................................ 55

3.3.4 Detailed project types ................................................................................................................... 58

3.3.5 Complexity ................................................................................................................................... 61

3.4 Key findings and discussion .............................................................................................................. 64
List of Figures

Figure 1- 1 Flowchart of study. ........................................................................................................... 18

Figure 2- 1 Cost overrun study. ......................................................................................................... 21
Figure 2- 2 Inaccuracy of cost estimates in 167 road and rail projects. ........................................ 31
Figure 2- 3 Inaccuracy of cost estimates in road projects (a) and rail projects (b). .................. 32
Figure 2- 4 Capital cost estimates – base estimate, inflation, and actual. ................................. 37
Figure 2- 5 Price indexes for Swedish road projects (a) and rail projects (b). ......................... 38

Figure 3- 1 Inaccuracy of cost estimates in Swedish transport projects over time, 1997-2009. 51
Figure 3- 2 Inaccuracy of cost estimates in road and rail projects over time, 1997-2009. ....... 52
Figure 3- 3 Inaccuracy of cost estimates in road and rail projects over time with SD, 1997-2009. ........................................................................................................................................ 54
Figure 3- 4 Forecasted costs and inaccuracy of cost estimates. ......................................................... 56
Figure 3- 5 Distribution of inaccuracy of cost estimates by project sizes. ................................... 56
Figure 3- 6 Distribution of inaccuracy of cost estimates in road projects (a) and rail projects (b) by detailed project types. ........................................................................................................... 59
Figure 3- 7 Complexity and inaccuracy of cost estimates. .............................................................. 62
Figure 3- 8 Complexity and inaccuracy of cost estimates by detailed project types. .................. 63

Figure 4- 1 Successive Calculation procedure. .................................................................................. 66
Figure 4- 2 Example of calculation details. ....................................................................................... 68
Figure 4- 3 Example of systematic and successive process. .......................................................... 69
Figure 4- 4 Example of S curve or cumulative distribution function related to the budget estimation. 71
Figure 4- 5 Example of the Successive Calculation result of the Stockholm bypass (Förbifart Stockholm). .......................................................................................................................... 74
Figure 4- 6 Standard deviation of normal distribution................................................................. 75
Figure 4- 7 Cost uncertainty during the project cycle. .................................................................... 76
Figure 4- 8 Phases of transport planning. ....................................................................................... 77
Figure 4- 9 Comparison between road project variances and actual outcome. .......................... 78
Figure 4- 10 Comparison between rail project variances and actual outcome. ......................... 79
Figure 4- 11 Road project variances in different project sizes. ...................................................... 82
Figure 4- 12 Rail project variances in different project sizes. ....................................................... 82

Figure 5- 1 Explanatory power of optimism bias and strategic misrepresentation, respectively, in accounting for forecasting inaccuracy as a function of political and organizational pressure..... 87
Figure 5- 2 Definition of Up-lifts within a certain class of transport scheme. .............................. 88
Figure 5- 3 Probability distribution of cost overruns for road projects according to the UK study. ..... 90
Figure 5- 4 Probability distribution of cost overruns for rail projects according to the UK study. ...... 91
Figure 5- 5 Probability distribution of cost overruns for fixed links projects according to the UK study. ........................................................................................................................................... 91
Figure 5- 6 Probability distribution of cost escalation for rail projects in the Netherlands. Based on reference class of 68 projects........................................................................................................................................... 94
Figure 5- 7 Road projects in Thailand............................................................................................................................................................................................................................................ 95
Figure 5- 8 Road projects in the Philippines............................................................................................................................................................................................................................................ 96
Figure 5- 9 COR histograms for highway and bridge projects in Vancouver Island.................................................................................................................................................. 96
Figure 5- 10 Inaccuracies of cost estimates of 167 road projects by the Erlang distribution. (database from Flyvbjerg et al., p.17[28])........................................................................................................................................... 103
Figure 5- 11 Inaccuracies of cost estimates of 58 rail projects by the Erlang distribution. (database from Flyvbjerg et al., p.17[28])........................................................................................................................................... 104
Figure 5- 12 Inaccuracies of cost estimations of the hybrid, RBE and convention methods by Lognormal distribution............................................................................................................................................................................ 104
Figure 5- 13 Examples of distributions in the thesis............................................................................................................................................................................................................................................ 106
Figure 5- 14 Stockholm bypass (Förbifart Stockholm)............................................................................................................................................................................................................................................ 108
Figure 5- 15 Västlänken............................................................................................................................................................................................................................................................................................................ 109
Figure 5- 16 Probability curves of road projects............................................................................................................................................................................................................................................................................................................ 111
Figure 5- 17 Distribution of cost overruns of road projects.................................................................................................................................................................................................................. 111
Figure 5- 18 Required uplift of road projects............................................................................................................................................................................................................................................................................................................ 112
Figure 5- 19 Cost estimations for road projects............................................................................................................................................................................................................................................................................................................ 113
Figure 5- 20 Probability curves of rail projects............................................................................................................................................................................................................................................................................................................ 115
Figure 5- 21 Distribution of cost overruns of rail projects............................................................................................................................................................................................................................................................................................................ 116
Figure 5- 22 Required uplift of rail projects............................................................................................................................................................................................................................................................................................................ 116
Figure 5- 23 Cost estimations for rail projects............................................................................................................................................................................................................................................................................................................ 117
List of Tables

Table 2- 1 Summary of reviewed studies. ................................................................. 29
Table 2- 2 Inaccuracy of transport project cost estimates by types of projects .......... 33
Table 2- 3 Comparison of cost overruns between Swedish transport projects and results of Flyvbjerg studies. .................................................................................. 34
Table 2- 4 Comparison of cost overruns between Swedish transport projects and result of other studies. .................................................................................. 34
Table 2- 5 Cost of major rail projects (2005-2009). .................................................. 39

Table 3- 1 Summary of factors contributing to cost overruns.................................. 47
Table 3- 2 Independent samples test by types of projects. ........................................ 50
Table 3- 3 F-statistic of average cost overruns in road and rail projects over time, 1997-2009. 52
Table 3- 4 F-statistic of average cost overruns in road and rail projects, before and after 2005. ................................................................................................. 53
Table 3- 5 Distribution of cost overruns in percent and in million SEK by project sizes........ 57
Table 3- 6 Distribution of cost overruns in percent and in million SEK by detailed project types. ................................................................................................. 60
Table 3- 7 Test of homogeneity of variance of detailed project types. .............. 60

Table 4- 1 Two examples of applications and results in cost estimation by the Successive Calculation ................................................................................................. 72
Table 4- 2 Descriptive statistics of road projects ..................................................... 78
Table 4- 3 Descriptive statistics of rail projects ....................................................... 79
Table 4- 4 Variances by project sizes ...................................................................... 81

Table 5- 1 Categories and types of projects used as basis for Reference Class Forecasting. ... 89
Table 5- 2 Required uplifts by the Reference Class Forecasting according to the UK study. .... 92
Table 5- 3 Descriptive statistics of estimation errors in the three samples in Australia .......... 97
Table 5- 4 COR probabilities of highway projects in Vancouver Island .................. 99
Table 5- 5 Cost overruns probability estimation comparison between Monte Carlo method and Beta Distribution Fitting method .................................................. 101
Table 5- 6 Goodness-of-fit Statistics of distributions of road and highway projects in VIHP. .... 102
Table 5- 7 Types of projects used as basis for Reference Class Forecasting of road projects. 110
Table 5- 8 Cost uplifts for selected percentiles of road projects ................................... 112
Table 5- 9 Types of projects used as basis for Reference Class Forecasting of rail projects .. 114
Table 5- 10 Cost uplifts for selected percentiles of rail projects .............................. 117
### List of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADB</td>
<td>Asian Development Bank</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of Variance</td>
</tr>
<tr>
<td>BART</td>
<td>San Francisco Bay Area Rapid Transit</td>
</tr>
<tr>
<td>CBA</td>
<td>Cost-Benefit Analysis</td>
</tr>
<tr>
<td>COR</td>
<td>Cost Overrun Ratio</td>
</tr>
<tr>
<td>CPI</td>
<td>Consumer Price Index</td>
</tr>
<tr>
<td>CV</td>
<td>Critical Value</td>
</tr>
<tr>
<td>FDOT</td>
<td>Florida Department of Transportation</td>
</tr>
<tr>
<td>FTA</td>
<td>Federal Transit Administration</td>
</tr>
<tr>
<td>HEATCO</td>
<td>Harmonised European Approaches for Transport Costing and Project Assessment</td>
</tr>
<tr>
<td>INDOT</td>
<td>Indiana Department of Transportation</td>
</tr>
<tr>
<td>JBIC</td>
<td>Japan Bank for International Cooperation</td>
</tr>
<tr>
<td>MLE</td>
<td>Maximum Likelihood Estimation</td>
</tr>
<tr>
<td>NPI</td>
<td>Net Price Index</td>
</tr>
<tr>
<td>PASW</td>
<td>Predictive Analytics Software</td>
</tr>
<tr>
<td>PPP</td>
<td>Public Private Partnership</td>
</tr>
<tr>
<td>RBE</td>
<td>Risk-Based estimating</td>
</tr>
<tr>
<td>RCF</td>
<td>Reference Class Forecasting</td>
</tr>
<tr>
<td>RRV</td>
<td>Riksrevisionsverket (Swedish National Audit Bureau)</td>
</tr>
<tr>
<td>SIKA</td>
<td>Swedish Institute for Transport and Communications Analysis</td>
</tr>
<tr>
<td>SD</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>SOC</td>
<td>Social Overhead Capital</td>
</tr>
<tr>
<td>SPSS</td>
<td>Statistical Package for the Social Sciences</td>
</tr>
<tr>
<td>SvD</td>
<td>Swedish newspaper Svenska Dagbladet</td>
</tr>
<tr>
<td>TIE</td>
<td>Transport Initiatives Edinburgh</td>
</tr>
<tr>
<td>UMTA</td>
<td>Urban Mass Transportation Administration</td>
</tr>
<tr>
<td>VIHP</td>
<td>Vancouver Island Highway Project</td>
</tr>
<tr>
<td>VTI</td>
<td>Swedish National Road and Transport Research Institute</td>
</tr>
<tr>
<td>WBS</td>
<td>Work Breakdown Structure</td>
</tr>
</tbody>
</table>
1 Introduction

1.1 Purpose and scope of this thesis

The purpose of this thesis is to study cost overruns in Sweden and internationally, factors affecting cost overruns and possible improvements of cost calculations.

The following topics are analysed in this thesis;

- The situation and trend for cost overruns in transport projects in Sweden and internationally.
- General descriptive statistics for transport projects in Sweden.
- Factors that affect cost overruns in transport projects.
- How cost calculations are currently carried out in Sweden, particularly how the Successive Calculation method is applied.
- A probability distribution of cost overruns for transport projects in Sweden. This is used to test the Reference Class Forecasting method for two case studies.

Transport projects contribute to the development of societies. The countries that provide efficient transport investments with good management have develop in social, economic and environment dimensions. Unfortunately, transport investments are plagued by cost overrun and benefit shortfall problems. The implementations of transport projects with such problems will lead to risks for the wrong prioritization and costly delay. These problems are found not only in transport projects but occur in other kind of projects such as water projects, power plants, ICT systems and urban and regional development projects [1][2].

The desire of decision makers is to improve the returns from the limited budget. In this process, Cost-Benefit Analysis (CBA) has been an important tool for evaluating and ranking transport investments for several decades [3]. The cost overrun and benefit shortfall problems lead to incorrect CBAs because CBA depends on cost and benefit forecasts. However, this thesis only concentrates on cost estimates.

In recent years there has been an increased interest in cost overruns in transport projects in Sweden and elsewhere e.g. in Norway, Australia, England and Canada. New methods, such as Successive Calculation and Reference Class Forecasting, have been developed to improve the cost calculations. These methods are intended to lead to better cost estimates and thus better CBAs as well as a better control over public investment budgets. However strong emphasis on decreasing cost overruns, may induce incentives to adapt in both desirable and undesirable ways. Examples of undesirable consequences could be;
• Adding reserve funds to projects may lead to heterogeneous assessments and deceptive representations of relative merits of projects. Such reserves may in turn induce lead risks for inefficient designs and overspending. Good incentives in the planning phase are therefore important.

• Undesired reductions of quality: With too much emphasis on reaching cost targets projects are likely to be constructed below at the intended standard levels. Some items for implementing on projects such as environmental and safety aspects risk being compromised or even being removed. In such cases the real costs can become higher as the deficiencies have to be corrected later.

• Incentives to postpone the construction of dimensions representing secondary objectives like environment and safety.

These potential consequences will also be addressed in this thesis. How to deal with cost overrun problem is the motivation to do this research. Lessons learnt from Swedish transport projects are studied in order to find the better methods and/or practical applications of cost calculation. Our hope is that this will consequently contribute to more efficient transport planning in the future. Moreover, the thesis can be a part of an evaluation of the first use of Successive Calculation in transport planning in Sweden.

1.2 Overview of state-of-the-art

This section gives an overview of the cost overrun problems and the methods to handle them. More about different aspects of this are found in each chapter of the thesis.

The definition of the cost overruns or inaccuracy in cost forecasts is measured as actual cost minus forecasted cost as a percentage of forecasted cost. An inaccuracy of zero shows that the forecasted cost for the project is correct and thus equals actual cost. Forecasted cost is the estimate made at the time of decision to build, or as close to this as possible. Actual cost is the output of construction cost measured after the project is completed. All costs are calculated in constant prices or inflation-adjusted currencies.

From the studies of Bent Flyvbjerg and others, key observations about the cost overruns in transport projects (based on a sample of 258 transport projects in 20 countries on 5 continents) are:

• 9 out of 10 projects have cost overruns;
• Overrun is found across the 20 nations and 5 continents;
The average cost overruns are 45%, 34% and 20% for rail projects, bridges and tunnels, and road projects, respectively;

Overrun is constant for the 70-year period and cost estimates have not improved over time.

Many transport projects have experienced cost overruns. To mention only a few:

- One is the Channel tunnel, the longest underwater rail tunnel in Europe, connecting France and the UK\textsuperscript{[2]}. It opened in 1994. The construction cost was 80% over budget and revenues were half of those forecasted. As a consequence, the internal rate of return was negative, at -14.5%. An economic and financial \textit{ex post} evaluation of the project concluded that “the British Economy would have been better off had the Tunnel never been constructed”\textsuperscript{[2]}.

- The next example is the Danish Great Belt rail tunnel, the second-longest underwater rail tunnel in Europe. It opened in 1998. The construction cost was 120% over budget and the project proved nonviable even before it opened\textsuperscript{[2]}. Only by cross-subsidizing the tunnel with revenues from a nearby motorway bridge was it possible to pay for the tunnel\textsuperscript{[2]}.

- Another example is the Ōedo Line, Tokyo's first linear motor metro line which allows it to use smaller cars and smaller tunnels\textsuperscript{[7]}. It opened in 2000. The original budget was ¥682.6 billion and the final cost of this construction ranged from ¥988.6 billion to over ¥1,400 billion or 105% of cost escalation\textsuperscript{[7]}.

- The next example is the Central Artery/Tunnel project (CA/T), known unofficially as the Big Dig. It is a megaproject in Boston that rerouted the Central Artery (Interstate 93) into a 5.6-km tunnel\textsuperscript{[8]}. Although the CA/T project cost was estimated in 1985 at US$6.0 billion (adjusted for inflation as of 2006), over US$14.6 billion had been spent as of 2006 or a 143% cost escalation\textsuperscript{[8]}.

- The last example is Edinburgh Trams. It is a tramway system which is currently under construction (2011) in Edinburgh, Scotland\textsuperscript{[9]}. With an original budget at a cost of £375 million in 2003, the cost of this tram system is now uncertain, and it is anticipated to be over £600 million\textsuperscript{[9]}. It was originally scheduled to enter service in February 2011 but had to postpone the opening to 2014 because of the budget problem. In February 2011, there was an article – “Time’s up: Today's the day the first tram should have been running in Edinburgh” published by Edinburgh Evening News\textsuperscript{[10]}. The article stated that “The German engineering giants charged with building the line have revealed that 72% of the construction work remains with just 38% of the budget left.”
The above projects are all examples of the cost overruns in transport megaprojects. However, it is also a serious problem for routine or small-sized transport projects. For instance, the average cost overrun was 5.9% for 127 road projects in the Vancouver Island Highway Project (VIHP) implemented during 1993-2003. Furthermore, a study of 620 Norwegian road projects constructed during 1992-1995 showed that there was 7.9% cost overrun on average. A recent study revealed an average 19% cost overrun for 36 road construction projects in Slovenia (1995-2007).

In Sweden, like other countries, the development of the infrastructure of the country is a major term in the state budget. In December 2008, the Swedish Parliament decided to allocate 417 billion SEK for the transport infrastructure projects in the period 2010-2021. Cost overrun problem is one of the key challenges for those investments since the problem has been large in previous Swedish transport projects. The studies about the cost overrun problems in Swedish transport projects are presented below.

The first study was conducted in 1994 by the Swedish National Audit Bureau (Riksrevisionsverket: RRV). They investigated the performance of 8 road projects built by the Road Administration (Vägverket) and 7 rail projects by the Rail Administration (Banverket). The average cost overrun for the road projects was 86% (2-182%) and the average overrun for the rail projects was 17% (-14-74%). In 2002, there was a study by the Swedish Institute for Transport and Communications Analysis – SIKA. They examined cost outcomes of road and rail projects which opened to service during 1997-1999. The key finding was that the average cost escalations were 10% and 20% for road and rail projects, respectively. SIKA also provided one more study about cost overruns in 2002. They concluded that the average cost overruns were 5% for the road projects and 14% for the rail projects which were opened during 1997-2000. SIKA recommended that the Transport Administrations should calculate expected cost overruns based on these outcomes including the standard deviations.

Recently, there were two audit reports provided by the Swedish National Audit Office (Riksrevisionen). These reports showed the different cost overruns compared with calculations made by the Transport Authorities. The main reasons for differences are found in assumptions such as price index and what decision is used as the estimated cost. The first audit report was published in 2010 and showed the results from road projects. One of the key results was that the total costs of road investments completed between 2005 and 2009 based on the Swedish National Audit Office calculations increased at 8-18% while based on the Vägverket calculations.

---

1 The Road and Rail Administrations were merged into a Transport Administration (Trafikverket) in April 2010.
they were around -1% (means cost saving or no cost overrun). The second study was published in 2011 and concerned rail projects \cite{14}. A finding was that the average cost overrun for the rail projects completed between 2005 and 2009 based on the Swedish National Audit Office calculations was 55% while based on the Banverket calculations they were on average 26%. Moreover, the Swedish National Audit Office concluded that the cost overruns problem as noted by the RRV 15 years ago still remained and they could not see that any real improvements had occurred in this aspect.

Responding to the last audit report by the Riksrevisionen, the Government has recently reported that several steps have been taken to create cost control for major rail investments \cite{19}. One of the most important steps was the introduction of the Successive Calculation. The Successive Calculation is a cost calculation method that was developed in the 1970’s in Denmark. It has had a limited use in Sweden, but was used to calculate the project costs in the latest national transport investment plan for the period 2010-2021. The government has emphasized that experiences are needed to be drawn from this first major use of the Successive Calculation method.

Beyond the studies of cost overruns problem above, the Swedish media also publish their own reports on this problem. For example, the Swedish newspaper Svenska Dagbladet (SvD) published an article on 6 April 2011 \cite{20} – “Uncontrollable note for light rail construction - Skenande nota för spårvägsbyggen”. SvD revealed that the final bill for the first part of the Stockholm ongoing light rail or city tram (between Hamngatan and Waldemarsudde) will be 380 million SEK which is 67% more than the planned cost (228 million SEK).

In sum, the cost overrun problem is one of the most important problems in transport planning. According to the literature, the cost estimates used in transport planning are systematically and significantly misleading. Cost overrun problems seem to be a global phenomenon. Thus, we cannot avoid confronting the problem.
1.3 Methodology

1.3.1 Research design

The study process has been as shown in Figure 1-1. It was initiated by doing literature studies about the cost overruns in overall.

We study the situation and trends for cost overruns in transport projects in Sweden and internationally. The material used for the study come from academic papers and independent government audits. Regarding cost overruns in Sweden, we also analyse transport project cost data from a database collected by VTI, Swedish National Road and Transport Research Institute (detail of the data is shown in next section) and complemented in this study. The output are general descriptive statistics such as average overruns and standard deviations (SD) for both road and rail projects. The general descriptive statistics are compared with the results from previous studies.

The next step is the study of what factors affect cost overruns for transport projects in Sweden and internationally. In the VTI data, we test whether different main project types (i.e. road and rail projects) perform differently with regard to cost overruns. We also test whether overruns varies with year of project completion, size of project, detailed project type and complexity.

Then, we study the current method used to estimate project costs in Sweden – the Successive Calculation. The general procedure of the method including applications and limitations is reviewed. Next, how the method has been applied in Swedish investment planning is studied by analysing a database on results of cost calculations using the method. To see the cost variance, we calculate cost at 85% confident limit (the probability of keeping to budget) minus cost at 50% confident limit as a percentage of cost at 50% confident limit. The average cost variances by this method are then compared to the actual outcomes. In addition, we analyse the cost variance in each planning phase.

Next, a new method called “Reference Class Forecasting (RCF)” is studied. As with the study of the Successive Calculation, the general procedure of the RCF method including applications and limitations is reviewed. The key concept of this method is to examine the experiences of a class of similar projects, lay out a rough distribution of outcomes for the reference class, and then position the current project in that distribution. Before demonstrating how to apply this method, we review which distribution provides the best fit for the VTI project cost database. Then, we apply the RCF method to estimate the project costs of our case studies. The case study for road projects is the Stockholm bypass (Förbifart Stockholm). The case study for rail
projects is the Västlänken tunnel in Göteborg. The cost estimations of the two case studies are compared with the total costs estimated by the Successive Calculation method. The cost variance results by the RCF method are compared with the cost variance analysed by the Successive Calculation method. Lastly, we conclude our findings and suggest future researches.

1.3.2 Data

In the thesis, there are two databases as shown in Figure 1-1. The first database is data on cost estimations and outcomes of road and rail projects in Sweden which were completed during 1997-2009. The data are published in yearly reports from Vägverket and Banverket which had originally been collected at the VTI and has then been complemented in our study. After refining the data, 102 road projects and 65 rail projects are used in our analysis. The second database is data on results of cost calculations using the Successive Calculation method. They are collected from the projects in the latest national investment plan that covers the period 2010-2021. For road projects, we receive the data from the Transport Administration calculation sheets. For rail projects, we find the data from the Transport Administration website. The data comprise 249 road projects and 46 rail projects.

1.3.3 Disposition

The thesis is organized as follow. This chapter serves as a first chapter which provides general information of this thesis including cost overrun background. Chapter 2 further describes the cost overrun problems in Sweden and internationally based on previous studies. This chapter also provides the general descriptive statistics for transport projects in Sweden. Chapter 3 presents factors that affect cost overruns in transport projects. Chapter 4 and Chapter 5 show the methods to improve cost estimations – the Successive Calculation and the Reference Class Forecasting, respectively. Finally, Chapter 6 gives conclusions and a discussion, as well as ideas for future researches.
Figure 1-1 Flowchart of study.
2. Cost overrun problem

Cost overrun, sometimes called ‘cost increase’ or ‘cost escalation’, in transport projects is one of the most important problems in transport planning as mentioned earlier. However, studies about cost overruns are rather few. Unlike other kind of studies, the rigorous study of cost overruns not only comes from academics but also independent government audits\(^{[21]}\) as shown in Figure 2-1. The conventional cost overrun study has been carried out by academics at universities and research institutes. The data of the studies are collected by the researchers themselves (published sources or direct contacts), or come from other studies and audit reports. The study results, including recommendations and new methods for better cost estimation are published in scholarly articles, books and journals. In some cases, the transport organizations or authorities employ academics and/or private consultants to study the cost overruns of their projects. In these cases, the data can be acquired directly from the transport organizations. Sometimes, the transport authorities do the cost overrun analysis by themselves.

The independent government auditors are responsible for monitoring the accountability, effectiveness and efficiency of public spending. They give recommendations to the Parliament on how to improve the use of the national budget. In some countries, government auditor provides detailed investigations to explain the causes of cost overruns and study the frequency and magnitude of cost overrun. Even though, the academics and the independent government auditors have the same main interest which is cost overruns in transport projects, they have different mandates, objectives and access to data. The academics use data from the independent government auditors to their studies. However, the academics mainly refer to other academic studies and often conduct less systematic analysis than independent government auditors. In recent years, it has been found that the independent government auditors more often refer to the academic studies than previously. An increased cooperation between the two groups is needed in the future for better understand cost overruns and how they can be avoided.

The definition of the cost overrun or inaccuracy in cost forecasts is measured as actual cost minus forecasted cost as a percentage of forecasted cost (Equation 2.1)\(^{[5]}\). Forecasted cost is the estimate made at the time of decision to build, or as close to this as possible if no estimate was available for the decision to build. Actual cost is the construction cost measured after the project was completed. All costs are calculated in constant prices or inflation-adjusted currencies. This definition is used in the majority of the scholarly studies\(^{[21]}\).
For academics, the early studies of cost overruns in transport projects are the studies of Merewitz in 1973 [22]. The Merewitz’s studies compared cost overrun in urban rapid transit projects including the San Francisco Bay Area Rapid Transit (BART) system. However, Flyvbjerg and others [22] claimed that these studies had not produced statistically valid results. Moreover, they stated that their study in 2002 was the first statistically significant study of cost overruns in transport projects. Based on the 258 transport projects in 20 nations, they found that 9 out of 10 transport projects had cost escalation. The average cost overruns were 45%, 34% and 20% for rail, fixed-link, and road projects, respectively. They concluded that the cost escalation appeared as a global phenomenon. The overrun was constant for the 70-year period and cost estimates had not improved over time.

Regarding independent government auditors, Siemiatycki [21] studied 13 audit reports (9 reports from United States, 2 reports from England and 2 reports from Canada). These audit reports cover rail, road and bridge projects completed between 1980 and 2007. The auditors studied the performances of delivering transport projects by variety of sample sizes and different types of investments. They systematically measured the frequency and magnitude of cost overruns in transport sector.

In this chapter, the previous cost overrun statistics results are presented and summarized. The VTI data is also analysed. Only statistics of the cost overruns are focused in this chapter such as average cost overruns and standard deviations.

\[
\% \text{Difference of cost} = \frac{(\text{Actual cost} - \text{Forecasted cost})}{\text{Forecasted cost}} \times 100 \quad \text{Equation 2.1}
\]
The Independent government auditors are responsible for monitoring the accountability, effectiveness, efficiency, and probity of public spending and providing recommendations to Parliament on how to improve using the national budget.

In some countries, government auditors provided detailed investigations to explain the causes of cost overruns and studied the frequency and magnitude of cost overrun.

Figure 2-1 Cost overrun study.

Note: Adapt from Siemiatycki [21], Flyvbjerg et al. [22] [23]
2.1 Previous studies

The summary of previous studies is as shown in Table 2-1. The following is the brief result of each study.

**Urban Mass Transportation Administration (UMTA), 1990** \[24\]

The study was carried out by the UMTA (currently the Federal Transit Administration: FTA). This study covered 10 major transit capital investment projects in the U.S. They were constructed with partial federal financing during the period 1971-1987, with a total value of US$ 15.5 billion at 1988 prices. 9 of 10 projects had cost overruns. The average cost overrun was 52% with the standard deviation (SD) \(^1\) 29%.

**Transport and Road Research Laboratory, 1990** \[25\]

The Transport and Road Research Laboratory in the England studied the performance and impact of Rail Mass Transit in 21 developing countries. Only 13 of the 21 metros capital cost overruns could be estimated. Almost all of the metro systems incurred higher costs than expected. Only the Hong Kong mass transit railway and the Proto Alegre metro in Brazil were completed within budget. 6 metros had overruns above 50%.

**Riksrevisionsverket, 1994** \[15\]

The Swedish National Audit Bureau (Riksrevisionsverket: RRV) investigated the performance of 8 road projects and 7 rail projects in Sweden with a total value of 13 billion SEK at 1994 prices. The average capital cost overrun for the road projects was 86%, ranging from 2 to 182%, and the average overrun for the rail projects was 17%, ranging from minus 14 to 74%. However, two thirds of the projects were still under construction when the study was carried out. Therefore the actual costs for these projects might be higher than the expected actual costs by the RRV \[26\].

**Office of Program Policy Analysis and Government Accountability, 1996** \[27\]

The Office of Program Policy Analysis and Government Accountability reviewed the Florida Department of Transportation’s (FDOT) performance in controlling cost overruns and delays for road and bridge projects. They investigated 3,969 construction contracts in Florida completed during 1980-1995. Cost overruns had increased from an average of less than 2% during fiscal years 1980-81 through 1984-85 to an average of 15% by fiscal year 1994-95.

---

\(^1\) The standard deviation (SD) is the square root of the average error between the mean and the observations made \[42\] (see more details in Appendix B).
Skamris and Flyvbjerg, 1997 [26]

This article studied inaccuracy of traffic forecasts and cost estimates on large transport projects. On seven large Danish bridge and tunnel projects since 1960, construction costs had been underestimated (average cost overrun was 14%) and traffic had been overestimated in the initial phases of planning. Cost overrun and benefit shortfall pattern is also found in studies from other countries of large transport projects. The result of this overoptimism in the initial phases of planning was that decisions were based on misleading forecasts that might lead to a misallocation of funds and underperforming projects.

Swedish Institute for Transport and Communications Analysis (SIKA), 2002 [16],[17]

In 2002, there was a study by the SIKA. They examined cost outcomes of road and rail projects in Sweden which were opened to service during 1997-1999. The key finding was that the average cost escalations were 10% and 20% for road and rail, respectively. SIKA also provided one more study about cost overruns in 2002. They concluded that the average cost overruns were 5% for the road projects and 14% for rail projects which were opened during 1997-2000. They recommended that the Transport Administrations should calculate expected cost overruns based on these outcomes including the standard deviations.


In 2002, Bent Flyvbjerg and others carried out the first statistically significant study of cost performance in transport projects. This is the only one study that covered the international comparison. Based on a sample of 258 transport projects in 20 nations completed in 1927-1998, they found that 9 out of 10 transport projects had cost overruns. On average, cost overruns were 45% (SD=38%), 20% (SD=30%) for 58 rail projects and 167 road projects, respectively. They concluded that cost overrun was a global phenomenon. The overrun was constant for the 70-year period and cost estimates had not improved over time.

Federal Transit Administration (FTA), 2003 [29]

As mentioned, the UMTA [24] studied 10 major transit projects in the U.S. which opened between 1975 and 1990. Later, the FTA investigated 21 rail and busway projects that opened between 1990 and 2002. The capital costs of projects on average were 21% greater than the initial estimates.
Odeck, 2004 [12]

The study showed the cost overruns of road construction based on data of 620 Norwegian road constructions during 1992-1995. The mean cost overrun was 7.9% ranging from -59% to 183% (SD=29.2%). An interesting finding was that cost overruns appeared to more predominant among smaller projects as compared to larger ones. Other factors found to influence the size of cost overruns included completion time of the projects and the region where projects are situated.

Indiana Department of Transportation (INDOT), 2004 [30]

The results of an agency survey showed that with regard to the problem of cost overruns, the INDOT had an average rank compared to other states. The overall rate for cost overrun amounts for the 2,668 road construction and maintenance projects by the INDOT between 1996 and 2002 was 4.5%. Moreover, 55% of all INDOT contracts experienced cost overruns. It was determined that the average cost overrun amount and rate differed by project type. The average cost overrun rates of bridge projects, road construction, road resurfacing, traffic projects and maintenance projects were 8.1%, 5.6%, 2.6%, 5.6% and 7.5%, respectively

Dantata et al., 2006 [31]

This paper compared the results of the UMTA report [24] to cost overruns of transit projects completed after 1990. The projects were on federally funded rail projects in the United States. From comparison with selected 16 transit rail projects, they concluded that there was evidence to suggest that cost overruns for projects completed before 1990 were different from that of projects complete after 1994 (i.e. cost overruns had become smaller from 52% to 30%). However, the data was not statistically proved this at a level of significance of 5%.


The study was analysed common risk factors and proposed analysis models for cost overrun risk analysis in transport investments. The Vancouver Island Highway Project (VIHP) database was used. There were 127 road and highway construction projects, 36 bridge and tunnel projects in Canada implemented during 1993-2003. The average cost overruns were 5.9% (SD=27%) and 5.2% (SD=23%) for road and bridge projects, respectively.
UK National Audit Office, 2007 [32]

The study was examined how the costs of building and improving roads were estimated and monitored from early forecasts through to the final cost of schemes. The UK Department for Transport had approved expenditure of over £11 billion between 1998 and 2021 for the development of new and existing trunk roads and motorways in the UK by the Highways Agency and under £1.7 billion on major road schemes which were proposed and developed by the local authorities. By 2006, the 36 schemes by the Highways Agency had been completed and had cost 6% more than estimated. By 2006 the 20 schemes by the local authorities completed had cost 18% more than initially estimated.

Flyvbjerg, 2007 [33]

The study presented the risks of cost and demand in urban rail projects. “Urban rail” was defined as rail in an urban area, including both heavy and light rail, which may be underground, at level or elevated. A sample of 44 urban rail projects was analysed (1966-1997). From the previous data of Flyvbjerg’s study [22], the 44 urban rail projects were a subset of the data (58 rail projects). 18 projects were located in North America, 13 projects were in Europe and 13 projects were in developing nations. The average cost overrun of the urban rail projects was 44.9% (SD=37.3%). It also reported the average cost overrun of the other rail projects (14 rail projects) was 44.1% (SD=43.3%).

Roxas Jr. and Chalermpong, 2008 [34]

They studied the cost forecasting inaccuracies in road and bridge projects in Thailand and the Philippine. It had been verified that there were no significant differences in cost forecasting inaccuracies between road and bridge projects. In Thailand, average cost overrun of 44 road and bridge projects was -10.8% (SD=30.5%)2. The negative value for the average cost overrun

---

2 The authors indicated that the forecast inaccuracies found in Thailand and the Philippines were much smaller in magnitude compared to those found from previous studies. However, due to the small sample size of the database in this study, it might not be a representative of the population used and thus the results should be interpreted with caution. Moreover, most of the projects considered in this research were funded by international agencies such as the Asian Development Bank (ADB), World Bank and the Japan Bank for International Cooperation (JBIC). Therefore, locally funded projects and Public Private Partnership (PPP) projects were not well represented in the database. They also explained that Flyvbjerg’s database was bigger and more diverse. There were more projects included and the project size in terms of the project cost was much bigger when compared to their study.
indicated that the actual cost was less than the forecasted cost. In Philippines, average cost overrun of 85 road and bridge projects was 5.4% (SD=35.9%).

Lee J.K., 2008 [35]

This study was carried out cost overrun and cause in Korean Social Overhead Capital (SOC) projects i.e. roads, rails, airports, and ports. There were data of 161 completed projects, which was including 138 road projects, 16 rail projects, 2 airport projects and 5 port projects from 1985 and 2005. The results indicated that 95% and 100% of road and rail projects, respectively, had a maximum cost overrun of 50%. The average cost overruns were 11% and 48% for road and rail projects, respectively.

Federal Transit Administration, 2008 [36]

The study based on the previous studies by the UMTA [24] in 1990 and the FTA [29] in 2003. It was conducted an analysis of the predicted and actual impacts of 21 major transit projects in the U.S. (opened 2003-2007). On average, the actual construction costs exceeded the inflation-adjusted estimates by 40.2%. The average cost overrun in this study was higher than the result of study in 2003. However, the cost estimates of the projects in 2003 and 2007 were more accurate than was found in the 1990 UMTA study. Thus, there was an improvement in cost estimations for major transit projects in the U.S.

Singh, 2009 [37]

This study investigated the delays and cost overruns in public funded infrastructure projects in India. The author concluded that the problem of time and cost overruns in India was widespread and severe. The 894 projects from 17 infrastructure sectors in the period 1992-2009 were analysed including 122 railway projects and 157 road and highway projects. The railway projects had average cost overruns at 95% (SD=179%) while road projects had cost overrun at 16% (SD=62%).
RGL Forensics, Faber Maunsell/Aecom and Frontier Economics, 2010 [37]

The paper provided for the European Commission on *Ex post* Evaluation of Cohesion Policy Programmes 2000-2006. They analysed 19 rail, 21 road and 7 urban transport projects located in Germany, Spain, France, Greece, Ireland, Italy, Poland and Portugal. It showed that average cost overruns were 26.9%, 9.4% and 45.4% for rail, road and urban transport, respectively. The study was also studied water and energy projects and the total average cost overruns was 21.2%.

Chevroulet and Reynaud, 2010 [38]

The study investigated 6 European high speed rail projects i.e. Frankfurt-Cologne ICE, Eurotunnel, Öresund Fixed Link, Paris-Lille TGV, Madrid-Sville AVE and Lyon-Marseilles TGV. The cost overruns ranged from 8% to 116% with 51% on average (SD=40%).


There were two audit reports provided by the Swedish National Audit Office (Riksrevisionen). They showed the differences of cost overruns compared with calculations made by the Transport Authorities. The main reasons are differences in assumptions such as price index and what decision is used as the estimated cost. The first audit report was published in 2010 and showed results from 35 road projects completed between 2005 and 2009 [18]. One of the key results was that cost overruns of road investments were at 8-18%. The second study was published in 2011 and concerned 28 rail projects completed between 2005 and 2009 [14]. A finding was that the average cost overrun for the rail projects was 55%.

Makovsek et al., 2011 [13]

In this study, the authors focused on cost performance of road infrastructure constructed through the National Motorway Construction Programme in Slovenia. It was found that 19% (SD=46.1%) cost overrun for 36 road projects which were completed in the period 1995-2007.
Summary

The summary of the literature above is shown in Table 2-1. The literature confirms that the cost overrun problem in transport projects is a global problem. Cost overruns in transport projects happen around the world – the U.S., Canada, the Philippines, South Korea, India, Sweden, England and Slovenia. Most of the studies focus on the problems on a national level and two of the studies analyses the problems on a continental level (in Europe). There is only one database used in the studies of Bent Flyvbjerg and others that shows international results of cost overruns in transport projects.

Almost all of the studies that examines both road and rail projects (6 of 7 studies) show average cost overruns for road projects that are lower than the cost overruns for rail project. In road projects, the mean cost overruns range between 4.5% (2,668 road projects in Indiana, USA) and 86% (8 road projects in Sweden). In rail projects, the cost overruns are between 14% (rail projects in Sweden) and 95% (122 railway projects in India). Cost overruns of less than 10% are found in 8 studies and all of them deal with cost overruns for road projects. In the next section, comparisons between cost overruns in Swedish transport projects and the results of the above studies are shown.

---

3 Only one study by Roxas Jr. and Chalermporn showed cost underrun but the study was not well represented results as explained.

4 There was only one study by Riksrevisionsverket in 1994 that indicated that the average cost overrun for road projects was higher than the average cost overrun for rail projects.
Table 2-1 Summary of reviewed studies.

<table>
<thead>
<tr>
<th>Number and types of project examined</th>
<th>Years covered by study sample</th>
<th>Countries covered in sample</th>
<th>Average estimated cost per project</th>
<th>Average difference of cost (%)</th>
<th>Standard deviation</th>
<th>Min. - Max. difference of cost (%)</th>
<th>Percentage of projects experiencing escalating project costs</th>
<th>Authors/Year of publish</th>
<th>Type of study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ten major transit projects</td>
<td>1971-1987</td>
<td>USA.</td>
<td>$1.5 billion</td>
<td>52%</td>
<td>29</td>
<td>-11% - 106%</td>
<td>90% UMTA /1990</td>
<td>Academic</td>
<td></td>
</tr>
<tr>
<td>15 projects: 7 rail; 8 road</td>
<td>Before 1994</td>
<td>Sweden</td>
<td>$130 million*</td>
<td>7%</td>
<td>N/A</td>
<td>-14% - 74% 2% - 182%</td>
<td>71% 100% Riksrevisionsverket/1994</td>
<td>Audit</td>
<td></td>
</tr>
<tr>
<td>3,969 construction contracts by FDOT</td>
<td>1980-1995</td>
<td>Florida, USA.</td>
<td>$1.7 million</td>
<td>7%*</td>
<td>N/A</td>
<td>0.8% - 15.1%</td>
<td>N/A Office of program Policy Analysis and Government Accountability /1996</td>
<td>Audit</td>
<td></td>
</tr>
<tr>
<td>Seven large bridge and tunnel projects</td>
<td>Since 1960</td>
<td>Denmark</td>
<td>N/A</td>
<td>14%*</td>
<td>N/A</td>
<td>-10% - 33%</td>
<td>71% Skamms, Flyvbjerg/1997 Academic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rail and road projects</td>
<td>1997-2000</td>
<td>Sweden</td>
<td>N/A</td>
<td>Rail: 14% Road: 5%</td>
<td>27</td>
<td>N/A</td>
<td>N/A SiKA/2002 Academic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>258 projects: 58 rail; 33 fixed-link; 167 road</td>
<td>1927-1998</td>
<td>20 countries on 5 continents</td>
<td>$348 million ($1.5 billion - $8.5 billion)</td>
<td>Overall: 28% 58 rail: 45% 33 fixed-link: 34% 167 road: 20%</td>
<td>39 38 62 30</td>
<td>N/A</td>
<td>N/A 86%</td>
<td>Flyvbjerg, Holm, and Buhl/2002, 2003 and Priemus, Flyvbjerg and Wee/2008</td>
<td>Academic Audit</td>
</tr>
<tr>
<td>21 rail and busway projects</td>
<td>1990-2002</td>
<td>USA.</td>
<td>$524 million ($98 million - $4.4 billion)</td>
<td>21%</td>
<td>N/A</td>
<td>-28% - 72%</td>
<td>76% Federal Transit Administration/2003 Academic*</td>
<td>Academic*</td>
<td></td>
</tr>
<tr>
<td>620 road projects</td>
<td>1992-1995</td>
<td>Norway</td>
<td>Less than $100 million</td>
<td>8%</td>
<td>29</td>
<td>-59% - 183%</td>
<td>52% Odeck/2004 Academic</td>
<td>Academic</td>
<td></td>
</tr>
<tr>
<td>2,668 road construction and maintenance projects</td>
<td>1996-2002</td>
<td>Indiana, USA.</td>
<td>N/A</td>
<td>4.5%</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A 55% INDOT/2004 Academic*</td>
<td>Academic*</td>
<td></td>
</tr>
<tr>
<td>16 urban rail</td>
<td>1995-2004</td>
<td>USA.</td>
<td>$486 million ($94 million - $1.625 million)</td>
<td>30%</td>
<td>39</td>
<td>-28% - 133%</td>
<td>81% Dantata et al./2006 Academic</td>
<td>Academic</td>
<td></td>
</tr>
<tr>
<td>127 road projects</td>
<td>1993-2003</td>
<td>Canada</td>
<td>N/A</td>
<td>127 road: 5.9% 36 bridge: 5.2%</td>
<td>27 23</td>
<td>4.8% - 23.4% 8.0% - 19.0%</td>
<td>82% 81% Qing Wu/2006 Academic</td>
<td>Academic</td>
<td></td>
</tr>
<tr>
<td>36 road projects by National Highway Agency</td>
<td>20 road projects by Local Authorities</td>
<td>2002-2006</td>
<td>$65 million $24 million</td>
<td>National Highway: 6% Local Authorities: 18%</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A 64%</td>
<td>UK National Audit Office/2007 Audit</td>
<td></td>
</tr>
</tbody>
</table>

1 The study was studied by academic and/or private consultant. It was sponsored by transport authority.
2 Adjust by current exchange rate (1 SEK = $0.15).
3 The average value was calculated from cost overruns in all fiscal years of the study.
4 The average construction cost overrun was only for the five completed projects.
<table>
<thead>
<tr>
<th>Number and types of project examined</th>
<th>Years covered by study sample</th>
<th>Countries covered in sample</th>
<th>Average estimated cost per project</th>
<th>Average difference of cost (%)</th>
<th>Standard deviation</th>
<th>Min. - Max. difference of cost (%)</th>
<th>Percentage of projects experiencing escalating project costs</th>
<th>Authors/Year of publish</th>
<th>Type of study</th>
</tr>
</thead>
<tbody>
<tr>
<td>44 urban rail projects</td>
<td>1966-1997</td>
<td>18 in North America 13 in Europe 13 in developing nations</td>
<td>N/A</td>
<td>44.9%</td>
<td>37.3</td>
<td>N/A</td>
<td>N/A</td>
<td>Flyvbjerg/2007</td>
<td>Academic</td>
</tr>
<tr>
<td>129 road and bridge projects</td>
<td>Before 2008</td>
<td>85 in Philippines 44 in Thailand</td>
<td>N/A</td>
<td>Philippines: 5.4% Thailand: -10.8%</td>
<td>36 30</td>
<td>-67% - 167% -59% -106%</td>
<td>N/A</td>
<td>Roxas Jr., Chalermpoon/2008</td>
<td>Academic</td>
</tr>
<tr>
<td>161 projects; 138 road 16 rail 2 airport 5 port</td>
<td>1985-2005</td>
<td>South Korea</td>
<td>Road $19 million Rail $ 455 million</td>
<td>Road: 11% Rail 48%</td>
<td>N/A</td>
<td>N/A</td>
<td>Road 87% Rail 94%</td>
<td>Lee J.K./2008</td>
<td>Academic</td>
</tr>
<tr>
<td>21 major transit projects</td>
<td>2003-2007</td>
<td>USA</td>
<td>$566 million ($58 million - $2.2 billion)</td>
<td>40%</td>
<td>N/A</td>
<td>-1% - 185%</td>
<td>94%</td>
<td>Federal Transit Administration/2008</td>
<td>Academic</td>
</tr>
<tr>
<td>894 projects from seventeen infrastructure sectors; 157 road and highway 122 railway transport projects; 19 rail 21 road 7 urban transport</td>
<td>1992-2009</td>
<td>India</td>
<td>N/A</td>
<td>157 road and highway: 16% 122 railway: 95%</td>
<td>62 179</td>
<td>N/A</td>
<td>54% 83%</td>
<td>Singh/2009</td>
<td>Academic</td>
</tr>
<tr>
<td>6 European high speed rail projects</td>
<td>Before 2010</td>
<td>Europe</td>
<td>$ 6 billion</td>
<td>51%</td>
<td>40%</td>
<td>8% - 116%</td>
<td>100%</td>
<td>Chevroulet, Reynaud/2010</td>
<td>Academic</td>
</tr>
<tr>
<td>35 road and 28 rail projects</td>
<td>2005-2009</td>
<td>Sweden</td>
<td>N/A</td>
<td>Road : N/A Rail : 55%</td>
<td>N/A</td>
<td>N/A</td>
<td>8% - 18%</td>
<td>N/A</td>
<td>Riksrevisionen/2010, 2011</td>
</tr>
<tr>
<td>36 road projects</td>
<td>1995-2007</td>
<td>Slovenia</td>
<td>N/A</td>
<td>19%</td>
<td>46</td>
<td>N/A</td>
<td>N/A</td>
<td>Dejen Makovsek et al./2011</td>
<td>Academic</td>
</tr>
</tbody>
</table>
2.2 Cost overruns in Swedish transport projects

2.2.1 Descriptive statistics

As mentioned, data on cost estimations and outcomes of road and rail projects in Sweden completed during 1997-2009 are analysed in this thesis. There was a regulation adopted in 1997 that made it compulsory to compare between forecast costs and actual costs of large transport project investments. Consequently, the data have been published in yearly reports in Vägverket and Banverket. After our data collection, it was found that a rail project completed in 1998, Vallstanäs-Rosendal line had extreme cost overrun. The estimated cost was 8 million SEK and the actual cost was 40 million SEK. Thus, the cost overrun was 400%. It has been regarded as an outlier. After refining the data, 102 road projects and 65 rail projects are used in this thesis (as shown in Appendix A).

Figure 2-2 shows a histogram of the distribution of inaccuracies of cost estimates. The distributions of inaccuracies of cost estimations divided on road and rail projects are shown in Figure 2-3. Table 2-2 shows the general descriptive statistics of the data.

![Figure 2-2 Inaccuracy of cost estimates in 167 road and rail projects.](image)
Figure 2-3 Inaccuracy of cost estimates in road projects (a) and rail projects (b).
Table 2-2 Inaccuracy of transport project cost estimates by types of projects.

<table>
<thead>
<tr>
<th>Project type</th>
<th>Number of Cases (N)</th>
<th>Cost escalation (%)</th>
<th>Standard deviation</th>
<th>Level of significance, p*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Minimum</td>
<td>Maximum</td>
<td>Average</td>
</tr>
<tr>
<td>Road</td>
<td>102</td>
<td>-46.6</td>
<td>134.4</td>
<td>11.1</td>
</tr>
<tr>
<td>Rail</td>
<td>65</td>
<td>-54.2</td>
<td>250.0</td>
<td>21.1</td>
</tr>
<tr>
<td>All projects</td>
<td>167</td>
<td>-54.2</td>
<td>250.0</td>
<td>15.0</td>
</tr>
</tbody>
</table>

* Level of significance from zero (see details in Appendix B).
** For all project types, average cost overrun is different from zero with very high significance.

From the statistics results above, there are cost overruns in both road and rail projects in average. Like in the literature results, the average cost overrun in road projects is lower than in rail projects. In road projects, the average cost overrun is 11.1% with very strong significant difference from zero ($p < 0.001$). The cost overruns in road projects range between -46.6% and 134.4%.

In rail projects, the average cost overrun is 21.1% with very strong significant difference from zero ($p < 0.001$). The cost overruns in rail projects are between -54.2% and 250.0%. It should be noted that the standard deviation of cost overruns in rail projects is very high (50.5%). As show in Figure 2-3, the distribution of inaccuracy of cost estimates in rail projects is widely spread from mean. For road projects, the standard deviation of cost overruns is not as high as in rail projects and the data mostly lies in the range of 0% to 25%.

Next, the comparisons of cost overruns in transport projects in Sweden and of the previous studies are conducted. There are two comparisons. The first one is the comparison with results from the most cited studies by Flyvbjerg and other [5] [6] [22] [28]. The studies cover transport megaprojects and the project characteristics are thus different from the projects in this thesis. The second comparison is with the other international studies that were published after the most cited studies. These studies analyse transport projects that are more similar to this thesis projects in terms of size of project.

Table 2-3 shows cost overrun comparison of cost overrun between Swedish transport projects and results of Flyvbjerg studies. Average cost overruns in Swedish road and rail projects are lower than the average cost overruns from Flyvbjerg studies with very strong significance ($p < 0.001$). Thus, the cost overrun problems in Sweden seem to be less serious than other countries in magnitude. The standard deviation of cost overruns in road projects is not much
different than the result of Flyvbjerg studies. The standard deviation of cost overruns in rail projects is however much higher than the result of the Flyvbjerg studies.

As shown in Table 2-4, the average cost overruns in the other studies are 8.1% and 45.7% for road and rail projects, respectively. In road projects, the average cost overrun is higher in Sweden than the average cost overrun in other studies. In rail projects, the average cost overrun is lower than the average cost overrun in other studies. Thus, the cost overrun problem of Swedish road projects seems to be slightly more severe than in other countries. For the rail projects, the conclusion is the same as when we compare with the Flyvbjerg studies.

**Table 2-3** Comparison of cost overruns between Swedish transport projects and results of Flyvbjerg studies.

<table>
<thead>
<tr>
<th>Project type</th>
<th>Number of Cases (N)</th>
<th>Average cost overrun</th>
<th>Standard deviation</th>
<th>Level of significance, p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Road</td>
<td>167 / (102)</td>
<td>20.4 / (11.1)</td>
<td>29.9 / (24.6)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Rail</td>
<td>58 / (65)</td>
<td>33.8 / (21.1)</td>
<td>38.4 / (50.5)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

* The test of the difference between means of the two studies (see details in Appendix B).

** ( ) Results of Swedish transport projects.

**Table 2-4** Comparison of cost overruns between Swedish transport projects and result of other studies.

<table>
<thead>
<tr>
<th>Project type</th>
<th>Number of Cases (N)</th>
<th>Average cost overrun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Road</td>
<td>3,988 / (102)</td>
<td>8.1 / (11.1)</td>
</tr>
<tr>
<td>Rail</td>
<td>300 / (65)</td>
<td>45.7 / (21.1)</td>
</tr>
</tbody>
</table>

* Other studies are the studies that were published after Flyvbjerg studies.

** ( ) Results of Swedish transport projects.

### 2.2.2 Differences in cost calculations

The comparison of observed costs with the costs estimated at the time of project approval gives results that are similar to international studies’ results but deserve some comments [39]. There are two main difficulties in cost overrun comparisons – price index and the decision used as the estimated cost\(^1\). These factors affect the project costs.

\(^1\) Another problem that is obvious from an ongoing project at VTI is that there are shortcomings in the follow-up of actual costs at the Transport Administration in Sweden since the costs are not always registered correct, or even at the correct project. It is however outside the scope of this thesis to discuss the reasons behind this and its possible consequences.
When cost overruns are studied, all costs (actual cost and forecasted cost) should be calculated in constant prices or inflation-adjusted currencies. Constant prices are obtained by deflating current prices to a certain base year. Thus, the inflation is neutralised. There are several possible indexes for such deflation for example Consumer Price Index (CPI). Therefore, the price index affects cost overrun. Figure 2-4 is an example of cost calculations of rail projects in the U.S. In this example, nominal outcome costs are as reported and are not adjusted. Forecasted costs are adjusted by price index. If the price index is changed and make the estimated cost increase, some projects may change from cost underrun projects to cost overrun projects. Conversely, if the price index is changed and make the estimated cost decrease, some projects may change from cost overrun projects to cost underrun projects.

Another difficulty is what decision is used as the estimated cost. As mentioned, the definition of the cost overrun or inaccuracy in cost forecasts is measured as actual cost minus forecasted cost as a percentage of forecasted cost. Forecasted cost is the estimate made at the time of decision to build, but when exactly is this time of decision? Bent Flyvbjerg et al. discussed this as follows:

"Ideally, we would calculate cost development on the basis of the cost estimate at the time of the real decision to build. However, in most cases, it is virtually impossibly to identify the specific, real decision date."

A study investigated 10 academic studies and 13 audit reports about cost overruns. It showed that the actual construction costs were not adjusted for inflation in 2 academic studies. Cost overruns were calculated as the difference between the price at the time of the contract award and the final construction cost in 2 academic studies. For audit reports, 5 of them explicitly accounted for inflation, while the other did not state whether their findings are inflation-adjusted. More than a third of the audit reports focused on escalating payments above winning bid or the contractually agreed upon price.

These difficulties were also found in cost overrun calculation for Swedish transport projects. Therefore, the cost overruns calculated by the audit office were different from what was calculated by the Transport Authorities. The first explanation is that the transport agencies use different price indexes which are shown in Figure 2-5. The road and rail authorities use their own price indexes while the audit office uses Net Price Index (NPI) which is similar to CPI (KPI in Swedish) as shown in Figure 2-5. The second explanation is the decision that is used as the estimated cost. The Transport Authorities choose the estimated cost in the latest plan. They
agree that the latest action plan is the plan that the Government later decided for the current plan period, hence it is relevant to follow up analysis. However, the audit office claims that if choosing the latest plan, cost comparison hides the cost increases that occur between the first and the last action plan. Thus, they use the estimated cost at the earliest plan.

Table 2-5 shows the costs of major rail projects in Sweden (opened between 2005 and 2009). Cost calculations differ by using different indexes and plan. For example, the estimated cost of the Öxnered-Trollhättan line was 682 million SEK in earliest plan cost by using NPI index. At the same time, the estimated cost was 958 million SEK in the latest plan by using BV-index or Banverket index. In this case, the estimated project’s cost by the Swedish National Audit Office is 276 million SEK lower than by the Banverket. The calculation of road projects uses the same concept as the rail projects. This is the reason why the cost overruns calculated by the Swedish National Audit Office are higher than the ones calculated by the Transport Authorities (Vägverket and Banverket).
Figure 2- 4 Capital cost estimates – base estimate, inflation, and actual.
Figure 2-5 Price indexes for Swedish road projects (a) and rail projects (b).
Table 2-5 Cost of major rail projects (2005-2009).

<table>
<thead>
<tr>
<th>Project / year opened to service</th>
<th>Cost (million SEK)</th>
<th>Percentage deviation, the final cost against the first plan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kävlinge–Lund 2005 BV-index 800</td>
<td>1263</td>
<td>963</td>
</tr>
<tr>
<td>Kävlinge–Lund 2005 NPI 537</td>
<td>941</td>
<td>902</td>
</tr>
<tr>
<td>Lekarekulle–Frillesás 2005 BV-index 896</td>
<td>458</td>
<td>680</td>
</tr>
<tr>
<td>Lekarekulle–Frillesás 2005 NPI 602</td>
<td>356</td>
<td>637</td>
</tr>
<tr>
<td>Sthlm S–Arstaberg 2005 BV-index 2479</td>
<td>1413</td>
<td>1610</td>
</tr>
<tr>
<td>Sthlm S–Arstaberg 2005 NPI 1665</td>
<td>1100</td>
<td>1508</td>
</tr>
<tr>
<td>Hallsberg rangerbangård 2005 BV-index -</td>
<td>-</td>
<td>467</td>
</tr>
<tr>
<td>Hallsberg rangerbangård 2005 NPI -</td>
<td>-</td>
<td>437</td>
</tr>
<tr>
<td>Oxnered–Trollhättan 2006 BV-index 1087</td>
<td>693</td>
<td>958</td>
</tr>
<tr>
<td>Oxnered–Trollhättan 2006 NPI 682</td>
<td>504</td>
<td>838</td>
</tr>
<tr>
<td>Mjölby bangårds-ombyggnad 2006 BV-index 103</td>
<td>295</td>
<td>218</td>
</tr>
<tr>
<td>Mjölby bangårds-ombyggnad 2006 NPI 64</td>
<td>215</td>
<td>191</td>
</tr>
<tr>
<td>Blekinge kustbanan 2007 BV-index -</td>
<td>-</td>
<td>894</td>
</tr>
<tr>
<td>Blekinge kustbanan 2007 NPI -</td>
<td>-</td>
<td>726</td>
</tr>
<tr>
<td>Torebo–Hedberg 2008 BV-index -</td>
<td>1259</td>
<td>1621</td>
</tr>
<tr>
<td>Torebo–Hedberg 2008 NPI -</td>
<td>825</td>
<td>1278</td>
</tr>
<tr>
<td>Vännäs–Bastuträsk 2008 BV-index -</td>
<td>-</td>
<td>1091</td>
</tr>
<tr>
<td>Vännäs–Bastuträsk 2008 NPI -</td>
<td>-</td>
<td>860</td>
</tr>
<tr>
<td>Arlöv–Malmö 2008 BV-index -</td>
<td>-</td>
<td>857</td>
</tr>
<tr>
<td>Arlöv–Malmö 2008 NPI -</td>
<td>-</td>
<td>676</td>
</tr>
</tbody>
</table>

2.3 Key findings and discussion

- The problems of cost overrun appear worldwide. The average cost overrun in rail projects is higher than in road projects.

- Cost overruns among the studies are not directly comparable. Although concept and definition of the cost overrun are commonly understood, a variety of assumptions are used in cost overrun studies especially in audit reports.

- Comparing cost estimations and outcomes of 167 road and rail projects in Sweden during the period 1997-2009, it is found that average cost overruns are 11% (SD = 24.6%) and 21% (SD = 50.5%) for road and rail projects, respectively.

- The average cost overrun in Swedish road projects (11.1%) is much lower than the result from the most cited international studies (29.9%). However, it is slightly higher than the result from the average of other studies (8.1%). Therefore, the cost overrun in road projects in Sweden is not obviously better or worse than in other countries.

- For Swedish rail projects, the average cost overrun (21.1%) is lower than the results from both the most cited studies (33.8%) and the average of the other studies (45.7%). Thus, cost overrun problems in rail projects in Sweden are less severe than in other countries.
However, the standard deviation of cost overrun in Swedish rail projects is very high (50.5%).

- Two main aspects that should be considered when comparing between cost overrun studies are price index and the decision that is used as the estimated cost. These factors make a significant difference in project costs. One recommendation is that cost calculations with and without adjusted price indexes should be provided by the transport administration.
3. Factors contributing to cost overruns

Before we begin to study the factors contributing to cost overrun, there are two words that we need to clarify – *causes* and *explanations*. From the definition in the study of Bent Flyvbjerg and others [40], the meaning by ‘cause’ is ‘to result in’; the cause is not the explanation of the result. Causes refer to the variables or factors that influence the cost overruns, such as the implementation period or the size of the project [41]. Explanations are more general and might comprise several causes [41]. There are four categories of cost overrun explanations – technical, economic, psychological and political explanations. However, this thesis only concentrates on causes (or factors) of cost overruns, not explanations.

The first section of this chapter reports the literature study about the factors contributing to cost overruns. Next, it gives a brief explanation of statistic methods which are used to analyse the VTI data. The results are shown in the last section. In the next chapter, the currently used method in Sweden for cost estimations – “Successive Calculation” - will be explained.

3.1 Previous studies

The summary of some previous studies is shown in Table 3-1. The followings are the brief results of each study.

**Urban Mass Transportation Administration, 1990** [24]

The study covered 10 major transit capital investment projects in the U.S. cities. The transit projects were constructed during the period 1971-1987. It was stated that the changing in project scope was not the factor contributing to cost overruns. Changes in the physical characteristics of these projects between their planning and construction stages were generally quite minor, and many of the changes that were made should have reduced than increased project costs. Most of projects with the unanticipated inflations had cost underruns. Since actual inflation proceeded more slowly than it was forecast, actual cost would have been lower than those forecast if construction of the project had processed according to its planned schedule. Both delays in the start of construction and lengthening of planned construction schedules were the factors contributing to cost overruns.

**Office of Program Policy Analysis and Government Accountability, 1996** [27]

The Office of Program Policy Analysis and Government Accountability reviewed the Florida Department of Transportation’s (FDOT) performance in controlling cost overruns and delays when building roads and bridges. To identify the reasons why FDOT had experienced cost overruns in
transport projects, they examined a sample of 132 projects that had experienced cost overrun problems (1993-1995). These projects had cost increases at 17%. They found that there were complex and interrelated reasons why FDOT experienced cost increases in transport projects. The main factors accounting for cost overruns were errors and omissions in design plan, inadequate coordination with local government and utility companies, problems in identifying the scope of work that needed to be done during project development, changes in project specifications after design plans had been completed and damages to construction site due to extreme weather.


The sample of this study was used covering 258 transport projects in 20 countries completed in 1927-1998. They discussed whether different types of projects (rail, fixed link and road) performed differently with respect to cost overruns. They set a null hypothesis that type of project had no effect on cost overrun. An F-test falsified the null hypothesis at a very high level of statistical significance. Thus, project type mattered.

They also tested whether cost overrun varied with geographical location among Europe, North America, and “other geographical areas” (a group of 10 developing nations plus Japan). Considering all project types, they found that the difference between geographical areas in terms of cost overrun is highly significant. Thus, geography mattered to cost overruns. However, they found that when considered in each project type, the difference between the geographical areas is nonsignificant or does not influence cost overruns. They explained it might be because of the limited number of observations. Next, they analysed an effect from time (year of decision and year of completion) on cost overrun. They concluded that cost overrun had not decreased over time. Cost overrun was constant for the 70-year period and cost estimates had not improved over time.

Flyvbjerg, Holm, and Buhl, 2004 [40]

From database in the previous studies by Flyvbjerg and others (a sample of 258 rail, bridge, tunnel and road projects in 20 countries completed in 1927-1998), they furthered their analysis about what caused cost overruns. This study focused on the dependence of cost overrun on the length of the project-implementation phase, the size of the project and the type of project ownership. They found that the cost escalation was strongly dependent on the length of the implementation phase.
Next factor was the size of the project. They found that only bridges and tunnels were statistically significant that larger projects had larger cost overruns, whereas this did not appear in road and rail projects. For all project types, bigger projects did not have a larger risk of cost escalation than did smaller ones.

Lastly, by comparing the cost overruns for three types of project - ownership-private, state-owned enterprise and other public ownership, they concluded that the data did not support the general claim that public ownership is problematic *per se* and private ownership a main source of efficiency in curbing cost escalation \[^{40}\]. They stated that the general claim was an oversimplification and the issue of project ownership was more complex than was usually assumed.

**Odeck, 2004** \[^{12}\]

The paper investigated the cost overruns of road construction using data from 620 Norwegian road constructions over the years 1992-1995. A regression model was conducted to explain the cost overruns. The coefficient of the factors used in the model showed the cost overrun causes. From regression model results, the factors that contributed to cost overruns were the size of estimated cost (or project size), completion time, and location. It was found that cost overruns were larger and more predominant among smaller projects. It was suggested that one possible reason was the amount of attention given to larger projects. Thus, larger projects were most probably better managed as compared to smaller ones.

For the correlation between completion time and cost overrun, the results showed that cost overrun increases within certain range of completion (about 200 weeks or 2.7 years) after which it declined. It might imply that cost overrun tended to be higher the shorter the completion time was expected to take. The study stated that a possible explanation was that accuracy in cost prediction was difficult the shorter the construction was expected to take, meaning that uncertainties decreased with time. Another explanation could be that project management got better oversight over causes of overrun and were able to control for them the longer the project lasted.

About location, only the region of middle Norway was found to be different from the base region of eastern Norway. Thus, locations of projects were correlated with cost overruns. The study stated shortly about a correlation between the year of project and cost overrun. It specified that considering the dummies for vintage year of project, only 1993 were found to differ from the base year of 1992 (Since no details about this factor in the study, this factor will not be shown in Table 3-1). Lastly, the study showed that type of project (bridge, tunnel or ordinary road) had no correlation with cost overruns.
Dantata et al., 2006 [31]

This paper studied the effect of project size (in terms of actual cost) on cost overrun. Data was collected on 37 rail transit projects in the U.S. These included the 10 projects in UMTA report [24] in 1990, the 16 project completed after 1990 and 11 other projects. The relationship between the project size and cost overrun showed no clear pattern between the two variables. The correlation coefficient between the actual capital cost and the cost overrun is 0.37. They concluded that this was a weak correlation and therefore the relationship between the project size and cost overrun was weak. Moreover, they concluded that there was evidence to suggest that cost overruns for projects completed before 1990 were different from that of projects complete after 1994 (i.e. cost overruns had become smaller from 52% to 30%), but they did not have sufficient data to statistically prove this at a level of significance of 5%.

Roxas Jr., Chalermpong, 2008 [34]

The paper studied the cost forecasting inaccuracies in 44 road and bridge projects in Thailand and 85 road and bridge projects the Philippines. It was verified that there were no significant differences in cost forecasting inaccuracies between road and bridge projects in both countries. They tested that whether project size, location (urban and rural) and source of funding influenced on the amount of cost overruns.

The researchers analysed road and bridge projects in Thailand into two groups - costs less than one billion baht and costs of one billion baht and more. The result was the size of the project significantly influences the amount of cost overrun incurred. Interestingly, the average cost overrun and standard deviation of the second group were less than those of the transport projects with smaller costs. They concluded that it implied that bigger projects have smaller cost overruns. Regarding location, they concluded that there were no significant differences in the cost overruns in transport projects from both the urban and rural areas. They suggested that projects were planned, evaluated, and implemented consistently, despite the different area classifications. For the source of funding (foreign and locally funded projects), the researcher’s hypothesis was if a project was funded by a foreign institution, then it is expected that the cost overruns were less pronounced due to the fact the there were less incentives or no incentives at all, for the foreign agencies to manipulate the figures in the project proposals. The result showed the source of funding significantly influences the cost overruns. Moreover, the average cost inaccuracy of foreign projects is smaller than those of locally funded projects, thus supported the researcher’s hypothesis.
For the road and rail projects in the Philippines, all the factors – project size (project cost less than 1 billion pesos or more), location (urban or rural) and source of funding had significant effects on the amount of cost overruns. Bigger or more expensive projects experienced lower cost overruns. One explanation is bigger projects often had international consultants and longer planning and evaluations. Projects from the urban areas has less cost overruns probably because of better monitoring, quality of contractors and many other factors. Lastly, they stated that projects funded by JBIC/PJHL had less cost overruns when compared to other funding agencies (The authors did not provide the meaning of “JBIC/PJHL” and “other funding agencies”, thus it is not clear about this conclusion whether it means the foreign and locally funded projects.).

Singh, 2009[37]

This study investigated the delays and cost overruns in publically funded infrastructure projects in India. The 894 projects from 17 infrastructure sectors in the period 1992 and 2009 were analysed including 122 railway projects and 157 road and highway projects. It was found that first, since 1980s the cost overruns had declined. Cost overruns had systematically declined not only in absolute terms but also as a percentage of project cost. Second, delays were one of the crucial correlations behind the cost overruns. Third, bigger projects had experienced much higher cost overruns compared to smaller ones. Fourth, percentage cost overruns also escalate with length of the implementation phase. Fifth, compared to other sectors, projects from road, railways, urban development sectors had significantly higher cost overruns. Sixth, compared to other states, projects located in southern states had experienced lower cost overruns. Performance of rich states was not significantly better than that of the poorer states.

All above findings were statistically significant and robust to regression techniques. However, the results were analysed from the infrastructure projects in India. They are not only transport projects but include other types of project such as petroleum, telecommunication, mine and coal. Thus, it should be noted that the study did not focus only on transport projects like this thesis.
Summary

From the summary of factors contributing to cost overruns in Table 3-1, it shows that project type, location, length of implementation and size of project are the common factors used to explain cost overruns in transport projects. Some studies find that project type matters while two studies find that the project type does not matter. Studies of Flyvbjerg and others \[5\] \[6\] \[22\] \[28\] state that cost overruns in road, fixed link (bridge and tunnel) and rail projects perform differently with respect to cost overruns. Another study \[34\] states that there are no significant differences in cost overruns between road and bridge projects in Thailand and the Philippines. Another study \[12\] also states that type of project (bridge, tunnel or ordinary road) does not influence to cost overruns in Norwegian road projects.

Furthermore, all studies find that location of projects affected cost overruns. However, one study \[34\] shows that location affect cost overruns in road and bridge projects in the Philippines but not in Thailand. All studies also indicate that length of implementation and delay are factors contributing to cost overruns. The cost overruns increase with the length of implementation and delay.

There are varying conclusions in the size of project aspect. One study \[31\] shows that the relationship between the project size and cost overrun is weak. One study \[37\] finds that bigger projects have experienced higher cost overruns compared to smaller ones. On the other hand, two studies \[12\] \[34\] state that more expensive projects experience lower cost overruns. One study \[40\] concludes in two directions – larger bridge and tunnel projects have larger cost overruns, whereas this appears not to be the case for ordinary road and rail projects.

One interesting factor is historical period or the time of the project. This factor shows whether cost overruns improve over time. Some studies \[5\] \[6\] \[22\] \[28\] \[31\] find that cost overruns have not statistically significant decreased over time while one study \[37\] states that since the 1980s the cost overruns of transport projects in India have declined.
Table 3-1 Summary of factors contributing to cost overruns.

<table>
<thead>
<tr>
<th>Authors/Year of publish</th>
<th>Number and types of project examined</th>
<th>Project type</th>
<th>Location</th>
<th>Historical period</th>
<th>Length of implementation/ Delay</th>
<th>Size of projects (cost)</th>
<th>Project ownership</th>
<th>Changing in project scope, design, regulatory requirement</th>
<th>Unanticipated Inflation</th>
<th>Poor coordination</th>
<th>Weather</th>
<th>Source of funding</th>
</tr>
</thead>
<tbody>
<tr>
<td>UMTA/1990</td>
<td>10 major transit projects in the U.S.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[✓]</td>
<td>[×]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Office of program Policy Analysis and Government Accountability/1996</td>
<td>132 road projects in Florida by FDOT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[+]</td>
<td>[+]</td>
<td>[+]</td>
<td>[+]</td>
<td></td>
</tr>
<tr>
<td>Flyvbjerg, Holm, and Buhl/2002, 2003 and Priemus, Flyvbjerg and Wee/2008</td>
<td>258 projects in 20 nations: 58 rail; 33 fixed-link; 157 road</td>
<td>[✓]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[✓]</td>
<td>[×]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flyvbjerg, Holm, and Buhl/2004</td>
<td>Same data as previous study</td>
<td>[+]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[+]</td>
<td>[✓]</td>
<td></td>
<td>[&lt;num&gt;2]</td>
<td></td>
</tr>
<tr>
<td>Odekk/2004</td>
<td>620 Norwegian road projects</td>
<td>[×]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[✓]</td>
<td>[num&gt;3]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dantata et al./2006</td>
<td>37 rail projects in the U.S.</td>
<td>[✓]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[✓]</td>
<td>[×]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roxas Jr., Chalermpon/2008</td>
<td>129 road and bridge projects 85 in the Philippines 44 in Thailand</td>
<td>[×]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[✓]</td>
<td>[×]</td>
<td></td>
<td>[num&gt;3]</td>
<td></td>
</tr>
<tr>
<td>Singh/2009</td>
<td>894 projects from 17 infrastructure sectors in India; 157 road and highway 122 railway</td>
<td>[✓]</td>
<td></td>
<td></td>
<td></td>
<td>[num&gt;3]</td>
<td></td>
<td>[✓]</td>
<td>[×]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 [✓] means that the factor affects cost overruns.
[×] means that the factor does not affect cost overruns.
[+] means that a higher factor value contributes to more cost overruns.
[-] means that a higher factor value contributes to less cost overruns.
[±] means that the more factor value contributes to more and less cost overruns.
[?] means that the factor is needed to further study whether it affects cost overruns.

2 Only fixed-link (bridges and tunnels) larger projects statistically had lager cost overruns.

3 It was found that the relationship between the project size and cost overrun was weak.
3.2 Statistical methods

As mentioned, the VTI data are analysed to find the factors affecting cost overruns for transport projects in Sweden (the results are shown in next section). This section explains the general statistical methods which are used to analyse the data. More details in each method are provided in Appendix B.

For statistical analyses, we use the SPSS or PASW software. Most of the analyses are general descriptive statistics (means and standard deviations). The data are explored by graphs such as bar charts and scatterplots. Regression models are also used. Analyses in each factor are explained in the following.

The first factor is project type (road and rail projects). The Levene’s test is used for this part. It tests the hypothesis that the variances in different groups are equal. It basically does a one-way ANOVA on the deviations (i.e. the absolute value of the difference between each score and the mean of its group). Therefore, if the Levene’s test is significant, we can be confident in the hypothesis that the variance are significantly different and that the assumption of homogeneity of variances has been violated. In our analysis, it means that project type matters.

Regarding significance, we use the convention of terms: very strong significance ($p < 0.001$), strong significance ($0.001 \leq p < 0.01$), significance ($0.01 \leq p < 0.05$), nearly significance ($0.05 \leq p < 0.1$) and non-significance ($0.1 \leq p$).

In the next step, we test whether cost estimates have improved over time. The scatterplot of the differences between actual and estimated costs against year of project completion is applied. The null hypothesis is that year (or time of project) has no effect on the difference between actual and estimated costs. The compare means in one-way ANOVA function is used.

Next, the size of the project measured in costs is considered. In each project, there are two costs – forecasted cost and actual cost. From study of Bent Flyvbjerg and others, it is recommended that forecasted costs should be used. There are two reasons. First, cost escalation is statistically confounded with actual construction costs being part of it, whereas forecast construction costs are not. Second, the decision about whether to go ahead with a given project is based on the forecasted construction costs. We scatterplot the percentage cost escalation against project size. We also make regression models to see the relationship or correlation between project type and cost overruns. It should be noted that correlation does not imply causation. Even if we can see that there is a correlation or relationship between size of the project and cost overrun, it does not imply that size of the project causes cost overrun.
After this, we investigate the relationship between project size and cost escalation. The distribution of the cost overrun against project size is used to divide the projects into different groups – small, medium, large and very large projects. We provide summary descriptive statistics of percentage and monetary amount of cost overruns in each group.

Then we analyse detailed project type. For road projects, they are categorized into major roads (including interchanges), motorways, secondary roads and urban roads. For rail projects, they are classified into new tracks in existing lines, new tracks in new lines, stations and rail yards and upgrading existing lines. Descriptive statistics of percentage and monetary amount of cost overruns in each group are provided. The Levene’s test is used to test whether there are different variances in each detailed project type.

Lastly, for road projects, complexity of project affects cost overruns is analysed. The complexity of project is approximated by the ratio of construction cost and length of project. The scatterplot of the cost difference against complexity of project is conducted. Regression models are used to find the relationship between cost overrun and complexity.

3.3 Factors of cost overruns in Swedish transport projects

3.3.1 Types of projects

There are two types of projects (102 road projects and 65 rail projects). Whether different types of projects perform differently in cost overruns is tested. Statistical analysis in Table 3-2 shows the Levene’s test falsifies the null hypothesis at a very high significance ($p < 0.001$). The assumption that homogeneity of variances or type of project has no effect on cost overrun has been violated. Therefore, it is verified that the variance are significantly different and project type does matter. Thus, in the analyses that follow in this thesis, each type of project will be considered separately.

The result is in line with findings in studies of Flyvbjerg and others [5] [6] [22] [28]. They found that there were significant differences in the practice of cost overruns among different types of projects - rail, fixed-link and road projects.

We conclude that the Swedish road and rail projects perform differently in terms of cost overruns. The different cost overrun performances between road and rail projects seem to be a general characteristic in transport projects as found in previous chapter that the average cost overruns in rail projects are often higher than in road projects.
Table 3-2 Independent samples test by types of projects.

<table>
<thead>
<tr>
<th>Cost escalation</th>
<th>Levene’s Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
<th>95% Confidence Interval of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
<td>t</td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>20.298</td>
<td>.000</td>
<td>-1.713</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>-1.492</td>
<td>83.601</td>
<td>.140</td>
</tr>
</tbody>
</table>

3.3.2 Years of project completion

In this section, we test whether the cost estimations have improved over time. We start by plotting a scatterplot between percentage of cost inaccuracies against the years of project completion for the 167 road and rail projects as shown in Figure 3-1. The diagram does not seem to indicate an effect from year of project completion on cost overrun.

Next, we conduct the scatterplot in different types of projects (road and rail) as shown in Figure 3-2. The interpolation lines in the diagram show that cost overruns have declined since 2005, especially in road projects. The trend is tested by a statistical analysis. If the cost estimations have improved over time, the average cost overruns in each year will be different. Thus, the null hypothesis is average cost overruns in each year are the same. The null hypothesis is analysed by the comparing means in one-way ANOVA function. The results are shown in Table 3-3. It is statistically significant ($p < 0.05$) for both road and rail projects to reject the null hypothesis. Thus, it can be concluded that the year of project completion has effect on cost overruns. Therefore, the statistical analysis supports that cost overruns have declined since 2005 especially in road projects.

However, it was stated in two audit reports [14] [18] by the Riksrevisionen that the Road and Rail Administrations have changed to use their own self-constructed price indexes since 2005. Thus, this is the reason why it is obvious that cost overruns had declined since 2005. But it is still not clear whether the cost estimations really have improved over time. To clarify about this, the further analyses are done separately into two groups – before and after 2005. The results are shown in Table 3-4. From the results of F-test of cost overruns after 2005, the null hypothesis cannot be rejected, for both road and rail projects. Thus, year of project completion has no clear effect on cost overruns.

Factors contributing to cost overruns
Regarding the results of F-test of cost overruns before 2005, it is statistically significant ($p < 0.05$) for both road and rail projects to reject the null hypothesis. It can be concluded that year of project completion has effect on cost overruns. However, as shown in Figure 3-3, it cannot explain the trend of cost overruns during 1997-2004, for both road and rail projects. For road projects, during 1999-2002, cost overruns increased, then decreased in 2003 and increased again in 2004. Moreover, average cost overrun was very high in 1998 because there were two outliers (cost overruns were 134% and 110%). These two projects were secondary roads which was a group where average cost overrun was high. For rail projects, the cost overruns were very high in 1999 and 2002 and were almost zero in the rest of years. There were outliers in each year (cost overruns were 250% in 1999 and 146% in 2002). The outlier projects were stations and rail yards which was a group where average cost overrun was high.

Figure 3-1 Inaccuracy of cost estimates in Swedish transport projects over time, 1997-2009.
Figure 3-2 Inaccuracy of cost estimates in road and rail projects over time, 1997-2009.

Table 3-3 F-statistic of average cost overruns in road and rail projects over time, 1997-2009.

<table>
<thead>
<tr>
<th>Road</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>19392.720</td>
<td>12</td>
<td>1616.060</td>
<td>3.44</td>
<td>.000</td>
</tr>
<tr>
<td>Within Groups</td>
<td>41764.741</td>
<td>89</td>
<td>469.267</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>61157.460</td>
<td>101</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rail</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>53762.565</td>
<td>12</td>
<td>4480.214</td>
<td>2.12</td>
<td>.031</td>
</tr>
<tr>
<td>Within Groups</td>
<td>109604.201</td>
<td>52</td>
<td>2107.773</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>163366.766</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Factors contributing to cost overruns
Table 3-4 F-statistic of average cost overruns in road and rail projects, before and after 2005.

<table>
<thead>
<tr>
<th>Project</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Road 05-09</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>209.284</td>
<td>4</td>
<td>52.321</td>
<td>.411</td>
<td>.799</td>
</tr>
<tr>
<td>Within Groups</td>
<td>3946.377</td>
<td>31</td>
<td>127.302</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>4155.661</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rail 05-09</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>9073.935</td>
<td>4</td>
<td>2268.484</td>
<td>1.095</td>
<td>.376</td>
</tr>
<tr>
<td>Within Groups</td>
<td>66277.638</td>
<td>32</td>
<td>2071.176</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>75351.573</td>
<td>36</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Road 97-04</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>10936.135</td>
<td>7</td>
<td>1562.305</td>
<td>2.396</td>
<td>.032</td>
</tr>
<tr>
<td>Within Groups</td>
<td>37818.364</td>
<td>58</td>
<td>652.041</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>48754.499</td>
<td>65</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rail 97-04</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>44673.884</td>
<td>7</td>
<td>6381.983</td>
<td>2.946</td>
<td>.027</td>
</tr>
<tr>
<td>Within Groups</td>
<td>43326.563</td>
<td>20</td>
<td>2166.328</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>88000.446</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Factors contributing to cost overruns

In sum, we conclude that cost overruns in Swedish transport project are more or less constant for the 13-year period and that cost estimates have not improved over time. It should be noted that this conclusion holds true regardless of the change of using price indexes by Road and Rail Administrations. The conclusion corresponds to the studies of Flyvbjerg and others [5] [6] [22] [28] but does not correspond to a study of transport projects in India which stated that cost overruns had declined. The conclusion also supports the statement in the recent two audit reports [14] [18] that the cost overruns problem in Swedish transport projects as noted by the RRV 15 years ago still remained and it could not be seen that any real improvements had occurred in this aspect.

**Figure 3- 3** Inaccuracy of cost estimates in road and rail projects over time with SD, 1997-2009.
3.3.3 Sizes of projects

In this section, we test whether the size of project in terms of forecasted cost influences cost overrun. First, the scatterplot of the percentage cost inaccuracy against project size is provided in Figure 3-4. From the diagram, it is not clear to conclude a relationship between the forecasted cost and cost inaccuracy. From the regression models, the coefficients of forecasted cost in both road and rail projects are negative. Thus, it can be concluded that the higher forecasted cost have less cost overrun both in road and rail projects. However, it should be noted that the correlations between project size and cost overruns are very low in both road and rail projects (adjusted $R^2 = 0.178$ and 0.065 for road and rail, respectively). Thus, it also can be concluded that the relationship between the project size and cost overrun is weak in both road and rail projects.

Then, the road and rail projects are divided into four groups - small (cost $\leq 100$ million SEK), medium ($100$ million SEK $< \text{cost} \leq 500$ million SEK), large ($500$ million SEK $< \text{cost} \leq 1000$ million SEK) and very large (cost $> 1000$ million SEK). Figure 3-5 shows the distribution of cost overruns in each group. From the graph, it is obvious that small transport projects ($< 100$ million SEK) have much higher cost overruns than large projects, especially in road projects. It also shows that cost overruns in rail projects have very high standard deviation.

Next, the descriptive statistics are conducted by sizes of projects. The results are shown in Table 3-5. It shows that the average cost overruns are 29% and 43% in small road and rail projects, respectively while the average cost overruns are 11% and 21% in all road and rail projects, respectively. That means cost overruns in small projects in both road and rail projects are higher than the average of them by more than two times. They also constitute 27% and 23% of the total costs in road and rail projects, respectively. Moreover, it also shows that the average cost overruns are very low in bigger projects especially in rail projects (only 2.1%).
Factors contributing to cost overruns

Figure 3-4 Forecasted costs and inaccuracy of cost estimates.

Road Projects
\[ y = -10.22\ln(x) + 64.491 \]
Adjusted \( R^2 = 0.178 \)

Rail Projects
\[ y = -11.17\ln(x) + 80.244 \]
Adjusted \( R^2 = 0.065 \)

Figure 3-5 Distribution of inaccuracy of cost estimates by project sizes.
We conclude that the small transport projects in Sweden (< 100 million SEK) have much higher cost overruns than large projects especially in road projects. However, the correlation between project size and cost overrun is weak.

The result corresponds with a study of 620 Norwegian road projects[^12]. In the study, it was suggested that one possible explanation was the amount of attention given to larger projects. Thus, larger projects were most probably under better management compared with smaller ones. This explanation may also be the case for road and rail projects in Sweden. The result also corresponds with a study of 129 roads and bridges in Thailand and the Philippines. However, the result is not in line with finding for transport projects in India which stated that bigger projects had experienced much higher cost overruns compared to smaller ones.

Furthermore, the result is the same as in a study of 37 rail projects in the U.S. finding that the relationship between the project size and cost overrun is weak.

---

[^12]: It is the average cost overrun of road projects.
[^1]: It is the average cost overrun of rail projects.
3.3.4 Detailed project types

In Section 3.3.1, we investigated whether road and rail projects perform differently in terms of cost overruns. The detailed project type is further analysed in this section. There are four categories in road projects – major roads, motorways, secondary roads and urban roads. For rail projects, they are divided into four groups – new tracks in existing line, new tracks in new line, stations and rail yards and upgrading existing line.

The results are shown in Figure 3-6 and Table 3-6. For road projects, the average cost overruns are lowest in motorways (only 3.0%) and highest in secondary roads (20.2%). The standard deviation of cost overruns in secondary roads is high (37.7%). Furthermore, motorways constitute more than half (51.3%) of the total amount of cost overruns in million SEK.

For rail projects, new tracks in new lines are underrun (-8.8%). This means the forecasted cost is higher than the actual costs in average. One explanation is that when constructing new tracks in new line corridors it is usually easier to control costs than when constructing in existing line corridors. There are no conflicts with the existing lines during construction. The highest average cost overrun is found in stations and rail yards (34.7%) with very high standard deviation (68.7%). The large standard deviations in rail projects are as interesting as the large average cost overruns. The size of standard deviations demonstrates that uncertainty and risk regarding cost overruns are large, indeed. Furthermore, construction of new tracks in existing line corridors constitutes a lot of the total amount of cost overruns (75.5%) in million SEK.
Figure 3-6 Distribution of inaccuracy of cost estimates in road projects (a) and rail projects (b) by detailed project types.

Factors contributing to cost overruns
Table 3-6 Distribution of cost overruns in percent and in million SEK by detailed project types.

<table>
<thead>
<tr>
<th>Road</th>
<th>Project types</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Major roads</td>
<td>Motorways</td>
</tr>
<tr>
<td>Number of projects</td>
<td>24</td>
<td>32</td>
</tr>
<tr>
<td>% of projects</td>
<td>23.5</td>
<td>31.4</td>
</tr>
<tr>
<td>Average cost per project (in mill. SEK)</td>
<td>252</td>
<td>719</td>
</tr>
<tr>
<td>Average cost overrun (%)</td>
<td>13.6</td>
<td>3.0</td>
</tr>
<tr>
<td>Standard deviation (%)</td>
<td>18.1</td>
<td>11.6</td>
</tr>
<tr>
<td>Sum overrun (in mill. SEK)</td>
<td>424</td>
<td>843</td>
</tr>
<tr>
<td>% of sum overrun (in mill. SEK)</td>
<td>25.8</td>
<td>51.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rail</th>
<th>Project types</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>New tracks, existing line</td>
<td>New tracks, new line</td>
</tr>
<tr>
<td>Number of projects</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>% of projects</td>
<td>21.5</td>
<td>10.8</td>
</tr>
<tr>
<td>Average cost per project (in mill. SEK)</td>
<td>843</td>
<td>779</td>
</tr>
<tr>
<td>Average cost overrun (%)</td>
<td>23.4</td>
<td>-8.8</td>
</tr>
<tr>
<td>Standard deviation (%)</td>
<td>28.6</td>
<td>13.1</td>
</tr>
<tr>
<td>Sum overrun (in mill. SEK)</td>
<td>1,462</td>
<td>-499</td>
</tr>
<tr>
<td>% of sum overrun (in mill. SEK)</td>
<td>75.5</td>
<td>-25.8</td>
</tr>
</tbody>
</table>

Table 3-7 Test of homogeneity of variance of detailed project types.

<table>
<thead>
<tr>
<th>Test of Homogeneity of Variance</th>
<th>Levene Statistic</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Road</td>
<td>Based on Mean</td>
<td>7.962</td>
</tr>
<tr>
<td>Rail</td>
<td>Based on Mean</td>
<td>1.982</td>
</tr>
</tbody>
</table>

As seen in the previous results there are different means of cost overruns in each detailed project type. The cost overruns in each detailed project type may statistically differ from the others and we should not combine analysis of each detailed projects together. Thus, we do a further test whether there are different variances in each detailed project type of road and rail projects. The null hypothesis is that the variances in different groups are equal. The Levene’s test based on mean (more explanations as shown in Appendix B) is applied. The result is shown in Table 3-7. The Levene’s test result is non-significant for the rail projects (\(p > 0.1\)) indicating

\(^1\) It is the average cost overrun of road projects.
\(^2\) It is the average cost overrun of rail projects.
that they are similar in variances. However, for the road projects, the Levene’s test is significant \( (p < 0.001) \) indicating that they are different in variances in each detailed project type. Thus, the road projects should be analyzed separately in each detailed project type (major roads, motorways, secondary roads, and urban roads). However, there are some constrains to do analysis separately in road projects. First, it will not be a statistically valid analysis because of the small sample size in each detailed project size (around 16-32 observation in each group). Second, it is impossible to find the reference studies to compare. None of our literature studies classifies the road into sub-groups as in this thesis. Therefore, even though it is significant that road projects are different in variance in each detailed project type, the road projects are not analysed separately in each detailed project type in this thesis.

3.3.5 Complexity

In this section, it is tested whether the complex projects have higher cost overruns. The complexity of project is measured somewhat crudely by the ratio of construction cost and length of project. In this thesis, the analysis of complexity is focused only on road projects because the lengths of projects are provided only in road projects. First, the scatterplot of the percentage cost inaccuracy against complexity of project is provided in Figure 3-7. From the diagram, it is seen that the more complex projects have lower cost overruns. This result is confirmed by the regression model, as shown in Figure 3-7. The coefficient of complexity is negative. However, the relationship between cost overrun and complexity is very weak (adjusted \( R^2 = 0.038 \)).
Figure 3-7 Complexity and inaccuracy of cost estimates.

The result shows that the more complex projects have lower cost overruns with low adjusted $R^2$. The next question is whether this trend is different in each type of road. The scatterplot of the percentage cost inaccuracy against complexity of road project in each type are shown in Figure 3-8. It shows that the correlation between cost overrun and complexity is weak in major roads. Moreover, there are no relationship between cost overrun and complexity in motorways, secondary roads and urban roads. Therefore, our conclusion is that complexity does not affect cost overruns. Note that our measure to complexity is crude and that we are not sure it deflects complexity very well.
Factors contributing to cost overruns

Figure 3-8 Complexity and inaccuracy of cost estimates by detailed project types.
3.4 Key findings and discussion

- The literature study shows that researchers commonly test the following factors contributing to cost overruns – project type, location, length of project implementation and project size. Most studies show that the location of projects affects cost overruns. The effect of such projects characteristics are hard to generalize since they will differ between countries as well as between urban and rural areas. However, there is one factor that has the same pattern of effect to cost overruns in all studies - the length of project implementation and delay. Higher length of project implementation and delay always increase cost overruns.

- It is obvious that road and rail projects in Sweden perform differently in terms of cost overruns. Moreover, this result supports the conclusion from previous studies that the average cost overruns are higher in rail projects than in road projects. The Swedish rail projects also have very high standard deviation of cost overruns. It shows that uncertainty and risk regarding cost overruns are large.

- The cost overruns in road and rail projects in Sweden are constant for the 13-year period and cost estimates have not improved over time. Note that this finding is regardless of the change in use of price indexes by Road and Rail Administrations.

- We find that small Swedish transport projects (< 100 million SEK) have much higher cost overruns than large projects. Moreover, they constitute around one fourth of the total amount of budget used for cost overruns in both road and rail projects. We also find that the average cost overruns are very small in bigger projects especially in rail projects.

- Average cost overrun in motorways is very low but they still constitute more than half of the total amount of cost overruns in million SEK. For secondary roads, the average cost overrun is high but they do not constitute much of the total budget for cost overruns.

- Among the rail projects, stations and rail yards have high average cost overrun with very large standard deviation. Therefore, planners should pay more attention to these investments. Moreover, it should be realized that constructing new tracks in existing line corridors constitute a lot of the total amount of cost overruns in million SEK.

- For road and rail projects in Sweden, the more complex projects do not seem to have a larger risk of cost overrun than the less complex projects. This finding is however uncertain.
4. Methods to improve cost estimations - Successive Calculation

This chapter explains one of the methods that are used for more realistic assessments of uncertainty in cost estimations. It is the Successive Calculation. The key concept of this method is that risk and uncertainty in cost estimation is considered by group analysis. In principle, the group should include individuals who can contribute to creativity and prevent optimism (and perhaps also too much pessimism). However, in practice, the group mainly includes people who relate with the projects. According to some researches in this area, applying this method still gives risk of optimistic cost estimation. The consideration from the planners or people who relate with the projects is mostly optimistic. Thus, this method is still what is called an inside view analysis. In the next chapter, we focus on another method which is the Reference Class Forecasting. It is an outside view analysis. The method examines the experiences of a class of similar projects, lays out a rough distribution of outcomes for this reference class, and then positions the considered project in that distribution [2]. In some studies, it is claimed that the resulting forecast from an outside view analysis is more accurate than an inside view analysis.

4.1 Successive Calculation method

The development of the Successive Calculation was initiated at the Technical University of Denmark by Steen Lichtenberg in the beginning of the 1970s [43]. The principle is originally a tool for fast, early cost estimating and scheduling in the construction industry. Later it has developed into a multi-purpose management instrument. From the 1980s onwards, it has functioned as a risk management tool and has been applied professionally [43]. Two important features of the method are using the group analysis and procedure of working top down.

4.1.1 General procedure

In this section, the general procedure of the Successive Calculation is summarized [44]. The procedure is shown in Figure 4-1. There are two main phases; the qualitative and the quantitative phases. The procedure applies a “top-down” philosophy. The main items of the project are considered first, followed by sub-items of the main items. The “top-down” philosophy forces the analysis group to proceed and analyse project systematically.
Methods to improve cost estimations – Successive Calculation

Figure 4-1 Successive Calculation procedure.
The qualitative phase

- **Establishing an analysis group**
  A properly balanced analysis group of key individuals is established to suit the specific purpose of the analysis. The analysis group implements the process outline below. The first task of the analysis group is an in-depth discussion of the actual programme, project or task, its objectives, characteristics, scope and other fixed preconditions.

- **General sources of uncertainty**
  All general or overall sources of potential uncertainty are identified, organized in groups, and defined according to relevant sub-routines. Brainstorming is used in this step.

The quantitative phase

- **Establishing and evaluating the basic structure**
  A set of main items or main activities is chosen and described. Their numerical values are evaluated, using the technique of “group triple estimating”. According to the probability theory\(^1\), the mean values can be calculated by using this formula:

\[
\text{Mean value } (M) = (\text{min.} + 3 \times \text{most likely} + \text{max.})/5 \tag{Equation 4.1}
\]

The set of items includes correction items which reflect the effect from the general sources of uncertainty mentioned above. The correction items are any other factors both external and internal, such as competition, organization etc.

Next, the uncertainty calculation is conducted. The uncertainty of a given parameter is either measured as the so-called “range”, defined as the difference between the minimum and the maximum limits of the triple estimate, or as the standard deviation. The standard deviation is approximately equal the formula below;

\[
\text{Standard deviation } (S) = (\text{max.} - \text{min.})/5 \tag{Equation 4.2}
\]

The standard deviation (S) reflects the local uncertainties of each parameter. However, they do not affect the grand total figure with their full value. They contribute only via square of S \((S^2)\) to the total sum of squares. This value is called the variance in statistical analysis. Here, the \(S^2\) is called the Priority Figure (P) because it illustrates the relative importance of the item to the uncertainty of the total result. Lastly, the global standard deviation (S) for the grand total result needs to be calculated by deriving the square root of the sum of the \(S^2\).

\(^1\) The mean value and standard deviation calculation above is based on “Erlang distribution”. However, if it is stressed that other distribution functions may be more relevant, the transformation formulas for the mean value and standard deviation must be adjusted accordingly.
The first numerical calculations of the total result are performed, using the statistical rules above. They result (1) in a mean value and its related variability, and notably (2) in a “Priority Figure” (P) for every discrete element which indicates its relative importance to the reliability of the total result. An example of calculation details is shown in Figure 4-2.

![Table of calculation details](image)

**Figure 4-2 Example of calculation details.**

As discussed earlier, each main item is evaluated, specified during the qualitative phase of the analysis. In the first line, the greatest optimist in the analysis group has evaluated that the costs of item 1 will certainly be higher than £10m. Similarly, the greatest pessimist has evaluated that the costs will not exceed £40m. The average or most likely value is £20m. The mean values (M) are calculated according to the Equation 4.1. For example, the first line, M = (10 + 3x20 + 40)/5 = 22.0. The standard deviations (S) are calculated by the Equation 4.2. Continuing with the first line as an example, S = (40-10)/5 = 6. The mean values and standard deviations are calculated for all items as well as their Priority Figures (S^2). The
mean value of the grand total is the sum of all mean values (£123m). The global standard deviation (S) is approximately equal to the square root of the sum of Priority Figures (140). In the example it is approximately £12m or around 10% of the mean value.

- The systematic, successive detailing process

The most crucial items are detailed successively. The above Priority Figure is the efficient guide. This successive detailing and re-evaluating process continues until the analysis group has approached what they believe is an unavoidable minimum level of uncertainty. In this step, the Priority Figure is reduced by considering sub-divisions in each parameter. In the parameters with high Priority Figure values, the statistic values in their sub-divisions are calculated. Then, a previous figure is cancelled and is replaced with the new set of detail items. An example of this process is in Figure 4-3.

**Figure 4-3** Example of systematic and successive process.
Figure 4-3 shows how to reduce the total uncertainty by a systematic, successive process. The items no.1 and no.2 are calculated in sub-divisions. In this case, they are relevant to the item no.3. Thus, the item no.3 needs to be re-evaluated after the calculation. The two new sub-divisions (item no.1 and no.2) and a re-evaluation (item no.3) of the major main items have reduced the total uncertainty significantly. In this case, the standard deviation reduces from £12m to £10m.

- Establishing conclusions

The analysis group finally identifies and ranks a set of suggested action plans or other conclusions which is likely to optimize the venture, including relevant measures against any major commercial risks.

The top ten risk items are summarized. Moreover, knowing the mean value and standard deviation from the previous step, it can provide the statistical S curve or cumulative distribution function. An example of the curve is shown in Figure 4-4. The normal distribution is a reasonable approximation. Only in case of a dominant skewed local uncertainty will the resulting curve be slightly skewed. In such cases, it is possible to make a more exact calculation of the resulting curve.

The statistical S curve or cumulative distribution function of the total result is the probability that the actual future costs will not exceed a certain given budget value. It shows the probability inherent in a given budget. This S curve allows management to develop safe and workable budgets while a conventional budget estimate consists of only one definite figure. The two levels of budgets that are reported as a result of the Successive Calculation are explained as follows.

- A working budget usually includes, as a minimum, the expected total sum (assumed base case conditions). In Figure 4-4, it equals £123.9m. It may typically be defined as being equal to the expected costs for the overall influences. Project managers are typically responsible for handling this.

- A higher-level budget is the cost for the higher level of responsibility, therefore requires a further margin or reserve in the budget in order to support for such risk and uncertainty of the project. For example in Figure 4-4, at a 90% confidence limit, the higher-level budget is £135.6m.
4.1.2 Applications and limitations

The Successive Calculation has been applied professionally since 1980\textsuperscript{[43]}\textsuperscript{[44]}. It is considered a multi-purpose management tool. It supports such processes as quality assurance of budgets, bid or tender estimates, and schedules, profitability analyses and other financial analyses. It is also applied in risk analyses, ranking of alternative solutions and team building. However, the application to cost estimations is focused in this thesis. Two examples of practical experiences and results in cost estimation are shown in Table 4-1.

From Table 4-1, the results of both projects are excellent by using the Successive Calculation. Nevertheless, this method has some limitations;

- The necessity of a modern, open management policy.
- A supplement to, not a substitute for, existing planning techniques.
- Only the overall total result is reliable, not each sub-item or activity.
- Focus on quantitative objectives.
- Catastrophes or disasters require a supplementary sub-procedure.
- Trained facilitators are required.
- Subjective uncertainty must be accepted.
- The implementation process requires effort, time and the support of senior management.
- An inside view, meaning optimism bias is likely to remain.
Table 4-1 Two examples of applications and results in cost estimation by the Successive Calculation.

<table>
<thead>
<tr>
<th>Projects</th>
<th>Original budget</th>
<th>Budget by Successive Calculation</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex high-tech, multi-purpose 10,000-seat arena, Oslo Spectrum in Norway</td>
<td>$45M</td>
<td>$80M ± $10M</td>
<td>$125M The official project account after the construction deviated by less than 1% from the calculated mean value or working budget.</td>
</tr>
<tr>
<td>Lillehammer Olympic Games in Norway</td>
<td>$1230M</td>
<td>$800M ± $90M</td>
<td>- The final official accounts equaled the analysis mean value of $800M, so the reserve fund was saved and was used to operate the facilities after the games.</td>
</tr>
</tbody>
</table>

4.2 Use in Swedish investment planning

4.2.1 Successive Calculation in transport investment planning

In Sweden, Transport Investment Plans are conducted by the Swedish Government via public agencies. The plans cover a period of around ten years and are typically revised every five years \[^3\]. The latest plan is the plan of the period 2010-2021. In this plan, the Road Administration and the Rail Administration carried out the entire planning process jointly during 2008-2009. The plan comprises one part with national road and rail investments and one part with 21 regional plans, one for each county region.

The Successive Calculation method was used in the latest plan. In 2008, there was a regulation stating that all transport projects with a total cost greater than 500 Million SEK needed to do the Successive Calculation \[^14\][^18][^45].

The Successive Calculation that has been applied in the Swedish investment planning actually is similar to the general procedure of the method shown in Figure 4-1. However, it is applied in two different ways – the full uncertainty analysis and the simplified uncertainty analysis \[^45\]. The summary of the two analyses is as follows.
The full uncertainty analysis

The full uncertainty analysis is used for large and complex projects. It is carried out by a cross-compound analysis group (8-15 members) for 2 days under the direction of a trained moderator. The result of the cost calculation is provided in a report (“Kalkyl-PM”). Before the uncertainty analysis is implemented, the project has developed its own basic calculation. The result of the basic calculation is not presented to analysis members before the analysis is carried out. The uncertainty analysis is made in accordance with the requirements of Annex 1 (“Kalkylbok”) and Annex 2 (“Kalkyl-PM”). After the full uncertainty analysis, a comparison with the basic calculation is done. Then, it is the responsibility of the project manager to update the estimate and implement appropriate measures to eliminate the largest uncertainties.

The simplified uncertainty analysis

The simplified uncertainty analysis follows the same methodology as the full analysis. However, it is done with a smaller analysis group and in less time. The analysis group should be at least 5 people, including at least 2 people who are not actively included in the project team. Before the analysis, the project develops its own basic calculation. The uncertainty analysis is carried out step by step according to Annex 1 (“Kalkylbok”) and Annex 2 (“Kalkyl-PM”). This analysis is performed in-house (with the support of regional staff).

Example of the Successive Calculation result

An example of a Successive Calculation result is shown in Figure 4-5. It is the result of Stockholm bypass (Förbifart Stockholm). From the left, the first results are the general issues or general sources of uncertainties. It shows the base case situation and the future situation. The results come from brainstorming. The second result is the basic hierarchical Work Breakdown Structure (WBS) of the project. This WBS is used in the analysis of the next step which is the top-level estimate sheet. An estimate sheet is a calculation sheet that shows the statistics values such as mean, Priority Figure and standard deviation.

The fourth one is the uncertainty profile or top ten risks. The fifth and the sixth results are the Bell-curve and S-curve of the cost estimation, respectively. The curves show the probability inherent in a given budget. For Swedish transport projects, they use cost at an 85% confident limit as the higher-level budget. The last result is the report of the Successive Calculation.

It should be noted that the cost at an 84.13% confident limit equals the probability that the cost will be within the estimate plus one standard deviation from the mean as shown in Figure 4-6[46]. Therefore, it can be simplified that the higher-level budget at the 85% confident limit in Swedish transport projects is the value at the one standard deviation above the mean.
Figure 4-5 Example of the Successive Calculation result of the Stockholm bypass (Förbifart Stockholm).
4.2.2 Cost variance analyses

The Successive Calculation method has only been applied since 2008 in Swedish transport projects, so there are not yet any evaluations on the outcomes. Instead, in this thesis, we compare the results of the method with the standard deviations of the actual outcomes. The actual outcomes are the standard deviations of road and rail projects completed during 1997-2009 that are reported in Chapter 2. As explained earlier, the budget at 85% confident limit is located at the one standard deviation from mean. Thus, the variances can be compared with the standard deviation of the actual outcomes. We assume that the distribution of cost estimate inaccuracy reflects the distribution of project cost. To see the cost variance, we calculate cost at 85% confidence limit minus cost at 50% confidence limit as a percentage of cost at 50% confidence limit as in Equation 4.3.

\[
\text{%Difference of cost} = \left(\frac{\text{Cost at 85\%} - \text{Cost at 50\%}}{\text{Cost at 50\%}}\right) \times 100
\]

Equation 4.3
Methods to improve cost estimations

Successive Calculation

The variances are calculated both in road and rail projects. The data of the budget at 85% and 50% confident limit are collected from the projects in the latest national investment plan. For road projects, we received the data from the Transport Administration calculation sheets. For rail projects, we find the data from the Transport Administration website. The variances are compared with the historical actual outcomes as explained above. Moreover, the variances in different planning phases are investigated. The results show whether cost estimates are more certain the longer an investment has come in the physical planning process. The cost estimates ought to be more certain the longer an investment has come in the physical planning process as shown in Figure 4-7[47].

![Project cycle diagram]

**Figure 4-7** Cost uncertainty during the project cycle.
The phases of planning for road and rail projects in Sweden are shown in Figure 4-8. We categorize road projects, in “Bristanalys” or very early planning, “Förstudie” or initial study, “Vägutredning” or feasibility study and “Arbetsplan” or design plan. Rail projects are categorized similarly. We have considered how for the projects had advanced at the time when the cost calculation was carried out.

After refining the data, 249 road projects and 46 rail projects are used. We calculate the differences of cost estimates or variances according to Equation 4.3 for each project. Then we analyse the descriptive statistic in each phase of planning. The mean value of each phase reveals the thinking of transport planners to the projects in each phase. Higher value of variance means they think the project more uncertainties. The results in each phase are as follows.
Table 4-2 Descriptive statistics of road projects.

<table>
<thead>
<tr>
<th>Road</th>
<th>N</th>
<th>Cost differences (%) between 85% and 50% confident limit</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Minimum</td>
<td>Maximum</td>
</tr>
<tr>
<td>Bristanalys</td>
<td>42</td>
<td>3.33</td>
<td>13.33</td>
</tr>
<tr>
<td>Förstudie</td>
<td>70</td>
<td>1.10</td>
<td>28.20</td>
</tr>
<tr>
<td>Vägutredning</td>
<td>76</td>
<td>1.70</td>
<td>18.30</td>
</tr>
<tr>
<td>Arbetsplan</td>
<td>61</td>
<td>.90</td>
<td>13.20</td>
</tr>
<tr>
<td>Total</td>
<td>249</td>
<td>.90</td>
<td>28.20</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>42</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4-9 Comparison between road project variances and actual outcome.
Table 4-3 Descriptive statistics of rail projects.

<table>
<thead>
<tr>
<th>Rail</th>
<th>N</th>
<th>Cost differences (%) between 85% and 50% confident limit</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Minimum</td>
<td>Maximum</td>
</tr>
<tr>
<td>Idéstudie</td>
<td>6</td>
<td>8.93</td>
<td>21.27</td>
</tr>
<tr>
<td>Förstudie</td>
<td>18</td>
<td>3.73</td>
<td>26.87</td>
</tr>
<tr>
<td>Järnvägsutredning</td>
<td>8</td>
<td>6.05</td>
<td>20.08</td>
</tr>
<tr>
<td>Järnvägsplan</td>
<td>14</td>
<td>3.69</td>
<td>24.26</td>
</tr>
<tr>
<td>Total</td>
<td>46</td>
<td>3.69</td>
<td>26.87</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4-10 Comparison between rail project variances and actual outcome.
For road projects, the differences of variances between each phase are small as shown in Figure 4-9. Table 4-2 shows descriptive statistics of variances in each phase. It shows that the variances are similar in means and standard deviations. The exception being that projects in the latest phase – Arbetsplan - have a somewhat smaller variance. Still, it can be concluded that the cost estimates in road projects are not more certain the longer an investment has come in the physical planning process. Moreover, we find that the average variance is significantly lower than actual outcome. The variance mean is only 9.1% while the actual outcome is 24.6% as shown in Figure 4-9.

For rail projects, there are no significant differences of variances between each phase as shown in Figure 4-10. Table 4-3 shows descriptive statistics of variances in each phase. The means and standard deviations of variances in each phase are also not significantly different. However, the highest uncertainty is found in “Idéstudie” phase. The “Järnvägsutredning” phase has higher uncertainty than the “Förstudie” phase. It may be because the projects in “Järnvägsutredning” phase are much larger and complex objects. Uncomplicated projects do not need “Järnvägsutredning” (they can go directly from Förstudie to Järnvägsplan). Furthermore, the average variance is much lower than the actual outcome. The variance mean is only 12.1% while the actual outcome is 50.5%.

In sum, we find that the differences between actual outcomes and variances are surprisingly small between projects in different planning stage in both road and rail projects. Therefore, we can conclude that planners believe that investments in early planning stages feel equally certain as in late planning stages. Moreover, their cost estimations seem optimistic. The actual outcomes in terms of cost overruns are considerably higher than the variances of the projects in the latest plan.

As size of projects is one of the factors identified in international studies that contributes to cost overruns, further analyses that consider different size of projects are conducted. The actual outcomes are the standard deviations which are stated in Chapter 3. The differences of cost estimates according to Equation 4.3 for each project size are calculated. Table 4-4 shows the results. As can be seen, for road projects the variance decrease slightly with project size. For rail projects the opposite is true, i.e. variances increases with project size. Figure 4-11 and Figure 4-12 are the graphs of variances in different project sizes for road and rail projects, respectively. The variances of small projects in both road and rail projects are much lower than the actual outcomes especially in rail projects. The difference is very large, 74% compared to 8%. For the large projects, the differences between actual outcomes and variances are much
lower. The actual outcome is even lower than in the Successive Calculation for very large rail projects. Our results are similar to those in a cost overrun study of Norwegian road projects [12]. The researcher explained that cost overruns were more predominant among smaller projects because more attention was given to cost estimates of large projects.

Table 4- 4 Variances by project sizes.

<table>
<thead>
<tr>
<th>Road</th>
<th>Transforms of forecasts of the cost (million SEK)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Small (cost ≤ 100)</td>
<td>Medium (100&lt;cost≤500)</td>
</tr>
<tr>
<td>Number of projects</td>
<td>116</td>
<td>118</td>
</tr>
<tr>
<td>Variances in latest plan (%)</td>
<td>9.7</td>
<td>8.7</td>
</tr>
<tr>
<td>Standard deviation (%)</td>
<td>4.1</td>
<td>3.4</td>
</tr>
<tr>
<td>Actual outcomes (%)</td>
<td>34.0</td>
<td>14.3</td>
</tr>
<tr>
<td>Differences between actual outcomes and variances in latest plan (Percentage units)</td>
<td>24.3</td>
<td>5.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rail</th>
<th>Transforms of forecasts of the cost (million SEK)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Small (cost ≤ 100)</td>
<td>Medium (100&lt;cost≤500)</td>
</tr>
<tr>
<td>Number of projects</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>Variances in latest plan (%)</td>
<td>8.1</td>
<td>11.0</td>
</tr>
<tr>
<td>Standard deviation (%)</td>
<td>3.7</td>
<td>5.6</td>
</tr>
<tr>
<td>Actual outcomes (%)</td>
<td>74.1</td>
<td>33.7</td>
</tr>
<tr>
<td>Differences between actual outcomes and variances in latest plan (Percentage units)</td>
<td>66.0</td>
<td>22.7</td>
</tr>
</tbody>
</table>
Figure 4-11 Road project variances in different project sizes.

Figure 4-12 Rail project variances in different project sizes.
4.2.3 Discussions

As the Successive Calculation has recently been applied, there are different opinions about the use of the method from people who have been involved in the planning. Some opinions are described below.

A planner who used to work at the Rail Administration thinks that the use of the Successive Calculation it will actually dramatically increase the cost of new infrastructure. He gives two explanations. Firstly, when everyone try to make cost estimates that they are more or less sure are not too high this will result in larger project budgets and these budgets will then be spent even if it would be possible to find more cost-effective solutions (since the project leaders have no incentives to keep costs down). Secondly, because the building companies know of the cost estimates, they will give high bids for constructing the infrastructure (which will increase their profits).

A researcher’s response to above opinion: This is an issue for how the uplift should be used. If the uplift is constructed as a completely general multiplier like 1.4 for railway projects this could lead to incentives for project managers of railway projects to use up the corresponding resources. Such fears could however be countered by bonuses paid to project managers for delivering cost underruns.

A planner at the Trafikverket explains about the Successive Calculation that they do not increase the budget for the project during the Successive Calculation. They only reduce the variance by the successive approach, making them more secure on each main item. This continues until they reach a low enough variance of approximate 10% between 50% and 85% confident limit.

If the above explanation is true, budgets will not increase because of the use of the Successive Calculation and thus the fears of the previous cited planner will not come true. However, there is a risk that the resulting variance is smaller than it will actually turn out in reality.
4.3 Key findings and discussion

- Using the Successive Calculation in cost estimation has advantage in evaluating uncertainty. However, the method has its limits. Some measures need to be incorporated with the use of the Successive Calculation. One recommendation is that the Trafikverket (or the Government) should start a systematic follow-up of the Successive Calculations. Incentive for the project manager who can efficiently control project cost is also an alternative.

- The variance analysis results show that the transport planners believe investments in different planning phases are equally certain. However, the cost uncertainty should decrease when the project advances towards later stages. Moreover, the variance is significantly lower than actual outcomes especially in small-sized projects. These results indicate that the transport planners still calculate the project costs too optimistically. Thus, we conclude that applying the Successive Calculation in Swedish transport projects do not put enough emphasis on uncertainties.
5. Methods to improve cost estimations – Reference Class Forecasting

As explained in the previous chapter, the Successive Calculation is one of the methods to improve cost estimations. Comparing to the conventional contingency approach, the Successive Calculation is better in terms of considering the risk and uncertainty in cost estimations. However, using group analysis in the Successive Calculation means it still has human bias. In this chapter, we present another method called “Reference Class Forecasting - RCF” which bypasses human bias. Human bias can be because of optimism bias or strategic misrepresentation. In some studies, it is claimed that the method promises more accuracy in forecasts or estimations by taking an “outside view” on prospects being estimated, while the conventional contingency method and also the Successive Calculation take an inside view. The key concept of this method is to examine the experiences of a class of similar projects, lay out a rough distribution of outcomes for this reference class, and then position the current project in that distribution. Apart from the RCF method, we discuss some similar methods, namely hybrid estimating approach and Risk-Based estimating (RBE).

5.1 Reference Class Forecasting method

Reference Class Forecasting is based on theories of decision making under uncertainty that won Princeton psychologist Daniel Kahneman the Nobel prize in economics in 2002. Later, the concept of RCF method was developed by Bent Flyvbjerg.

This section shows the general procedure, applications and limitations of the RCF method. First, we explain why using the reference class forecasting can make the cost estimations more accurate. In general, the main explanations of cost overruns are technical, psychological and political-economic. As was specified in some studies, psychological and political-economic explanations seem to be the main reason for inaccurate forecasts. Psychological explanations account for inaccuracy in terms of optimism bias, that is, a cognitive predisposition found with most people to judge future events in a more positive light than is warranted by actual experience. Political-economic explanations explain inaccuracy in terms of strategic misrepresentation. Here, when forecasting the outcomes of projects (in this thesis – cost estimations), forecasters and planners deliberately and strategically overestimate benefits and underestimate costs in order to increase the likelihood that it is their projects, and not the competitors, that gain approval and funding. Optimism bias and strategic misrepresentation are both deception, but where the latter is intentional, the first is not since optimism bias is self-deception. Although the two types of explanation are different, the result is the same: inaccurate estimations. As shown in Figure 5-1, explanations in terms of optimism bias have their relative...
Methods to improve cost estimations – Reference Class Forecasting

merit in situations where political and organizational pressures are absent or low, whereas explanations in terms of strategic misrepresentation have their relative merit where political and organizational pressures are high. Competition between projects and authorities creates political and organizational pressures. An example of the political pressures is an interview with a planner at UK local transport authority which was published in a study[^2]:

> You will often as a planner know the real costs. You know that the budget is too low but it is difficult to pass such a message to the counsellors [politicians] and the private actors. They know that high costs reduce the chances of national funding.

Both optimism bias and strategic misrepresentation are examples of human bias. The bias can be found in the inside view decision. This is because experts and the other team members make forecasts by focusing tightly on the case at hand, considering its objective, the resources they need to carry it out, and the obstacles to its completion. They construct in their minds scenarios of their coming progress and extrapolate current trends into the future. The resulting forecasts, even the most conservative ones, have often been shown to be overly optimistic. Even applying the Successive Calculation, cannot fully eliminate the human bias. We may reduce the strategic misrepresentation by choosing the group analysis members who are not directly involved in the project. However, they still have optimism bias because people make decisions based on their intuition or on past experiences.

The Reference Class Forecasting method bypasses human bias by cutting directly to outcomes. Thus, it is an outside view decision. It completely ignores the details of the project at hand, and it involves no attempt at forecasting the events that influence the project’s future course. Instead, it examines the experiences of a class of similar projects, lays out a rough distribution of outcomes for this reference class, and then positions the current project in that distribution.
5.1.1 General procedure

The Reference Class Forecasting for a particular project requires the following three steps:

**Step 1** Identification of a relevant reference class of past or similar projects

The class must be broad enough to be statistically meaningful but narrow enough to be truly comparable with the specific project.

**Step 2** Establishing a probability distribution for the selected reference class

This requires access to credible, empirical data for a sufficient number of projects within the reference class to make statistically meaningful conclusions.

**Step 3** Comparing the specific project with the reference class distribution

This step is in order to establish the most likely outcome for the specific project.

Figure 5-2 shows a stylized example of the probability distribution for budget increase in a selected reference class of projects. It also illustrates the link between the observed ex-post cost increases for historical projects and the required up-lift for a new project to ensure that the probability of the final cost being higher than the initial budget plus the up-lift is less than a given threshold level.

If the new project is similar to the projects in the reference class (same type of transport scheme) and the initial budget is established in a similar manner (not including budget
Methods to improve cost estimations

– Reference Class Forecasting

contingencies reflecting the risk of cost overruns above the level of contingencies used in the reference cases), the project should be placed at the point marked “initial budget”.

It should be expected that the final budget – on average – will exceed the initial budget by the average budget increase. This also implies that there is 50% chance of the budget increase being less than the average budget increase and 50% of the budget increase being higher than the average budget increase. If it is not acceptable that there is a 50% chance of the realized cost being higher than the budget at the average budget increase in Figure 5-2, the up-lift needs to be higher. Figure 5-2 shows an example of the necessary uplift to ensure that the probability of a realized cost above the budget (including up-lift) is below a given threshold (x%).

![Necessary up-lift to ensure that probability of higher cost is less than X%](image)

**Figure 5-2** Definition of Up-lifts within a certain class of transport scheme.

5.1.2 Applications and limitations

- **Applications**

The first application of the RFC approach was carried out 2004-2005 under the auspices of HM Treasury and the UK Department for Transport [2] [4]. In 2005, this method was also officially endorsed by the American Planning Association and since then it has been used by governments and private companies in the UK, the Netherlands, Denmark, Switzerland, Australia, and South Africa, among others [2]. In this literature study, the application for cost
estimations in transport projects is reviewed. We have studied examples of applying this method in the UK, the Netherlands, South East Asian, Australia and Canada.

The first study in the methodology for systematic, practical Reference Class Forecasting for transport projects was developed 2003-2004 and published by the UK Department for Transport in 2004\(^4\)\(^{47}\). Based on the study\(^47\), the Treasury and the Department for Transport decided to employ Reference Class Forecasting as part of project appraisal for large transport projects in the UK.

In the study, they started to do step 1 in the three-step procedure for RFC method described earlier. The specific categories and the types of project allocated to each category are shown in Table 5-1\(^1\). Projects are placed in the same reference class if tests showed no significant differences between projects.

**Table 5-1** Categories and types of projects used as basis for Reference Class Forecasting.

<table>
<thead>
<tr>
<th>Category</th>
<th>Types of projects</th>
<th>Source of optimism bias uplifts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roads</td>
<td>Motorway/Highways</td>
<td>Reference class of 172 road projects (of which 128 are British)</td>
</tr>
<tr>
<td></td>
<td>Trunk roads</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Local roads</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bicycle facilities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pedestrian facilities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Park and ride</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bus lane schemes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Guided buses on wheels</td>
<td></td>
</tr>
<tr>
<td>Rails</td>
<td>Metro</td>
<td>Reference class of 46 rail projects (of which 3 are British)</td>
</tr>
<tr>
<td></td>
<td>Light rail</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Guided buses on tracks</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Conventional rail</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High speed rail</td>
<td></td>
</tr>
<tr>
<td>Fixed links</td>
<td>Bridges</td>
<td>Reference class of 34 bridge and tunnel projects (of which 4 are British)</td>
</tr>
<tr>
<td></td>
<td>Tunnels</td>
<td></td>
</tr>
</tbody>
</table>

For each category of projects, a reference class of completed, comparable transport projects is used to establish probability distributions for cost overruns for new projects, as required by step 2 in Reference Class Forecasting. The histograms of cost overruns are initiated first and then the probability distributions are provided. Based on the probability distributions, the required

\(^1\) In this thesis, we focus only in road, rail and fixed link projects. However, the study also shows results of building projects, IT projects, standard civil engineering and non-standard civil engineering.
uplifts are needed to carry out step 3 in Reference Class Forecasting. Figures 5-3 to 5-5 show the results from the UK study of step 2 and 3 for roads, rails and fixed links, respectively. The required uplifts share the same basic S-shape, but at different levels, demonstrating that the required uplifts are significantly different between project categories for a given level of risk of cost overrun.

From the required uplifts, the lower the acceptable risk for cost overrun, the higher the uplift. For instance – a rail project (Figure 5-4), with a willingness to accept a 50% risk for cost overrun in a rail project, the required uplift for this project is 40%. If the decision planners are willing to accept only a 10% risk for cost overrun, then the required uplift is 68%.

Figure 5-3 Probability distribution of cost overruns for road projects according to the UK study.
Methods to improve cost estimations – Reference Class Forecasting

Figure 5-4 Probability distribution of cost overruns for rail projects according to the UK study.

Figure 5-5 Probability distribution of cost overruns for fixed links projects according to the UK study.
Table 5-2 Required uplifts by the Reference Class Forecasting according to the UK study.

<table>
<thead>
<tr>
<th>Category</th>
<th>Types of projects</th>
<th>Applicable optimism bias uplifts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>50% per-centile</td>
</tr>
<tr>
<td>Roads</td>
<td>• Motorway</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>• Trunk roads</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Local roads</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Bicycle facilities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Pedestrian facilities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Park and ride</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Bus lane schemes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Guided buses on wheels</td>
<td></td>
</tr>
<tr>
<td>Rail</td>
<td>• Metro</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>• Light rail</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Guided buses on tracks</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Conventional rail</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• High speed rail</td>
<td></td>
</tr>
<tr>
<td>Fixed links</td>
<td>• Bridges</td>
<td>23%</td>
</tr>
<tr>
<td></td>
<td>• Tunnels</td>
<td></td>
</tr>
</tbody>
</table>

Table 5-2 presents an overview of applicable optimism bias uplifts for the 50% and 80% percentiles for projects in Table 5-1. The 80% percentile corresponds to a risk of cost overrun of 20% - is the level of risk that the UK Department for Transport is typically willing to accept.

From Table 5-2, if a group of planners are preparing the business case for a new motorway, and if they or their client decide that the risk of cost overrun must be less than 20%, they should then use an uplift of 32% on their estimated capital expenditure budget. Thus, if the initially estimated budget was £100 million, then the final budget-taking into account the bias at the 80%-level is £132 million. If the planners or their client decide instead that a 50% risk of cost overrun is acceptable, the uplift is 15% and the final budget is £115 million.

The results from the study above are referenced in many documents such as in a study by Kim Bang Salling and other [48], Developing Harmonised European Approaches for Transport Costing and Project Assessment (HEATCO) [49].

In October 2004, the first instance of practical use of the optimism bias uplifts was recorded, in the planning of the Edinburgh Tram Line 2. Ove Arup and Partners Scotland had been appointed by the Scottish Parliament’s Edinburgh Tram Bill Committee to provide a review of
the Tram business case developed on behalf of Transport Initiatives Edinburgh (TIE). TIE is project promoter and is a private limited company owned by the City of Edinburgh Council established to deliver major transport projects for the Council. The business case had estimated a base cost of £255 million and an additional allowance for contingency and optimism bias of £64 million – or 25% - resulting in total capital costs of approximately £320 million. Using the result from the study above, Arup then calculated the 80th percentile value for total capital costs (the value at which the likelihood of staying within budget is 80%) was £400 million (i.e. £255 million x 1.57, see Table 5-2). The 50th percentile for total capital costs (the value at which the likelihood of staying within budget is 50%) was £357 million (i.e. £255 x 1.40).

Arup remarked that these estimates of total costs were likely to be low, because the optimism bias uplifts should be applied to the budget at the time of decision to build, which typically equates to business case submission. The Tram had not yet even reached the outline business case stage, indicating that risks are substantially higher at this early stage and that this would affect uplifts. On that basis Arup concluded that the optimism bias uplifts may had been underestimated. Moreover, Arup mentioned that the study does allow for optimism bias to be adjusted downward if strong evidence of improved risk mitigation can be demonstrated. However this was not the case for the Tram. Thus the overall conclusion of Arup was that the promoter’s capital cost estimate of app. £320 million was optimistic. Most likely Tram Line 2 would cost significantly more [4]. Today the line is constructed and Arup has been proven right. As mentioned in Chapter 1, the cost of this tram system is now uncertain and it is anticipated to be over £600 million.

The first Reference Class Forecasting in the Netherlands was carried out in 2006 for cost and traffic projections on the proposed high-speed rail Zuiderzee Line project[50]. As step one of the reference class forecast of costs – and in order to establish the outside view – a reference class of 68 comparable rail projects was established. As step two, the probability distribution of cost overrun in the reference class was found. The distribution is shown in Figure 5-6. Finally, on the basis of this distribution and the base cost (€4,084 million), it was established that the expected construction cost of the Zuiderzee Line project was €5,660 million (or approximately 38% uplifts) at 50% risk of cost overruns. If the decision maker was willing to accept only a 20% risk of cost overrun, then the budget should be €6,084 million (or approximately 64% uplifts). The project has since been cancelled; therefore these forecasts cannot be evaluated.
For South East Asia, a study about the application of the RCF method to transport projects in selected South East Asian countries (Thailand and the Philippines) was carried out in 2008[34]. For Thailand, there are a total of 44 transport projects, 15 bridges and 29 roads. For the Philippines, there are a total of 85 transport projects, 25 bridges and 60 roads. In both courtiers, from statistic testing, there is no necessity to separate the reference class to road and bridge. Figure 5-7 shows the results of the RCF application for transport projects in Thailand. If we allow 10% risk that cost overrun may happen, a 15% uplift of the budget is needed. Figure 5-8 shows the results of the RCF application for transport project in the Philippines. Here 22% uplift is needed for the same risk. However, findings from the study should be used and interpreted with caution due to the small sample size and limited types of projects are considered in the study as mentioned in Chapter 2.
Methods to improve cost estimations – Reference Class Forecasting

Figure 5- 7 Road projects in Thailand.
Next, we summarize a study in 2010 that presents the results of estimation accuracy of a sample of road projects (between 2000 and 2005) conducted by the Australian State Road & Traffic Authority using a hybrid estimating approach blending primarily RCF with fixed contingency approach \cite{51}. To clarify about the hybrid estimating, there is an Authority’s estimating manual which specify the contingency value for each project depending on the planning stage and the project type of the estimate. For example, in the manual, contingency for detailed estimates needs to be within the range of 10-20% and 20-35% at the concept estimate stage. Contingencies outside the ranges specified in the manual need to be justified. The contingency ranges are determined based on historical performance data - a practice that mirrors the RCF approach. Thus it is called hybrid estimating. For the conventional project
sample, it is a sample of Australian construction projects (before 1999-2000) using the conventional fixed contingency approach.

The last sample is the Risk-Based Estimating (RBE) sample. It contains 11 water infrastructure projects in Australia (between 2003 and 2007) using the RBE approach. The RBE has recently been introduced in Australia. The RBE models the cost of individual components with base estimates and stochastic risk contingencies. The distribution of the overall project cost is derived by summing the stochastic cost.

The study finds that the average accuracy of the sample using the RCF (or hybrid approach) compares favorably to historical results. They analyzed cost overruns of a variety of projects in the past and got an average cost overrun at 34.7% and standard deviation 37.8%, which those projects typically are used the fixed contingency approach. A comparison with the conventional contingency approach shows that the hybrid approach can produce more conservative estimates than the conventional contingency approach, either by less overbudgeting or, more likely, underbudgeting. Moreover, a comparison with a small RBE sample provides the preliminary evidence that RBE can in turn be more accurate than the hybrid approach. The descriptive statistics of cost estimations are shown in Table 5-3.

Table 5-3 Descriptive statistics of estimation errors in the three samples in Australia.

<table>
<thead>
<tr>
<th>Sample</th>
<th>N</th>
<th>Mean</th>
<th>Std. deviation</th>
<th>Std. error mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation errors</td>
<td>Hybrid</td>
<td>36</td>
<td>−9.71a</td>
<td>14.48</td>
</tr>
<tr>
<td></td>
<td>Conventional</td>
<td>38</td>
<td>3.88</td>
<td>10.09</td>
</tr>
<tr>
<td></td>
<td>RBE</td>
<td>11</td>
<td>−3.49</td>
<td>5.95</td>
</tr>
</tbody>
</table>

*a* A negative value denotes under budget and a positive figure indicates over budget.

Recently, a related study was published (in January 2011) – Incorporating Risk of cost overruns into transportation capital projects decision-making [52]. In the study, the distribution-fitting method was used and the concept is similar to the RCF method. The main objectives of the study are to propose methods for estimating the risk of cost overruns and suggest ways to incorporate it into project decision-making. They proposed distribution-fitting models to create

---

2 The conventional contingency approach adds a percentage to the base estimate to account for risks. The contingency amount is typically determined based on an estimator’s intuition or past experience. The contingency percentage depends on the type and life cycle phase of the project. [51]
the probabilities of cost overruns for the Vancouver Island Highway Projects (VIHP). The VIHP database comprises of 163 projects in Vancouver Island (1993-2003), 127 highway projects and 36 bridge and tunnel projects. They used the Cost Overrun Ratio (COR) instead of the traditional cost overrun calculation. The COR is the ratio of real costs to planned budget. For example, if the budget for a project is $1 billion and the actual cost is $1.2 billion, COR = 1.2; if the actual cost is only $0.8 billion for the $1 billion budget, then COR = 0.8. The COR histograms for the highway and bridge projects in Vancouver Island are shown in Figure 5-9. They applied a variety of distributions and concluded that the Cauchy distribution is best fit to the COR distribution of the projects in Vancouver Island. The three risk levels are categorized: low (COR ≤ 1), medium (1 < COR ≤ 1.2), and high (COR > 1.2). The result of three different cost overrun risk levels in each distribution is shown in Table 5-4 (only for highway projects). It is comparable with the probability distribution in step two of the RCF method.

![Histogram of COR for All Projects](image1)

**Figure 5-9** COR histograms for highway and bridge projects in Vancouver Island.
Methods to improve cost estimations – Reference Class Forecasting

Table 5-4 COR probabilities of highway projects in Vancouver Island.

<table>
<thead>
<tr>
<th>COR</th>
<th>Observed simulation</th>
<th>Normal</th>
<th>Gamma</th>
<th>Beta</th>
<th>Cauchy</th>
</tr>
</thead>
<tbody>
<tr>
<td>COR ≤ 1</td>
<td>33.07%</td>
<td>42.33%</td>
<td>42.38%</td>
<td>43.73%</td>
<td>42.92%</td>
</tr>
<tr>
<td>I &lt; COR ≤ 1.2</td>
<td>60.63%</td>
<td>42.33%</td>
<td>38.95%</td>
<td>41.30%</td>
<td>33.06%</td>
</tr>
<tr>
<td>COR &gt; 1.2</td>
<td>6.30%</td>
<td>15.34%</td>
<td>18.67%</td>
<td>14.97%</td>
<td>24.02%</td>
</tr>
</tbody>
</table>

- Limitations

The RFC method was officially published in 2004. However, the real practical applications of this method are scarce. The real challenge in doing a Reference Class Forecasting lies in assembling a valid dataset that will allow a reliable forecast. Such datasets are rare in real-life policy-making and planning. Large publicly financed projects usually take a very long time from decision to build until open to service. So reconstructing the actual total costs of a public project is typically difficult and complex. For private projects, the cost information tends to be kept only within companies or owners. And for both publicly and privately financed projects, project owners are not encouraged to show cost escalations data of their projects. It is also quite time consuming to produce such data. For example, four years of data collection and refinement were needed for the study of Bent Flyvbjerg et al. to establish a sample of 258 transport projects with data on both actual construction costs and estimated costs at the time of decision to build. In sum, establishing reliable data on actual costs for even a single transport project is often highly time consuming or simply impossible.

Choosing the right reference class of comparative past projects becomes more difficult when evaluating initiatives, for which precedents are not easily found, for instance the introduction of new and unfamiliar technologies. Moreover even the same technology, but implemented in different areas or countries, may give different results.

The potentials for and barriers to Reference Class Forecasting will be different in situations where (1) optimism bias is the main cause of inaccuracy as compared to situations where (2) strategic misrepresentation is the chief reason for inaccuracy. In the first type of situation, the potential for using the outside view and Reference Class Forecasting will be good. Forecasters can be encouraged to use any methods that will improve their forecasts. However in the second type of situation, the potential for Reference Class Forecasting is low and barriers are high. That means
the forecasters may not be interested in other estimation methods to improve their forecasts because inaccuracy of the forecasts is deliberate. In order to lower barriers, incentives much be aligned to reward accurate forecasts and punish inaccurate ones.

One weakness of the RFC methods is that it is based on historical outcomes. Thus it may fail to predict or estimate extreme outcomes, that is, those that lie outside all historical precedents. However, this is a problem that is shared with all methods.

Another disadvantage is that the method is based on aggregated follow-ups of completed projects. No information is therefore obtained on which components of a project constitutes the largest risks for cost overruns. Since a new project will never be exactly the same as a previous one, there is a risk that the reference class is not fully relevant.

Lastly, it could be argued that in some cases the use of Reference Class Forecasting may result in such large reserves set aside for a project that this would in itself lead to risks of inefficiencies and overspending in the same way as when the Successive Calculation which was discussed in Chapter 4. It is therefore important to combine the introduction of Reference Class Forecasting and optimism bias uplifts with tight contracts, maintained incentives for promoters to undertake good risk assessment, and prudent control during procurement and project implementation.

5.1.3 Probability distributions

For both road and rail projects, in this thesis, it is quite straightforward to do the step 1 and step 3 of RCF method. For step 2, we first need to find the most suitable distribution to fit the VTI database. To determine this, we use the Maximum Likelihood Estimation (MLE) method and review the previous studies. The following is the summary of the literature study.

The first document is a master thesis – “Transportation Infrastructure Project Cost Overrun Risk Analysis” in 2006[11]. The cost overrun ratio (COR) distribution fitting of Vancouver Island Highway Projects (VIHP) data was provided. The Distribution Fitting method (by Beta distribution) and Monte Carlo method were used. According to the study, the advantage of the Distribution Fitting method is that it is based on real data and therefore its results are expected to be more objective and realistic. Hence, they recommended using the Distribution Fitting method. Furthermore, they focused on Beta Distribution Fitting, since COR is expected to follow Beta distributions. The result of estimating COR probabilities from both methods is shown in Table 5-5. As explained earlier, the COR is the ratio of real costs to planned budgets. For example, if the budget for a project is $1 billion and the actual cost is $1.2 billion, COR = 1.2; if the actual cost is only $0.8 billion for the $1 billion budget, then COR = 0.8. From Table 5-5, 57.9% of the projects had cost overruns
based on Monte Carlo Simulation and Beta Distribution Fitting. At the same time, 87.6% and 88.2% of the projects had the ratio of actual costs to budgets higher than 0.75 based on Monte Carlo Simulation and Beta Distribution Fitting, respectively.

Table 5- 5 Cost overruns probability estimation comparison between Monte Carlo method and Beta Distribution Fitting method.

<table>
<thead>
<tr>
<th>For Project Cost Overrun Ratio</th>
<th>Monte Carlo Simulation</th>
<th>Beta Fitting Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>98.7%</td>
<td>99.0%</td>
</tr>
<tr>
<td>0.75</td>
<td>87.6%</td>
<td>88.2%</td>
</tr>
<tr>
<td>1</td>
<td>57.9%</td>
<td>57.9%</td>
</tr>
<tr>
<td>1.25</td>
<td>22.8%</td>
<td>22.8%</td>
</tr>
<tr>
<td>1.5</td>
<td>3.7%</td>
<td>4.3%</td>
</tr>
<tr>
<td>1.75</td>
<td>0.1%</td>
<td>0.3%</td>
</tr>
<tr>
<td>2</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>2.25</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

The COR probabilities of VIHP data were also studied in another study. It is “Incorporating risk of cost overruns into transportation capital projects decision-making” [52] in 2011. In the study, they fitted the following four distributions:

- **the Normal distribution:** $N(\mu, \sigma)$, where the mean $\mu$ and the standard deviation $\sigma$ are the location and scale parameters, respectively;
- **the generalized Gamma distribution:** $\text{Gamma}(\theta, \beta, \gamma)$ where $\theta$ is the location parameter (which reflects the minimum value of the distribution’s variable), $\beta$ and $\gamma$ are the scale and shape parameters, respectively;
- **the Beta distribution:** $\text{Beta}(P, Q, A, B)$, where $P$ and $Q$ are shape parameters, and $A$ and $B$ are scale parameters, reflecting the minimum and maximum values of the distribution’s variable;
- **the Cauchy distribution:** $\text{Cauchy}(a,b)$ where $a$ is the location parameter (indicating the median of the distribution’s variable), and $b$ is the scale parameter.

The Goodness-of-Fit Statistics are the log-likelihood (LL), the Kolmogorov-Smirnov (KS) and the Chi-square ($x^2$) test. For the road and highway project data, for example, the Goodness-of-Fit Statistics of each distribution is shown in Table 5-6. It is obvious that the Cauchy distribution is the best by providing the maximum Log-Likelihood. Regarding to the goodness-of-Fit test, the null hypothesis that a given distribution “correctly” represents the data is rejected when the test statistics for the KS and $x^2$ are less than the Critical Value (CV) at a given level of significance. The test results in Table 5-6 indicate that the values for KS and $x^2$ for the Cauchy distribution are the smallest compared with the other three distributions tested. The results further show that the $p$-value for the Cauchy distribution model is significantly higher than that for the other three
models. They concluded, therefore, that the Cauchy distribution best fits the COR distribution of the VIHP database. This conclusion was further ascertained by running the same tests on the Flyvbjerg database.\textsuperscript{[22]}

**Table 5-6** Goodness-of-fit Statistics of distributions of road and highway projects in VIHP.

<table>
<thead>
<tr>
<th>Fitting statistics</th>
<th>Normal</th>
<th>Gamma</th>
<th>Beta</th>
<th>Cauchy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log-likelihood (LL)</td>
<td>34.7382</td>
<td>55.3480</td>
<td>105.9150</td>
<td>167.3666</td>
</tr>
<tr>
<td>Kolmogorov-Smirnov test</td>
<td>KS 0.3236</td>
<td>0.2871</td>
<td>0.2405</td>
<td>0.1388</td>
</tr>
<tr>
<td>p-value</td>
<td>3.04e–012</td>
<td>9.92e–010</td>
<td>1.63e–013</td>
<td>0.0133</td>
</tr>
<tr>
<td>Chi-square test</td>
<td>$\chi^2$ 1.47e+005</td>
<td>9.21e+003</td>
<td>2.58e+003</td>
<td>7.6401</td>
</tr>
<tr>
<td>p-value</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0219</td>
</tr>
</tbody>
</table>

*Note: KS test CV: 0.1073 ($\alpha = 10\%$); 0.1191 ($\alpha = 5\%$); 0.1429 ($\alpha = 1\%$). Chi-square test CV: 4.6052 ($\alpha = 10\%$); 5.9915 ($\alpha = 5\%$); 9.2103 ($\alpha = 1\%$).*

From a PhD thesis – “Assessment of transport Projects – Risk analysis and decision support” in 2008 \textsuperscript{[53]}, the following four conditions for estimating construction costs with probability distributions had been proposed:

- Upper and lower limits which the analyst is relatively certain the values do not exceed. Consequently, a closed-ended distribution is desirable.
- The distribution must be continuous.
- The distribution will be unimodal; presenting a most likely value.
- The distribution must be able to have a greater freedom to be higher than lower with respect to the estimation – skewness must be expected.

There are three distributions that fulfill the conditions above - Triangular distribution, Beta-PERT distribution and Gamma distribution. Moreover, Steen Lichtenberg, who has developed the principle of the Successive Calculation, further documents the applicability of an Erlang distribution for the estimation of the construction costs which corresponds to the conditions above. The properties of the Erlang distribution require a shape ($k$) and a scale ($\theta$) parameter. The Erlang distribution is a special case of the Gamma distribution where the shape ($k$) is an integer. In the Gamma distribution, this parameter is not restricted to the integers. Based on experience it is found that a shape parameter in the range of $k = 5$-15 matches the distribution of the variability uncertainty for construction costs. In fact the Erlang function with $k = 1$ is identical to the exponential distribution. Using $k = 5$ the function resembles a Lognormal distribution. Finally, when $k \geq 10$, the distribution is closer to the Gaussian distribution (Normal distribution).
In conclusion from the study, the Erlang distribution is recommended to apply for the construction costs derivation in transport appraisal. Figure 5-10 and 5-11 illustrate the fits of using the Erlang distribution for the data defined from Bent Flyvbjerg study \(^{[28]}\). For road projects, the distribution function is fitted with a shape parameter of \(k = 8\) and a scale parameter of \(\theta = 0.09\). For rail projects, the distribution function is fitted with a shape parameter of \(k = 23\) and a scale parameter of \(\theta = 0.075\).

Lastly, a study – “The accuracy of hybrid estimating approaches: A case study of an Australian State Road & Traffic Authority” in 2010 states that past studies showed that Lognormal distribution is appropriate for modeling cost estimation performance \(^{[51]}\). Figure 5-12 shows the lognormal distribution curves of cost accuracies for each case in the study.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fit_comparison}
\caption{Inaccuracies of cost estimates of 167 road projects by the Erlang distribution. (database from Flyvbjerg et al., p.17\(^{[28]}\))}
\end{figure}
Figure 5-11 Inaccuracies of cost estimates of 58 rail projects by the Erlang distribution. (database from Flyvbjerg et al., p.17 [28])

Figure 5-12 Inaccuracies of cost estimations of the hybrid, RBE and convention methods by Lognormal distribution.
Summary

A Distribution Fitting method using Maximum Likelihood Estimation (MLE) is applied in this thesis while the Monte Carlo method is not considered. The MATLAB Distribution Fitting Tool (Dfittool) function is used for the analyses. The Distribution Fitting Tool is a function for fitting a variety of distributions to data. There are 21 main distributions in the function such as Beta, Exponential, Gamma, Lognormal, Normal and Poisson. However, not all of the distributions are available for all data sets. The function determines the extent of the data (nonnegative, unit interval, etc.) and displays appropriate distributions in the distribution drop-down list. For our data, the function displays the following appropriate distributions; extreme value, generalized extreme value, generalized Pareto, logistic, non-parametric, normal and t location-scale distribution. Figure 5-13 shows examples of each distribution except the normal distribution. In the next section we will discuss which distribution best fits our data.
Methods to improve cost estimations – Reference Class Forecasting

Figure 5-13 Examples of distributions in the thesis.
5.2 Application in case studies

In this section, we apply the RCF method to estimate the project costs of our case studies. There are two case studies. The case study for road projects is the Stockholm bypass (Förbifart Stockholm). The case study of rail projects is the Västlänken. We create the uplifts for road and rail projects by using the VTI database as the reference class. The uplifts results are compared with the previous study by Bent Flyvbjerg. Moreover, the cost estimations of the two case studies are compared with the total costs estimated by the Successive Calculation.

5.2.1 Case study projects

The Stockholm bypass (Förbifart Stockholm) is a new motorway linking southern and northern Stockholm (as shown in Figure 5-14) resulting in a new route for the European highway (E4) past Stockholm.

The new link west of Stockholm has been under investigation for several decades and a large number of different alternatives have been studied\textsuperscript{[54]}. To reduce the impact on sensitive natural and cultural environments, 18 km of the total of 21 km of the project are in tunnels.

The construction work is planned to start in 2012 and it will take 8 to 10 years to finish. When the link opens for traffic it will be one of the longest road tunnels in the world. By 2035, the Swedish Transport Administration (Trafikverket) estimates that the project will be used by 140,000 vehicles per day.

The second case study is the Västlänken. It is about 8 km double-track railway tunnel in Gothenburg. It will increase the accessibility of the city with two new stations, at Haga and Korsvägen (as shown in Figure 5-15). The Trafikverket is planning for it. The total construction period will be 9 to 10 years. When the link opens to service, it will connect the commuter rail service to the existing lines and open up the new stations in the city.
Figure 5-14 Stockholm bypass (Förbifart Stockholm).
Route of Västlänken

Station

Figure 5-15 Västlänken.
5.2.2 Results of road projects

For road projects, we use the 102 road projects to be the reference class. The following is the details in each step of the RCF method.

Step 1 Identification of a relevant reference class of past or similar projects

We do not divide our reference class to a number of categories. As mentioned in Section 3.3.4, it is impossible to establish valid optimism bias uplift for each category. Hence, the reference class for road projects is shown in Table 5-7.

Table 5-7 Types of projects used as basis for Reference Class Forecasting of road projects.

<table>
<thead>
<tr>
<th>Category</th>
<th>Types of projects</th>
<th>Source of reference class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roads</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Major road</td>
<td>Reference class of 102 road projects</td>
</tr>
<tr>
<td></td>
<td>• Motorway</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Secondary road</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Urban road</td>
<td></td>
</tr>
</tbody>
</table>

Step 2 Establishing a probability distribution for the selected reference class

The probability curves of each distribution are shown in Figure 5-16. The Likelihood of location-scale, normal and generalized extreme value distributions are -440.42, -470.94 and -453.99 respectively. Thus the location-scale distribution is best fit for our road project data. Figure 5-17 shows the cumulative probability of cost overruns by the distribution. On the basis of the probability distribution, required uplifts are calculated as shown in Figure 5-18. The summary of uplifts for selected percentiles for road projects is shown in Table 5-8. With a willingness to accept a 50% risk for cost overrun in a road project, the required uplift will be 5%. If a planner is willing to accept only a 10% risk for cost overrun, then the required uplift will be 24%. Compared with the previous study by Bent Flyvbjerg and COWI consultant [47], the road projects in our study require lower uplifts.

We use the required uplift results to compare with the action of planners. As shown in Figure 5-19, using the RFC method, the different between cost estimation at 85% confident and 50% confident is approximately 13%. From the previous chapter, we find that the average difference is around 9% (when the Successive Calculation is applied in cost calculations). It shows that planners seem to estimate project costs with either optimism bias or strategic misrepresentation.

3 For Stockholm bypass (Förbifart Stockholm), it was 5%.
However, we cannot conclude that using the RCF will result in a higher estimated project cost because we do not know the project initial budgets.

**Figure 5-16** Probability curves of road projects.

**Figure 5-17** Distribution of cost overruns of road projects.
Methods to improve cost estimations – Reference Class Forecasting

Figure 5-18 Required uplift of road projects.

Table 5-8 Cost uplifts for selected percentiles of road projects.

<table>
<thead>
<tr>
<th>Category</th>
<th>Types of projects</th>
<th>Required uplifts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>50% percentile</td>
</tr>
<tr>
<td>Roads</td>
<td>Major road</td>
<td>5% (15%)</td>
</tr>
<tr>
<td></td>
<td>Motorway</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Secondary road</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Urban road</td>
<td></td>
</tr>
</tbody>
</table>

Note: ( ) is the required uplift from Bent Flyvbjerg study.
**Step 3** Comparing the specific project with the reference class distribution

By using the Successive Calculation for the Stockholm bypass, the expected project cost is 27.9 billion SEK at 50% risk of cost overruns. At 15% risk of cost overruns, or 85% confident that costs will be within budget, the cost is 29.2 billion SEK.

We want to apply the RFC method to estimate cost of the Stockholm bypass project. One constraint is that the initial budget is not provided. We therefore assume that the cost at 50% confident is the same as calculated by the Successive Calculation. From the result in step 2, the required uplift at the likelihood of 50% of staying within budget is 5%. Thus the assumed initial budget is 26.6 billion SEK (27.9 billion SEK / 1.05). As the required uplift for the 85% confident is 19%, the estimated cost equals 31.7 billion SEK (i.e. 26.6 billion SEK x 1.19). In this case, the project cost at 85% confident level by using the RCF method is higher than the corresponding cost using the Successive Calculation.
5.2.3 Results of rail projects

For rail projects, we use the 65 rail projects in the VTI database to be the reference class. The following is the details in each step of the RCF method.

**Step 1 Identification of a relevant reference class of past or similar projects**

As for road projects, we do not divide our reference class to a number of categories. As mentioned earlier, it is statistically significant that the rail projects are similar in variances in each detailed project type. Hence, we can combine analysis of each detailed project type together. The reference class for rail projects is shown in Table 5-9.

**Table 5-9** Types of projects used as basis for Reference Class Forecasting of rail projects.

<table>
<thead>
<tr>
<th>Category</th>
<th>Types of projects</th>
<th>Source of reference class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rails</td>
<td>• New track, existing line</td>
<td>Reference class of 65 rail projects</td>
</tr>
<tr>
<td></td>
<td>• New track, new line</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Station and rail yard</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Upgrading existing line</td>
<td></td>
</tr>
</tbody>
</table>

**Step 2 Establishing a probability distribution for the selected reference class**

The probability curves of each distribution are shown in Figure 5-20. The Likelihood of t location-scale, normal and generalized extreme value distributions are -332.42, -346.69 and -327.67 respectively. Thus the distribution by generalized extreme value is the best fit for our rail project data. Figure 5-21 shows the cumulative probability of cost overruns by the distribution. On the basis of the probability distribution, required uplifts are calculated as shown in Figure 5-22. The summary of uplifts for selected percentiles for rail projects is shown in Table 5-10. With a willingness to accept a 50% risk for cost overrun in a rail project, the required uplift will be 11%. If a planner is willing to accept only a 10% risk for cost overrun, then the required uplift will be 77%. Comparing with the previous study by Bent Flyvbjerg and COWI consultant [47], the rail projects in our study require lower uplifts at low confidence levels but higher uplifts at high confidence. The explanation of this result is that our rail project data is very high in variance (SD = 50.5%).

We use the required uplift results to see the action of planners. As shown in Figure 5-23, using the RFC method, the difference between cost estimation at 85% confident and 50% confident is approximately 44%. From the previous chapter, we find that the difference is only 12% when the Successive Calculation is applied in cost calculations. It shows that planners seem to estimate...
project costs with very high optimism bias or strategic misrepresentation. The optimism bias is much bigger in rail project cost estimations than in road projects. However, we cannot conclude that using the RCF will result in higher estimated project costs because we do not know the project initial budgets.

**Step 3** Comparing the specific project with the reference class distribution

The Västlänken project is our case study for the rail project. By using the Successive Calculation, the expected project cost was 16.2 billion SEK at 50% risk of cost overruns. At 15% risk of cost overruns or 85% confident that cost will be within budget, the cost is 19.3 billion SEK.

We want to apply the RFC method to estimate the cost of the Västlänken project. As with the road project case study, the initial budget is not provided. In response to this constraint, we assume that the cost at 50% confident is the same as calculated by the Successive Calculation. Thus the assumed initial budget is 14.6 billion SEK (16.2 billion SEK / 1.11). The required uplift for the 85% confident is 60%, so the estimated cost will be 23.4 billion SEK (i.e. 14.6 billion SEK \(x\) 1.60). In this case, the project cost by using the RCF method is much higher than the project cost by using the Successive Calculation.

![Probability curves of rail projects](image-url)

*Figure 5-20* Probability curves of rail projects.
Methods to improve cost estimations – Reference Class Forecasting

Figure 5-21 Distribution of cost overruns of rail projects.

Figure 5-22 Required uplift of rail projects.
Table 5-10 Cost uplifts for selected percentiles of rail projects.

<table>
<thead>
<tr>
<th>Category</th>
<th>Types of projects</th>
<th>Required uplifts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>50% percentile</td>
</tr>
<tr>
<td>Rails</td>
<td>New track, existing line</td>
<td>11% (40%)</td>
</tr>
</tbody>
</table>

Note: ( ) is the required uplift from Bent Flyvbjerg study.

Past Conventional Contingency Method (65 observations of completed projects)

Estimated Cost (EC) + Contingency = Initial Budget (IB)

\[
\begin{align*}
\text{At 50% confident} & : 1.21 \times \text{IB} \\
\text{At 85% confident} & : 1.82 \times \text{IB}
\end{align*}
\]

Diff. ≈ 50.5%

Actual

21% Cost escalation or 1.21 x IB

Present Successive Calculation (46 observations of planned projects)

Estimated Cost (EC) + Contingency = Initial Budget (IB)

\[
\begin{align*}
\text{At 50% confident} & : \text{Factor} \times \text{IB} \\
\text{At 85% confident} & : \text{Factor} \times \text{IB}
\end{align*}
\]

Diff. ≈ 12.1%

Should be better than past

Reference Class Forecasting (Based on 65 completed projects)

Estimated Cost (EC) + Contingency = Initial Budget (IB)

\[
\begin{align*}
\text{At 50% confident} & : 1.11 \times \text{IB} \\
\text{At 85% confident} & : 1.60 \times \text{IB}
\end{align*}
\]

Diff. = 44.1%

Future outcome

Unknown in this study

Figure 5-23 Cost estimations for rail projects.
5.3 Key findings and discussion

- With the first reference class forecasts carried out 2004-2005, it is still difficult to draw conclusions about the impacts of adopting the practice of Reference Class Forecasting. This is an area for further research. However, it is clear that in Sweden planners are increasingly aware of risks and uncertainties in cost estimations of transport projects and it is likely that the international studies referenced in this report have contributed to this.

- The RFC method has some limitations. Conducting the reference class is often highly time consuming or simply impossible. However, it worked well in our study. In our study, as well as in the international ones, it may however fail to predict or estimate extreme outcomes, that is, those that lie outside all historical precedents.

- If the RFC method is to be introduced as a standard tool in Sweden, it is important to combine the introduction of the method with other measures such as tight contracts, maintained incentives for promoters to undertake good risk assessment, and prudent control during procurement and project implementation.

- By the Reference Class Forecasting method, we find that with a willingness to accept a 50% risk for cost overrun in a road project, the required uplift will be 5% and 11%, for road and rail projects respectively. If a planner is willing to accept only a 10% risk for cost overrun, then the required uplift will be 24% and 77%, for road and rail projects respectively.

- For method used for improving transport cost estimations in Swedish transport projects is the Successive Calculation. However, compared with uplifts suggested by the RFC method, we find that planners still estimate project costs with optimism bias or strategic misrepresentation.

- For both case studies, the anticipated project costs by using the RCF method are higher than the project costs by using the Successive Calculation at the equal cost at 50% confident.

- The two case study projects chosen are more complex than the projects in the reference classes. The resulting required uplifts are therefore somewhat uncertain and likely to be too low. However the reference classes collected will be very useful, should the method be applied more frequently in future Swedish transport planning. The material is also large enough to be subdivided further, at least for road projects. Thus reference classes more similar to the project in question can be constructed. Moreover, the costs of each component of completed projects should be collected.
6. Conclusion and future research

- Based on cost estimations and outcomes of 167 road and rail projects in Sweden during the period 1997-2009, the average cost overruns are 11% (SD = 24.6%) and 21% (SD = 50.5%) for road and rail projects, respectively. In Sweden, the average cost overrun in road projects is similar to other countries, while the average cost overrun in rail projects is lower than in other countries. However, the standard deviation of cost overruns in Swedish rail projects is very high.
- The small Swedish transport projects (< 100 million SEK) have much higher cost overruns than large projects. Moreover, the average cost overruns are very low in bigger projects especially in rail projects.
- The cost overruns in road and rail projects in Sweden have been constant for the 13-year period and cost estimates have not improved over time. However this finding is regardless of the change in use of price indexes by Road and Rail Administrations.
- Using the Successive Calculation has advantage in evaluating uncertainty. However, the method has its limits. Some measures need to be incorporated with the use of the Successive Calculation. One recommendation is that the Trafikverket (or the Government) should start a systematic follow-up of the Successive Calculations. Incentive for the project manager who can efficiently control project cost is also an alternative.
- The variance analysis results show that the transport planners believe investments in different planning phases are equally certain. However, one would expect the cost uncertainty to decrease when the project advances towards later stages. Moreover, the variance is significantly lower than actual outcomes - especially in small-sized projects. These results indicate that the transport planners still calculate the project costs too optimistically. Thus, we conclude that applying the Successive Calculation in Swedish transport projects do not put enough emphasis on uncertainties.
- In our opinion, the Successive Calculation method is not likely to significantly reduce the variance of the cost overruns, even though it may reduce the average cost overruns. Special policies for the high variance project types such as small-sized projects ought to be developed.
- By the Reference Class Forecasting method, we find that with a willingness to accept a 50% risk for cost overrun of a project, the required uplift will be 5% and 11%, for road
and rail projects respectively. If a planner is willing to accept only a 10% risk for cost overrun, the required uplift will be 24% and 77%, for road and rail projects respectively.

- For both case studies, the anticipated project costs by using the RCF method are higher than the project costs by using the Successive Calculation at the equal cost at 50% confident.

- The two case study projects chosen are more complex than the projects in the reference classes. The resulting required uplifts are therefore somewhat uncertain and likely to be too low. However, the reference classes collected will be very useful, should the method be applied more frequently in future Swedish transport planning. The material is also large enough to be subdivided further, at least for road projects. Thus reference classes more similar to the project in question can be constructed.

- An interesting way to improve cost calculation might be to develop a cost estimation method which considers the risks of the costs in each individual component based on the experiences of a class of similar projects. This is the same concept as the RBE method used in Australia. It combines advantages from both the Successive Calculation and the RFC method.

- Potential future researches;
  - Further analyses in factors that contribute to cost overruns such as whether cost overruns in Swedish transport projects differ in different areas and cost overrun effects from several factors by multiple regression.
  - Explanations of the cost overruns in Swedish transport projects such as why the largest overruns are found in small projects. Lessons learnt from the success projects should be studied.
  - In-depth analysis of how and why project costs develop over time.
  - Starting building a database of completed project costs with the costs in each individual component.
  - Analysis of the result of the first use of Successive Calculation in terms of general statistics description and also ways to improve the method. Moreover, one assumption that should be investigated is whether Successive Calculation increases the project cost.
  - Developing a decision support system that covers all in Cost Benefit Analysis (CBA) with evaluating uncertainty. Cost estimation is just one component of CBA. To better transport planning, it is very important to develop a tool for all processes of CBA.
## Appendix A: Data from VTI

<table>
<thead>
<tr>
<th>Road 102 projects</th>
<th>Rail 65 projects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Years</strong></td>
<td><strong>Number of projects</strong></td>
</tr>
<tr>
<td>2009</td>
<td>3</td>
</tr>
<tr>
<td>2008</td>
<td>9</td>
</tr>
<tr>
<td>2007</td>
<td>7</td>
</tr>
<tr>
<td>2006</td>
<td>5</td>
</tr>
<tr>
<td>2005</td>
<td>12</td>
</tr>
<tr>
<td>2004</td>
<td>9</td>
</tr>
<tr>
<td>2003</td>
<td>12</td>
</tr>
<tr>
<td>2002</td>
<td>7</td>
</tr>
<tr>
<td>2001</td>
<td>9</td>
</tr>
<tr>
<td>2000</td>
<td>9</td>
</tr>
<tr>
<td>1999</td>
<td>6</td>
</tr>
<tr>
<td>1998</td>
<td>9</td>
</tr>
<tr>
<td>1997</td>
<td>5</td>
</tr>
</tbody>
</table>

| Average actual cost per project (Million SEK) | 368 | Average actual cost per project (Million SEK) | 440 |
| Maximum actual cost (Million SEK) | 4,000 | Maximum actual cost (Million SEK) | 2,438 |
| Minimum actual cost (Million SEK) | 50 | Minimum actual cost (Million SEK) | 6.6 |
| Total actual cost (Million SEK) | 37,568 | Total actual cost (Million SEK) | 28,625 |

| Percentage of projects experiencing escalating project costs¹ | 68% | Percentage of projects experiencing escalating project costs¹ | 57% |

¹ The projects that estimated cost equal actual cost are not included.
Appendix B: Statistical analyses

- **Standard deviation**
  The standard deviation is the square root of the variance. It is a widely used measurement of variability of diversity used in statistics and probability theory. It shows how much variation or dispersion there is from the average (mean). A low standard deviation indicates that the data points tend to be close to the mean, whereas high standard deviation indicates that the data are spread out over a large range of values.

  \[
  SD = \sqrt{\frac{\sum(x_i-\bar{x})^2}{N-1}} \quad \text{Equation (1)}
  \]

- **The independent t-test**
  This test is used to test for a difference between two independent groups on the means of a continuous variable.

  \(\mu_1, \mu_2=\) the true mean of database 1 and 2.

  \(\bar{x}_1, \bar{x}_2=\) the average of database 1 and 2.

  \(N_1, N_2=\) the size of the sample of database 1 and 2.

  \(s_1, s_2=\) the standard deviation of database 1 and 2.

  The Null Hypothesis, \(H_0\)

  \(H_0: \mu_1 - \mu_2 = 0 \) (or \(\mu_1 = \mu_2\)); that is true mean of both data base are the same. The result of the t-test is to see if we can refute this hypothesis.

  The Alternative Hypothesis, \(H_a\)

  \(H_a: \mu_1 > \mu_2\); that is the mean of database 1 is higher than database 2. The test of this \(H_a\) is called a one-tailed test because it considers only one end of the distribution.

  The test statistic \(t\) is calculated from Equation (2):

  \[
  t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}} \quad \text{Equation (2)}
  \]

  The above equation is true only when the sample sizes are equal. When we compare two groups that contain different numbers of participants, we use the pooled variance estimate t-test instead which takes account of the difference in sample size by weighting the variance of each group.
sample. We actually weight by the number of degrees of freedom, which is the sample size minus 1). Therefore, the pooled variance estimate is Equation (3):

$$S_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$$  \hspace{1cm} \text{Equation (3)}

The resulting weighted average variance is then just replaced in the Equation (1):

$$t = \frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}}$$ \hspace{1cm} \text{Equation (4)}

Then, the $p$-value is calculated by using the $t$ distribution. $P$-value is the area under the curve of $t$-score. If the result $p$-value is less than critical value (in this thesis is $p \leq 0.1$), the null hypothesis can be rejected and it is concluded that true mean of both data base are not the same.

- **Levene’s test**

Levene’s test is an inferential statistic used to assess the equality of variances in different samples. It tests the null hypothesis that the population variances are equal (called homogeneity of variance). If the result $p$-value of Levene’s test is less than critical value (in this thesis is $p \leq 0.1$), the obtained differences in sample variances are unlikely to have occurred based on random sampling. Thus, the null hypothesis of equal variances is rejected and it is concluded that there is a difference between the variances in the population.

- **Compare means by one-way ANOVA**

One-way ANOVA is an analysis of variance in which there is only one independent variable. It is used to test the hypothesis that several means are equal. This technique is an extension of the two-sample t-test.

- **Regression analysis**

Regression analysis is the analysis of how the typical value of the dependent variable changes when any one of the independent variables is varied, while the other independent variables are held fixed. The coefficient of determination ($R^2$) is the proportion of variability in a data set that is accounted for by the statistical model. For linear regression, the $R^2$ is the square of the sample correlation coefficient. The $R^2$ ranges from 0 to 1. Adjusted $R^2$ is a modification of $R^2$ that adjusts for the number of explanatory terms in a model. The adjusted $R^2$ can be negative, and will always be less than or equal to $R^2$. 

---

Appendix B: Statistical analyses 123
Appendix C: Data from latest Transport Investment Plan

Road 249 projects

<table>
<thead>
<tr>
<th>Road1</th>
<th>Phases</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bristanalys</td>
<td>Förstudie</td>
</tr>
<tr>
<td>Number of projects</td>
<td>42</td>
<td>70</td>
</tr>
<tr>
<td>Average forecasted cost per project (Million SEK)</td>
<td>108</td>
<td>139</td>
</tr>
<tr>
<td>Maximum forecasted cost (Million SEK)</td>
<td>392</td>
<td>773</td>
</tr>
<tr>
<td>Minimum forecasted cost (Million SEK)</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>Total forecasted cost (Million SEK)</td>
<td>4,518</td>
<td>9,755</td>
</tr>
</tbody>
</table>

Rail 46 projects

<table>
<thead>
<tr>
<th>Rail1</th>
<th>Phases</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Idéstudie</td>
<td>Förstudie</td>
</tr>
<tr>
<td>Number of projects</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>Average forecasted cost per project (Million SEK)</td>
<td>945</td>
<td>746</td>
</tr>
<tr>
<td>Maximum forecasted cost (Million SEK)</td>
<td>3,243</td>
<td>5,198</td>
</tr>
<tr>
<td>Minimum forecasted cost (Million SEK)</td>
<td>83</td>
<td>28</td>
</tr>
<tr>
<td>Total forecasted cost (Million SEK)</td>
<td>5,667</td>
<td>13,432</td>
</tr>
</tbody>
</table>

1 All costs are the costs at 50% confident level.

Appendix C: Data from latest Transport Investment Plan 124
References


