Appraisal of Transport Projects

Assessing Robustness in Decision making

Jensen, Anders Vestergaard

Publication date:
2012

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Jensen, A. V. (2012). *Appraisal of Transport Projects: Assessing Robustness in Decision making*. DTU Transport. PhD-2012-02

Appraisal of Transport Projects: Assessing Robustness in Decision Making

PhD Thesis

Anders Vestergaard Jensen

September 2012

DTU Transport
Department of Transport

Appraisal of Transport Projects:

Assessing Robustness in Decision Making

Anders Vestergaard Jensen

September 2012

DTU Transport
Department of Transport

PhD-2012-02
Department of Transport
Technical University of Denmark

Published by: DTU Transport
Bygningstorvet 116B
DK-2800 Kongens Lyngby
Denmark

Order at: http://www.transport.dtu.dk
Tel +45 45256500, Fax +45 45936412
transport@transport.dtu.dk

ISSN: 1601-9458 (Electronic version)
ISBN: 978-87-7327-238-1 (Electronic version)

ISSN: 1600-9592 (Printed version)
ISBN: 978-87-7327-239-8 (Printed version)

PREFACE

The following thesis completes my PhD study entitled "Appraisal of Transport Projects: Assessing Robustness in Decision Making". The study has been carried out at the Department of Transport at the Technical University of Denmark in the period from 2008 to 2012. The study has been partly financed by an institute scholarship, and has been supervised by Professor Steen Leleur.

The PhD study has been a fascinating and challenging journey both geographically and personally and it has taken me several places. At the conferences and the international PhD courses attended, I have met many inspiring people who in one way or another have contributed to this PhD study. Participation in various research projects during the study period has been learning and fruitful experiences, which have made the case studies possible and through collaboration with other research environments, they have given me new insight in the research field of appraisal.

Considering the many people who deserve gratitude for contributing to this thesis, one may wonder why there is only one author on the title page. I cannot thank everyone. Be as it may, I would like to thank my (current and former) colleagues at DTU Transport: Michael Bruhn Barfod for rewarding discussions concerning MCDA and decision conferencing, co-authorship in one of the papers and for good company on several conferences and intensive PhD courses; Sara Lise Jeppesen for inspiration and guidance to soft systems approaches; Inga Ambrasaite for co-authorship in one of the papers and for helpful feedback on this thesis; Kim Bang Salling for useful assistance with Monte Carlo Simulation. To my advisor Steen Leleur, I gratefully acknowledge his assistance and encouragement. Thanks Steen for your confidence and support since the beginning of my PhD studies.

I find it important to mention that attending the two COST Action IC0602 International Doctoral Schools: "Algorithmic Decision Theory: MCDA, Data Mining and Rough Sets", in April 2008, Troina, Italy and "Applying Decision Analysis to Real Problems" in April 2011, Manchester, UK. These two courses have been of great value to me, as the learning environment, and the lectures by and discussions with Professor Simon French and Professor Theodor Stewart have been very inspiring.

I dedicate this thesis to the three most important persons in my life, together they make it all worthwhile. My loving wife Ditte and our two children Petra and Aske, they embody my idea of happiness and joy.

Anders Vestergaard Jensen, June 2012

ABSTRACT

The evaluation of transport projects has traditionally been based on quantitative descriptions of selected impacts by the use of cost-benefit analysis (CBA). However, environmental impacts are now taking into consideration when decisions have to be made, regarding which large transport projects to implement. This is handled by using environmental impact assessment (EIA).

In most real-world policy situations there are many possible alternatives, many uncertainties, many stakeholders, and many consequences of interest. In spite of this, public decision-makers (DMs) have a responsibility to develop and implement policies that have the best chances of contributing to economic development, health, safety, and well-being of their constituencies. DMs are confronted with the difficult task of evaluating potential outcome and choosing policies to achieve the desired outcome in the presence of increasing complexity. Optimisation has been replaced with satisficing, where satisficing is defined as finding an acceptable or satisfactory solution to a problem instead of a socio-economical optimal solution.

This calls for new approaches of appraising transport projects that goes beyond what the classical CBA, seems suitable for. As a result of the above mentioned issues regarding conflict of interests, multiple impacts, participation of stakeholders, and sustainability, additional approaches for appraisal have been suggested. Multi-criteria decision analysis (MCDA) has been introduced in the transport planning system, where the appraisal method is an MCDA priority model, with a CBA nested within. However, some concern needs to be addressed: In CBA, the weights (that is, the monetary valuation) are determined on the basis of the best available evidence. In MCDA, weights can reflect evidence, expert opinion, or various preferences. However, defining weights on the basis of preferences introduces subjectivity in the analysis and in the extreme defies its purpose. Therefore it is essential to be able to cope with this introduced subjectivity in order to provide justifiable decision support. This can be dealt with by the use of sensitivity analysis and robustness measures. This also calls for an appraisal methodology which is transparent and able to include the different preferences among experts, the public, and stakeholders.

This PhD thesis has three main focuses concerning appraisal of transport projects:

- How to reconcile socio-economic and public acceptance to obtain decisions that are both rational and legitimate
- How to involve stakeholders in the appraisal
- How to apply robustness as mean to improve the appraisal

The examination of these main focuses has been two-folded: For each question first its relevance has been argued and afterwards been treated more specifically in regards to methodology and suggestions.

The main concern in the PhD study has been to apply robustness as a mean to improve the appraisal of transport projects. In order to accomplish this, sensitivity analysis methods and robustness measures have been elaborated, which examine the subjective part of the MCDA (in form of criteria weights) and its role in decision support making. For this purpose both deterministic and stochastic sensitivity analyses have been developed.

In addition, the focus has been formulating a framework which provides a transparent analysis that can be communicated to the DMs. The framework involves CBA for assessing the socio-economic part of the decision problem, EIA to include environmental issues, MCDA to embrace various and often conflicting criteria, and sensitivity analysis for taking into account the interests and preferences of different stakeholders. These various interests and preferences have been revealed by the use of decision conferencing, which engage the stakeholders and provide a common platform for understanding the decision problem. Leading up to this framework, this thesis has also examined the issues of how to provide decision support which both support decisions that are rational in a socio-economic manner as well as being able to legitimatise decisions for the public. Applications of the developed methods are demonstrated by three case studies which are described in four papers written during the PhD study.

The potential for transport appraisal demonstrated in this PhD study can be listed in the following main findings.

- To reconcile socio-economic analysis and public acceptance it can be recommended to widen the appraisal methodology.
- When widening the appraisal methodology, e.g. replacing CBA with EIA+CBA+MCDA+DC, there is a need to balance between the use of sophisticated modelling and emphasising transparency for all.
- The conduct of a DC with the participants representing different stakeholders is seen as a way of combining technical, analytical solutions with social group processes aiming for engaging the different stakeholders with consideration of their variability of their preferences.
- The different sensitivity analysis and robustness measures developed in the technical part of this PhD study are found to be relevant in supporting the DC group process, as robustness of a recommended solution is a major concern in the final steps of decision making.

As mentioned, the appraisal of transport projects is a complex issue involving conflict of various interests and this call for new approaches to the practice of appraisal. The

presented appraisal framework is a tool for analysing conflicting preferences, and it is believed that the efficient use of available information from CBA and EIA, combined with the MCDA approach and the robustness measures, establishes a strong decision aid tool for the practical evaluation of larger transport projects.

ABSTRACT IN DANISH (DANSK RESUMÉ)

Evalueringen af transportprojekter har traditionelt været baseret på overvejelser af kvantitative beskrivelser af udvalgte effekter ved hælp af cost-benefit analyse (CBA). Miljøpåvirkninger er nu også medtaget i betragtningerne, når beslutninger om gennemførelsen af større transportprojekter skal tages. Denne del bliver varetaget af vurdering af virkninger på miljøet (VVM).

Ved beslutninger om transportprojekter er der ofte adskillige alternativer, mange usikkerheder, mange interessenter og mange konsekvenser af interesse. På trods af dette har de offentlige beslutningstagere et ansvar for at udvikle og gennemføre en politik, der har den bedste chance for at bidrage til den økonomiske udvikling, sundhed, sikkerhed og trivsel for samfundet. De er konfronteret med den vanskelige opgave at vurdere mulige resultater og vælge mellem alternativer for at opnå de ønskede resultater i tilstedeværelse af stigende kompleksitet. Optimering er blevet erstattet med satisficering, som er defineret ved at finde en acceptabel eller tilfredsstillende løsning på et problem i stedet for en socio-økonomisk optimal løsning.

Dette kræver nye tilgange til vurdering af transportprojekter, som går ud over, hvad den klassiske CBA synes egnet til. Som følge af ovennævnte problemstillinger med transportvurdering er yderligere metoder til vurdering blevet foreslået for at kunne håndtere f.eks interessekonflikter, flere effekter, inddragelse af interessenter og bæredygtighed. Multi-kriterie beslutningsanalyse (MCDA) er blevet indført i transportplanlægningssystemet, hvor evalueringsmetode er en MCDA model med en CBA indlejret. Men det kræver at visse problemstillinger behandles; i CBA bestemmes de vægte (den monetære værdiansættelse) på grundlag af den bedst tilgængelige viden. I MCDA derimod kan vægte afspejle viden, ekspert bedømmelser eller forskellige præferencer. Men at definere vægte på grundlag af præferencer introducerer subjektivitet i analysen og i det ekstreme kan det trodse analysens formål. Det er vigtigt, at være i stand til at håndtere denne subjektivitet for at give et forsvarligt beslutningsgrundlag. Dette kan behandles ved brug af følsomhedsanalyse og robusthedsmål. Yderligere kræver en sådan tilgang også en vurderingsmetode, der er gennemsigtig og i stand til også at omfatte de forskellige præferencer der findes blandt eksperter, offentligheden og interessenter.

Denne ph.d.-afhandling har tre primære fokus områder omhandlende vurdering af transportprojekter:

- Hvordan kan man opnå transportpolitiske beslutninger, der forener samfundsøkonomisk forsvarlighed og offentlig accept?
- Hvordan kan interessenter involveres i en vurderingsproces af transportprojekter?

- Hvordan kan et mål for robusthed forbedre vurderingsprocessen?

Undersøgelsen af disse primære fokus har været tve-delt: For hvert spørgsmål er først dets relevans fremført, og bagefter behandlet mere specifikt med hensyn til metode og forslag.

Det primære fokus i ph.d.-studiet har været anvendelsen af robusthed som et middel til at forbedre vurderingen af transportprojekter. For at opnå dette, er der udarbejdet metoder til følsomhedsanalyse og robusthedsmål, der undersøger den subjektive del af MCDA (i form af kriterievægte) og dennes rolle i beslutningsprocessen. Til dette formål er både deterministiske og stokastiske følsomhedsanalyser blevet udviklet.

Fokus har desuden været på at formulere en struktur, der giver en gennemsigtig analyse, der er kommunikerbar. Strukturen indeholder: cost-benefit analyse til at vurdere den socio-økonomiske del af beslutningsproblemet; VVM til at vurdere miljøspørgsmål; MCDA til at behandle forskellige og ofte modstridende kriterier, og følsomhedsanalyse for at tage hensyn til interesserenter og deres præferencer. Disse forskellige præferencer er blevet afdækket ved anvendelse af beslutningskonferencer (DC), som engagerede interessenterne og skabte fælles platform for forståelse af beslutningsproblemet. Afhandlingen har ligeledes belyst spørgsmålet om hvorvidt man kan yde beslutningsstøtte der både kan understøtte beslutninger, der er rationelle i et socio-økonomisk perspektiv og som samtidig er i stand til at legitimatise beslutninger for offentligheden. Anvendelsen af de udviklede metoder er præsenteret i tre casestudier, der er beskrevet i fire artikler skrevet i forbindelse med ph.d.-studiet.

Potentialet for transportvurdering demonstreret i dette ph.d.-studie kan opgøres i de følgende resultater.

- For at opnå både god overordnet samfundsøkonomi og offentlig accept kan det anbefales at udvide vurderingsmetoden til også at omfatte miljømæssige og strategiske hensyn.

- Ved en sådan udvidelse er det vigtigt, at der lægges stor vægt på formidlingen af resultaterne herfra, og at metoden ikke bliver for kompleks. Dette skal sikre at metoden er transparent og at metodens anbefalinger kan accepteres.

- Ved anvendelse af beslutningskonferencer med deltagelse af forskellige interessenter er det muligt at kombinere tekniske, analytiske løsninger med en social interaktion, der sigter mod at engagere interessenterne.

- De udviklede mål for følsomhedsanalyse og robusthedsanalyse anses for at være relevante i at understøtte processen i beslutningskonferencer, idet robustheden af en anbefalet løsning er et vigtigt anliggende.

Som nævnt, er vurdering af transportprojekter et komplekst spørgsmål, der vedrører adskillige interesser, hvilket kræver nye tilgange til at praktisere projektvurdering. Den fremlagte vurderingsstruktur er et værktøj til at analysere modstridende præferencer.

En effektiv udnyttelse af tilgængelige oplysninger fra CBA og VVM, kombineret med et MCDA perspektiv og robusthedsforanstaltninger vil etablere et stærkt beslutnings-støtteværktøj for den praktiske evaluering af større transportprojekter.

"It is only with the heart that one can see rightly. What is essential is invisible to the eye."

Antoine de Saint-Exupéry, The Little Prince

TABLE OF CONTENTS

PART 1 INTRODUCTION

PART 2 DECISION SUPPORT

PART 3 ROBUSTNESS

PART 4 CONCLUSIONS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY OF SELECTED ACRONYMS

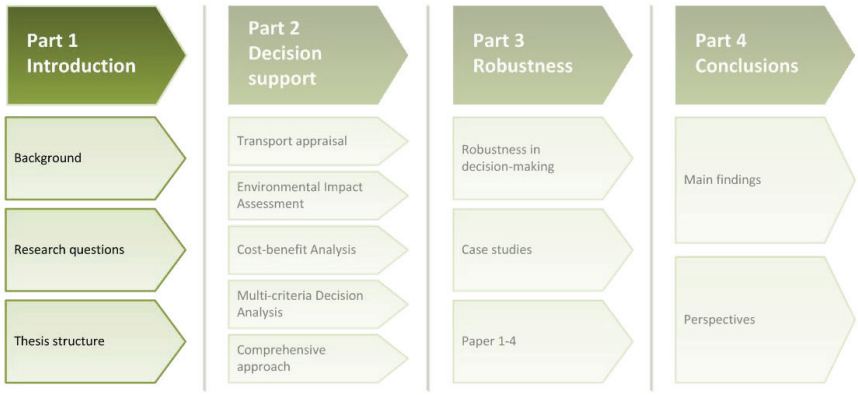
The acronyms found below are used throughout this thesis. They are presented by their full name when they are first used, but will later on be referred to by their acronym. In the list the acronyms are presented in alphabetical order. The list is intended as a help for the reader for refreshing the full name of an acronym after its introduction should occur.

Acronym	Full name
AHP	Analytic Hierarchy Process
BCR	Benefit Cost Ratio
CBA	Cost-Benefit Analysis
CI	Consistency Index
CR	Consistency Ratio
COSIMA	COmpoSIte Model for Assessment
CWs	Criteria Weights
DA	Decision Analysis
DC	Decision Conference
DM	Decision-Maker
DS	Decision Support
DSS	Decision Support System
EIA	Environmental Impact Assessment
EM	EcoMobility
IRR	Internal Rate of Return
MAVT	Multi-Attribute Value Theory
MAUT	Multi-Attribute Utility Theory
MCA	Multi-Criteria Analysis
MCDA	Multi-Criteria Decision Analysis
MCS	Monte Carlo Simulation
NPV	Net Present Value
RA	Robustness Analysis

Acronym	Full name
REMBRANDT	Ratio Estimations in Magnitudes or deci-Bells to Rate Alternatives which are Non-DominaTed
ROC	Rank Order Centroid
ROD	Rank Order Distribution
RP	Revealed Preference
RR	Rank Reciprocal
RS	Rank Sum
PVs	Performance Values
PNS	Post-Normal Science
SA	Sensitivity Analysis
SEA	Strategic Environmental Assessment
SIMDEC	risk SImulation and Multi-criteria decision analysis in combination for DECision support
SMART	Simple Multi-Attribute Rating Techniques
SMARTER	SMART Exploiting Ranks
SP	Stated Preference
TRR	Total Rate of Return
WSI	Weight Stability Interval
WTA	Willingness To Accept
WTP	Willingness To Pay

Part I

Introduction



1 INTRODUCTION

1.1 BACKGROUND

Decision-makers are confronted with the difficult problem of evaluating and choosing among different alternatives for transport projects[1]. In order to do this they include a variety of techniques and procedures that seek to predict and evaluate the consequences of these different alternatives. Commonly allocation of funds for transport projects have been based on cost-benefit analysis (CBA), and this appraisal methodology is widely recognised to be helpful for making good decision on which transport projects to fund. Basically it aims to identify which projects offer the best value for money – one of the core criteria for making decisions. The economic worthiness of a project is determined by calculating the Net Present Value (NPV) defined as the difference between discounted value of benefit and costs. If NPV of a project is positive, then the benefit exceeds the costs and the project is considered economically worthwhile. To rank projects, the Benefit-Cost Ration (BCR) defined as the projects NPV divided by its discounted costs (financed through government funding) is commonly used. The advocacy for the use of CBA for evaluating the qualities of transportation projects is so immense that virtually all western European countries and the United States have formalised CBA as tool for evaluating transport projects (Martinsen et al., 2010). However, the practical relevance of CBA does not always live up to its appeal in principle. One issue is, that there is not always agreement about what to include as benefits and costs, and therefore the concept of value for money is not always fully transparent. Another important problem is, that value for money is only one of many criteria for decision making, which often lead to disagreement about the relative importance of the results from CBA compared to other inputs in the decision making process (OECD, 2011). Environmental impact assessment (EIA) is one of the other important inputs to the decision making process. EIA is a systematic procedure for gathering information about the environmental impacts of a project or policy, however, EIA is not a comprehensive evaluation procedure, since it ignores non-environmental impacts and costs. Actually EIA is often an important input to the CBA. In many countries CBA and EIA are required for investments in infrastructure. Proposals have been suggested in the literature for an integrated appraisal approach. E.g. Archibugi (1989) stress the importance of comprehensive planning procedures and Morrissey et al. (2012) proposes a strategic project appraisal framework for ecologically sustainable urban infrastructure. However, there exists no formalised procedure of how to integrate CBA and EIA to provide a comprehensive appraisal consisting of all the information from these two methodologies. The output from the two evaluation procedures are often presented to the decision-makers (DMs) separately, even though there are overlap between them.

[1] The terms 'project', 'policy' and 'options' are used in this thesis with the same meaning, that is undertakings of different kind, size and purposes, to be built, implemented, organized, for a purpose, objective, goal or target

How is the knowledge from CBA then used in the decision making? Evidence have shown, that DMs may use subjective judgement in combining CBA and non-monetised impacts (e.g. included in EIA) when decisions are made (Martinsen et al., 2010). Nellthorp and Mackie (2000) developed a model which reproduces the pattern of decision making in the UK and from which the implicit weights or values on the variables in summary of the impacts of the project, can be derived from. Their findings include, among other things, that the NPV do not have significant influence on the decisions made, furthermore, they conclude that this implies that the DMs do not thrust the time savings benefits either, because they believe they are overestimated or overvalued. The conclusions Nellthorp and Mackie reaches are based on examining trunk road projects, and as they state, other transport projects can very well be even more complex and difficult to assess. Similar results of limited connections between CBA results and actual decisions can be found in McFadden (1975; 1976), Brent (1979), Nilsson (1991)(Odeck, 1996), Odeck (1996), Fridström and Elvik (1997), Nyborg (1998) and Sager and Ravlum (2005). Sager and Ravlum found that political DMs gather information and do not use it – they make decision first and look for relevant information afterwards and also collect and process a great deal of information that has little or no direct relevance to decisions. A similar critic has recently been presented in Denmark (Østergaard and Andersen, 2012). More recently Eliasson and Lundberg (2012) have examined how CBA results were used in the shaping of the Swedish National Transport Plan 2010-2012. They found some correlation between CBA results and investment decisions, however, with some interesting exceptions. They found that planners seems to use CBA primarily as a screening tool to avoid investments with negative net benefits, while it had little weight as a ranking criterion for investments with positive net benefits. Most importantly they concluded that planners' and politicians' selection of investments reveals, implicitly, preferences and considerations that may differ from those of the CBA methodology. These findings reveal a gap between the used appraisal methodologies and the actual decision making. The CBA alone does not clearly give a direction of which transport projects are to be financed or not. The decision making is influenced by other means as well. Furthermore, Damart and Roy (2009) question the value of CBA as a guide for DMs and conclude that CBA is not compatible with relevant and constructive debate. So a reasonable question would then be: Is it possible to provide decision support that describes the actual decisions made better than the ones used today?

It seems like decision making is affected primarily from three different perspectives illustrated in Figure 1-1; the economic part, including the CBA and the formal appraisal. A 'popular' part in which the perspective from a broad range of stakeholders and the public in general matters stated through e.g. consultation, focus groups and public opinions. A political aspect, in where the policy agenda and political processes play a role.

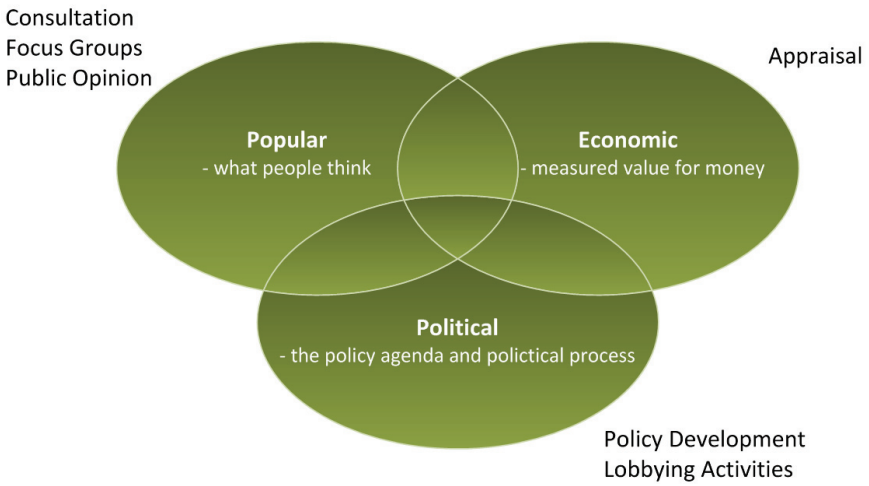


Figure 1-1 The appraisal context (Mackie, 2010)

Today's DMs have to legitimate the use of public resources and comply with stakeholders' acceptance of the choices made. Clearly the CBA could be questioned for not being sufficient for making decisions and therefore also, not possible to use to legitimatise the decisions made. This, and the methodology has been questioned for its applicability as means for public involvement and debate, promotes the search for an appraisal approach, which is more comprehensive and meets the needs of both the DMs and stakeholders. Multi-criteria decision analysis (MCDA) has been promoted to be used in the transport sector to overcome some of the shortcomings of CBA (Browne and Ryan, 2011; De Brucker et al., 2011; Haezendonck, 2007; Macharis et al., 2009; Tsamboulas and Mikroudis, 2000). In Figure 1-1 it is illustrated that appraisal is influenced from economic, popular and political aspects, however, on top of that one could add environment.

To provide solid and transparent decision support for decision making in the transport sector, and in any other sector, is a complex and difficult task. Adding the use of participative methods for revealing preferences among experts, citizens, politicians etc. makes the task even more difficult. This should be seen in the light, that there exist very thoroughly researched methodologies for conducting project evaluation. However, one setback of these, very complex and sophisticated approaches, is that they are not very informative about their foundations, and then there is the risk of the user not trusting the outcome of these models.

An important objective of a more comprehensive appraisal approach, which takes into account a broader perspective than that included in the CBA, and also being able to accommodate public participation, is robustness. Especially within the public sector where there is a clear need for legitimising the decisions made, robustness is very important. If the DMs can provide robust decisions which can be proven to be robust

for uncertainty in the input parameters and different preferences, there is a greater chance for public acceptance. Furthermore, a sensitivity analysis (SA), for assessing the robustness of the decision problem, can provide more insight about the sensitive parameters affecting the decision making.

To address the above concerns, an appraisal approach is required, which will enable allocation of transport project funds based on a transparent and comprehensive appraisal of both monetary and non-monetary impacts. This approach must be able to communicate the complexity and associated uncertainties of the decisions, and to allow for broad stakeholder participation, while integrating different aspects of the situation involved.

1.2 Research questions

The main aim of this thesis is to develop a framework for assessing the robustness of decisions, concerning transport projects which make use of both CBA and MCDA. Leading up to this framework, the thesis also examine the issues of how to provide decision support which both can support decisions that are rational in a socio-economic manner and as well is able to legitimatise the taken decisions for the public. As a response to the growing demand for participatory decision support the thesis also treats the question of how to involve stakeholders in the appraisal.

The central research questions are:

A. How to reconcile socio-economic and public acceptance to obtain decisions that are both rational and legitimate?

B. How to involve stakeholders in the appraisal?

C. How to apply robustness as a mean to improve the appraisal?

The examination of these research questions is two-fold: For each question first its relevance is argued and afterwards it is treated more specifically in regards to methodology and suggestions.

In order to answer these research questions, the following sub-questions will be addressed:

- What are the strengths and weakness of today's appraisal approach in transport planning?
- How are existing appraisal methodologies used today in transport planning?
- What is applicable decision support in transport planning?

- How can a process make use of quantitative information while giving proper weight to qualitative information?
- How can complicated sensitivity analysis be communicated?
- How can the discourse proceed in ways that are respectful of all viewpoints?

1.3 THE STRUCTURE OF THE THESIS

The thesis is presented in order to answer the research questions stated above and has been divided into four parts; introduction, decision support, robustness and conclusions. The first part seeks to introduce the problem and the thesis' objectives and structure. The second part is dedicated to describe and analyse decision making in a transport planning context, comprising of a study of the most important appraisal methodologies; environmental impact assessment (EIA) in chapter 3 and CBA in chapter 4 along with the proposed MCDA in chapter 5. The second part ends up with a description of comprehensive appraisal methodology exemplified with a case study in chapter 6. The third part deals with the theoretical considerations about the concept of sensitivity and robustness analysis in chapter 7 along the empirical findings of the PhD study documented by four papers in chapter 8. Finally, the last part concludes the thesis in chapter 9 by summarising the answers to the research question based on the results of the papers and the theoretical evaluation of the appropriateness of robustness measures of MCDA for decision making. Conclusions are drawn and ideas for further research are presented. Figure 1-2 gives an overview of the structure of the thesis.

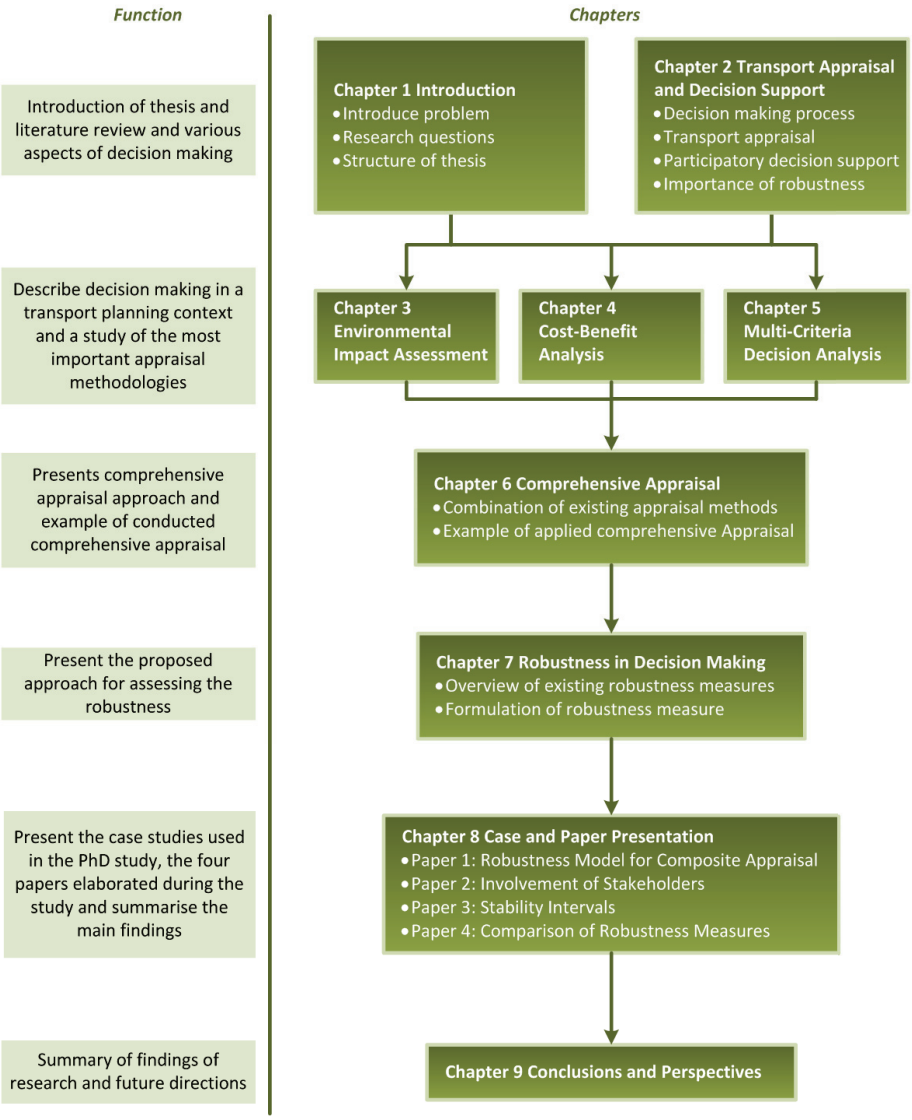


Figure 1-2 Flow chart summarising contents of thesis

In order to answer the research questions a number of different concepts of appraisal approach for transport planning is examined. An overview is given in Figure 1-3.

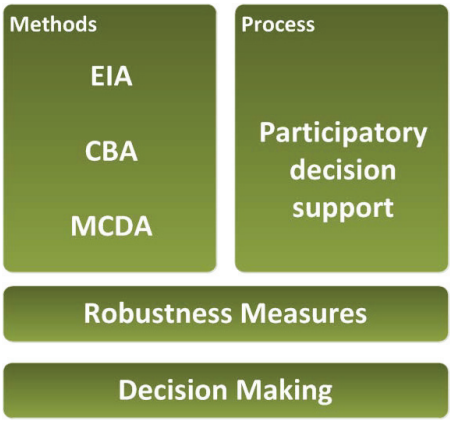


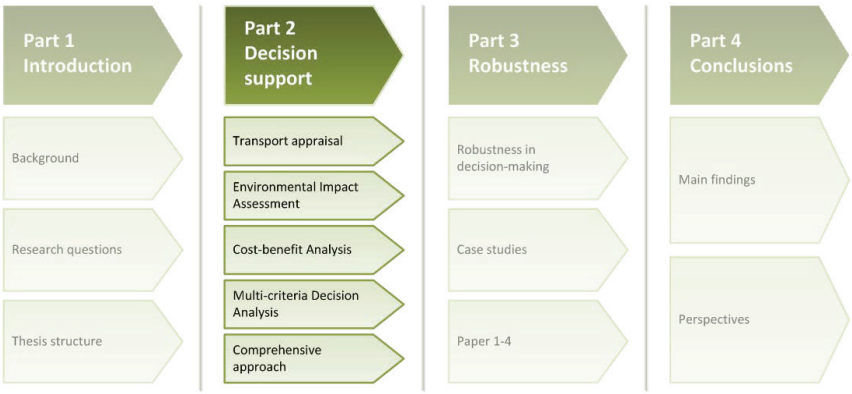
Figure 1-3 Concepts treated in this thesis as an appraisal approach for transport planning

Within the sub-level 'Methods' EIA, CBA and MCDA is processed. For 'process' primarily the concept of decision conference (DC) is treated. These two overlaying parts forms the basis for developing a robustness measure, to be used in the decision making.

This thesis sets out to evaluate a comprehensive approach to transport appraisal with the above mentioned aspects using EIA, CBA and MCDA. EIA and CBA are usually conducted after the project and its alternatives (including the no action option) are proposed. Therefore, this thesis will focus on treating decision making regarding transport projects, where a project has been proposed and a finite set of alternatives have been identified.

Part II

Appraisal Methods and Processes



2 TRANSPORT APPRAISAL AND DECISION SUPPORT

The purpose of this chapter is to set the scene for decision support and clarify some of the most important aspects in appraisal of transport projects. First the concept of decision support is described in a transport appraisal context. Following is a description of an alternative approach to appraisal (Post normal science), aspects of participatory decision support, how appraisal is conducted with regards to transport planning, and finally, the importance of robustness measure.

2.1 DECISION MAKING PROCESS

The overall obvious purpose of decision support (DS) is to improve decision making. More in detail, DS is deployed to help DMs to better deal with knowledge of various types and in various gradations, with the aim of fostering better performance (Holsapple, 2008). Providing decision support can be seen as a consultation process that attempts to focus DMs' attention on the important aspects of a problem. In order to support the DMs in making sound decisions, the process can be divided up into three phases (Figure 2-1).

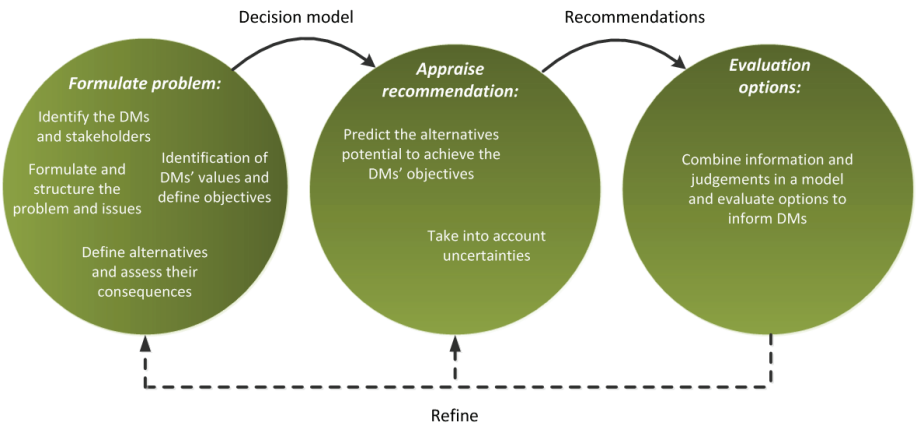


Figure 2-1 Different phases of the decision process (adapted from French et al., 2009)

The initial phase concerns the formulation of decision model that reflects the decision problem. This includes a definition of who the DMs and the stakeholders are and what their values are. Identification of DMs' values and a definition of the objectives, either reflecting DMs' values or some broader political objectives are needed. This initial phase can be crucial for determining the success of the DSS. An important issue in this

phase is to formulate the problem properly in order to solve the right problem – "we fail more often because we solve the wrong problem than because we get the wrong solution for the right problem" (Walker, 2000, p. 15). There is a tendency among analysts to focus on the used models rather than on the problem they are trying to solve. However, the models are merely the tools for the analyst; they are a mean to an end, not the end itself. "Fit the model to the problem, not the problem to the model" (Walker, 2000, p. 16). These aspects highlight the importance of formulating and structuring the problem in order to solve it. Formulating the problem is not straightforward and often it is an iterative creative process that requires much effort and thoughts as all potential alternatives (or strategies) seldom appear immediately.

The second phase consists of appraising the alternatives' abilities to achieve the objectives of the DMs. This is done by predicting the consequences of each possible alternative. This process often needs to consider uncertainties in these predictions and the aim of the phase is to guide the DMs towards an alternative that can meet the objectives.

The final phase is to appraise the recommendations. This appraisal involves interpreting the performance of the recommended alternatives as a real-world action. Output from the decision model is often a mathematically measure, which may not be meaningful for the DMs. Maybe more important is, that it is a model, meaning that it is a simplification of the real world and the DMs need to reflect on the recommendation and judge, if it makes sense in light of all the complexity the model might not have captured. The DMs need to assess whether the model has enlightened them enough to make their decision or further refinement of the models and analysis is necessary in order to get to a decision. This process continues until the DMs have enough confidence in the components, structure and values of the decision model describes the decision problem at stake (French et al., 2009).

As it is clear from Figure 2-1, a decision support system (DSS) is much more than just a decision model. It is important to notice that there cannot be one universal DSS that can assist any DM in any decision problem. A more or less tailored DSS is needed for any problem which again can vary from DM to DM.

In a transport related context, the decisions that are made potentially have a huge impact on the public in general, and therefore there are several issues to address. In principle better decision making processes should lead to better policy decisions. These decisions are by nature challenging, requiring a careful balance of the diverse public interest which is not easy to determine. The rapid rate of technological development, global interconnectedness and increasing reliance on private capital to drive economic development all contribute to make the identification of public interests ever more complex. In the face of complexity, elected officials cannot abdicate their responsibility for making policy decisions to analysts. But, if governments are going to produce coherent and effective policies, it is increasingly important that political DMs have the best advice and evidence available (OECD, 2009, p. 12). Good governance processes are important to manage policy complexity and to assist governments to support an

administrative culture that is well equipped to consider the consequences of policy options and to promote the economic, environmental, and social welfare goals that serve the interests of citizens (Ibid.).

In the decision making process a range of different actors are in play, more than it is possible to present here. Figure 2-2 illustrates a simplified picture of the players in a decision. There might exist more than one single DM, which is often the case for transport related decision problems, where decisions are the responsibility of a group of DMs. In order to conduct good decision making the DMs often consult with experts, engineers, accountants and other relevant subject experts. Thus many will contribute to the decision making process.

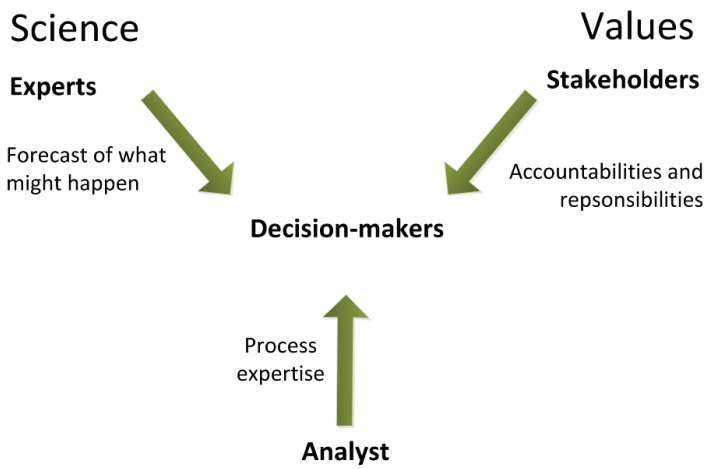


Figure 2-2 Players in the decision making process (French et al., 2009)

The DMs are the ones responsible for making the decision. Making and implementing a decision requires the DMs to hold appropriate responsibility, authority and accountability[2]. Individuals or groups are *responsible* for a decision if it their task to see to that the choice is made and implemented. Individuals or groups have the *authority* to make a decision if they have power over the resources needed to analyse and implement the choice. Individuals or groups are *accountable* for a decision if they are the ones who take the credit or blame for the decision process and for the choice that is made, how it is implemented and the final outcome of that choice.

The stakeholders share, or at least perceive that they share, the impacts from the decision and the DMs are accountable to some, but not necessarily all, of the stakeholders. They can be defined as those who are either both affected by and

[2] The terms responsibility and accountability are often used interchangeably, however here a distinction between the two terms is useful. In a summary distinction, "responsibility" focuses for the most part upon all of the elements of duty up to the point of decision, and "accountability" focuses for the most part upon all of the elements of duty after the decision is made.

affecting the problem, those who are acting on behalf of 'clients' and representing them, or those who do not have any direct control over resources viewed as relevant for the problem (e.g. future generations will be effected by today's policies) (Banville et al., 1998). The stakeholders have a claim, therefore, that their perceptions and values should be taken into account. The DMs are stakeholders, if only by virtue of their accountabilities, however, stakeholders are not necessarily DMs. In the transport sector, the government and its agencies have many stakeholders; these are e.g. the public, industry, users or political parties.

Experts often provide useful knowledge and advices about economic, marketing, scientific and other areas to the DMs. This knowledge is used by the DMs to better understand, formulate and assess the decision problem and furthermore used in modelling and forecasting the outcome of potential decisions; e.g. experts in transport modelling provide transport models which again can provide forecast of the impact of transport project on the transport flows in the system. Other experts may advise stakeholder groups, informing their perceptions and hence influencing the decision making.

In another dimension of the decision making process, the analysts develop and conduct the analyses, which draw together the empirical evidence and expert advice, to assess the likelihood of possible outcomes. The analysts work closely together with the DMs in order to clarify and elicit their uncertainties and values. When stakeholders also play a role in the decision, the analysts are also concerned with a synthesis of their values and perceptions. E.g. if decision conference (DC) is used for this, the analysts can have an important role in order to facilitate the process. Analysts provide process skills, helping structuring the analysis and interpret the conclusions, whereas experts provide information about the content of the decision. Because of this distinction analysts are also referred to as process experts. One important concern here is, that if the analysts or experts are too involved in the decision itself, there is the danger of introducing bias to the process and therefore it is often recommended that they are dissociated from the issues.

The quality of the advice provided e.g. from analysts, depends largely on having robust analytical processes that are integrated with the policy making apparatus and capable of communicating information to DMs at a time when it can have a positive influence. In political systems, which rely on the exercise of delegated powers, it is reasonable for citizens to expect, that policy decisions take into account a prior consideration of the anticipated impacts, and are informed by the views of stakeholders that are likely to be affected by these decisions (OECD, 2009).

Porter (1995) states that CBA was intended from the beginning as a strategy for limiting the play of politics in public investment decisions. This statement address one paradox of making more sophisticated appraisal methodologies for public investment decision, such as transport, on one side the politicians are in need of well-founded decision support tools, but on the other hand they do not want to be dictated by these tools (and probably should not). However, introducing more comprehensive decision

support tools, encompassing a wider range of impacts, the decision space wherein the politicians can make decision is narrowed down.

Sustainability raises a set of issues based on civil rights of current and future generations as well as respect for ecological systems. Understanding that human systems are complex, and will never be fully understood, have led to the development of approaches which favour adaptive behaviour and learning processes over optimal solutions. Closely linked to this view, is the acknowledgement of uncertainty of future events and their impacts on human and non-human systems. All these factors redefine the role of experts, the meaning of knowledge and how decision processes need to be designed to make more effective policy (de Montis et al., 2006). This leads us to a different perspective on appraisal framework and what is needed in order to provide a solid basis for sound justifiable decisions.

As a consequence of the above mentioned issues a decision making process must consider both the decisions and the individuals making them. The field of behavioural decision research provides such accounts. It entails three forms of research (Fischhoff, 2010, p. 724): (1) Normative, identifying the best possible choice, given the state of the world and DMs' values; (2) Descriptive, characterising how individuals make decisions, in terms comparable to the normative standard; and (3) Prescriptive, attempting to close the gap between the ideal (normative) and the reality (descriptive). Figure 2-3 provides an overview of the three forms of research.

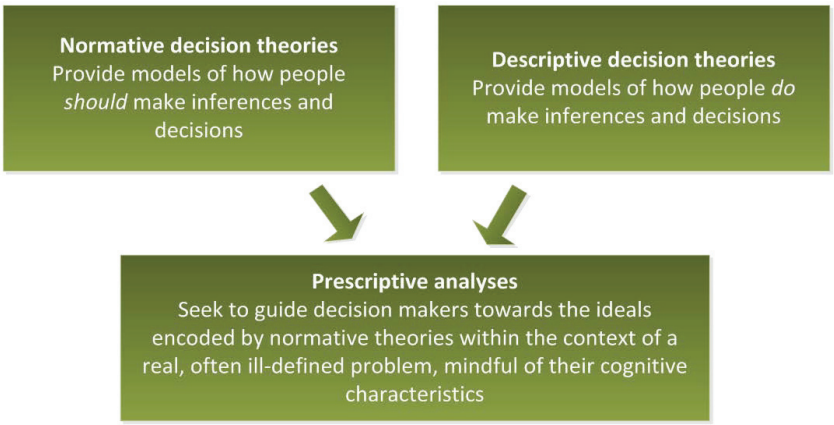


Figure 2-3 Different types of analysis and their particular roles (French et al., 2009, p. 58)

Normative models suggest how people should make decisions; descriptive models describe how they actually do. This thesis draws on these aspects of decision making to build prescriptive decision analysis and support. Prescriptive analysis seeks, among other things to guide DMs towards a decision by providing models that capture aspects

of the issues before them and of their beliefs and value judgements, while at the same time reflecting some rules of rationality (French et al., 2009). These models provide the DMs with perspectives on the issues, which give them an understanding of the issues; through this their judgements evolve and they can reach a decision. In communicating with the DMs and in eliciting their beliefs and value judgements an analyst needs to understand descriptive studies of how people make inferences and decide intuitively, because that is what they will do in a decision process. Thus, both normative and descriptive models contribute to prescriptive analysis.

Prescriptive analysis seeks to guide DMs when faced with a real problem of choice towards the rationality set in normative theories while being aware that their instinctive judgements may not satisfy the views and rules of normative decision theory. The key to forming this circle is to recognise that DMs are not unchanging beings. They can reflect; they can learn; they can understand. As a result, their preferences and beliefs can change. Prescriptive analyses guide this evolution of the DMs' perception and judgements. During prescriptive analyses these change and it is the purpose of the analysis that they should. People do not have clearly defined and stable preferences between all possible impacts in all plausible decision contexts. Rather, they construct them as they consider and deliberate on specific decision. The role of prescriptive decision analysis is to help the DMs construct their preferences and also get a closer insight into their uncertainties (Ibid.). Therefore the modelling process should be seen as a creative, dynamic, and cyclic process. The DMs' beliefs and preferences are assessed and modelled. The models are explored, leading to insights and a revision of the DMs' judgments, and hence revision of the models used. The analyst's role in this is to mediate and smooth this discussion by handling technical aspects. The process cycles until no new insight is found. This process has been termed requisite modelling by Phillips (1984). A model is requisite if its form and content are sufficient to solve the problem at hand. In other words, everything required to solve the problem is represented in the model or can be simulated by it. A requisite model is a simplified representation of a shared social reality.

Note here that the modelling process has a different character from that of scientific modelling. Scientists seek to understand the world in all its detail and they model behaviour in as fine detail as they can manage, building in more and more complexity to their models. For many decisions, more rough and ready models may be sufficient. Though, one must be careful not to use too simple a model to support a decision. Requisite modelling is about using a model that is fit for its purpose: sufficiently detailed to bring understanding to the DMs, but simple enough to still be transparent to the DMs (French et al., 2009). The concept of decision conference (DC), which is used in all four papers in this thesis for eliciting preferences and judgments among stakeholders, is a type of requisite modelling (Phillips and Bana, 2007). It is important to note here that the requisite model itself does not necessarily prescribe action, and it

is rarely descriptive, normative, optimal, or satisficing[3]. The focus in creating requisite models is on analysis.

Sensitivity and robustness analysis plays an important role in developing requisite models. Altering individual assessments or criteria sets allows disagreements between individuals to be examined to see if they make a difference in the final results. Changing one or more assessments such as pair wise comparisons or criteria weights over ranges of plausible values helps to identify crucial variables in the model (Phillips, 1984). One purpose of this type of analysis is to refine the decision model with the objective of obtaining a requisite model. This will be further elaborated in Chapter 7 and 8.

2.2 Post-normal science

Decision making in the transport sector is often characterised as one where facts are uncertain (the outcome of proposed alternatives), values are in dispute (lack of proper valuation techniques for non-market impacts and who's values count?), stakes are high (problems which affects the public, economy and the environment) and decisions are urgent (decisions have to be made now to cope with the problems of e.g. growth in traffic and its consequences). In such a situation we would be mistaken if we were to trust that true scientific facts determine the correct policy conclusions (Funtowicz and Ravetz, 1993). In order to deal better with two crucial aspects of decision making: uncertainty and value conflict, Funtowicz and Ravetz have developed a new epistemological framework called Post-Normal Science (PNS). The framework perspective challenges the assumption that in any dialogue, all valuations or numéraires should be reduced to a single, one-dimensional standard.

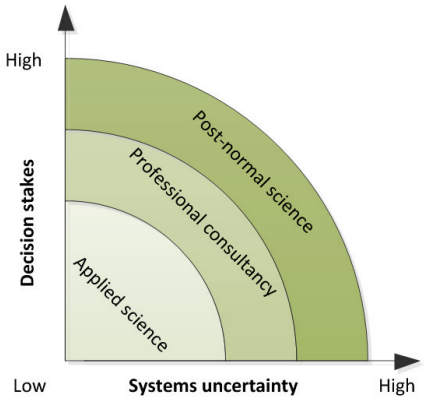


Figure 2-4 Post-normal science diagram

[3] Satisficing was given its meaning by Herbert A. Simon and can be defined as: "People will satisfice when they make a decision that satisfies and suffices for the purpose. This satisfactory sufficiency enables decision making which is good enough, rather than the absolute best – that which satisfices, while not ideal, will suffice to satisfy requirements." (Brown, 2004, p. 1241).

PNS can be described in relation to the more traditional problem-solving strategies by means of Figure 2-4. The PNS diagram has two axes, 'system uncertainty' and 'decision stakes'. The term 'system uncertainty' communicates the principle that the problem is not concerned with finding a particular fact, but with the understanding or management of an inherently complex reality. 'Decision stakes' covers all the various costs, benefits and value commitments that are involved in the problem through the stakeholders. Applied science is a problem-solving strategy where expertise is fully effective and is involved when both systems uncertainty and decision stakes are low. When either is medium, then the application of routine techniques is not enough. Skill, judgement, sometimes even courage, are required to adjust the "general knowledge" available to the "specific situation". This is called professional consultancy. Finally in cases where uncertainties are either of the epistemological or ethical type, or when decision stakes reflect conflicting purposes among stakeholders, and legitimate opposing views are openly used to challenge scientific arguments, we are in the realm of post-normal science (Funtowicz and Ravetz, 1993; Munda, 2008).

Whenever a problem is recognised as post-normal, even the routine research exercises take a new character. E.g. passenger transport has traditionally been seen as a straightforward engineering problem of maximising mobility – subject to the constraints of optimising costs and e.g. safety. Today transport policies are strongly influenced by several environmental aspects such as sustainability. Moreover, consumer demand for passenger transport depends directly on lifestyles. E.g. Americans want to be able to drive their car anywhere and Europeans want better conditions for bicycling. In terms of PNS, passenger transport presents the severe systematic uncertainties of climate change, combined with the crucial decision stakes in conceptions of the good life, along considerations of equity between people and generations. The entire population of passenger transport user (which is close to everyone) has effectively become an extended peer community. The success of sustainable transport technologies will depend on the effectiveness of the public's commitment to the values of the global environment (Funtowicz and Ravetz, 2008).

As PNS includes the idea of involvement of non-scientists, it appeals for a participatory approach to enhance either the quality, relevance or social acceptability of scientific knowledge (Hage et al., 2010).

Given that post-normal science concerns itself with issues that present diversity of perspectives (i.e. values held by the various stakeholders about the high stakes involved) and large uncertainties, its methodology aims at dealing with these specific issues. This methodology focuses on the management of both uncertainty and the quality assurance.

The participative multi-criteria process stems from critical and reflective thought (post-normal science) in contrast to a precisely defined process (normal science). It accepts that technical and social incommensurability by moving away from substantive rationality towards procedural rationality (Functowicz and Ravetz, 1992).

The Netherlands Environmental Assessment Agency made a transition from a technocratic model of science advising to the paradigm of post-normal science (Petersen et al., 2011).

2.3 Participatory decision support

In 1990 Healey predicted: "Citizens will demand more say in how the environment is managed and expect accountability from all agencies, i.e. democracy will be strongly asserted. But this will be in a context of substantial conflict between different interests arising from a more pluralist society, as well as conflicts between economic, social and environmental objectives. Resolving these conflicts in democratic ways will be a critical issue" (Healey, 1990 cited in Booth and Richardson, 2001, p. 142). In order to accommodate this participatory involvement, which facilitate dialogue between different actors and different preferences, to address and manage conflict, together with enabling people to express their concerns, requires planning processes (Booth and Richardson, 2001). A more philosophical viewpoint is given by Spagnoli (2003, p. 327): "A group of individuals is more intelligent than the sum of the individual intellects. Massive participation means massive criticism and this improves the quality of a proposal which can survive this massive criticism."

Only a few decades ago all societal decisions were taken by the authorities. Today many still are, however, an increase in public participation is occurring. In the 1980s new laws for public participation was introduced, however, in practice the participation only took place once the decision about a given project had been made and confirmed by the technical studies (Damart and Roy, 2009). Later in 1990s and 2000s directives specified measures for linking the socioeconomic evaluation phase of a project with the public debate phase (Ibid.).

How to reconcile rationality (socio-economically optimal) and legitimacy (public acceptance) to obtain decisions that are both rational and legitimate? When DMs choose among alternative investment project, they reveal their priorities, and these priorities must be perceived as legitimate. This calls for an appraisal methodology which is transparent and able to include the different preferences among experts, the public, and stakeholders. One approach for legitimizing the decisions made, is to involve various preferences and values in a participatory decision support context. However, the DMs should also be responsible to the economy of the society and make decisions which are economically sound and reasonable.

Participatory decision support could lead to an improvement of public confidence in government or agencies and provide information about the needs and preferences of society. By allowing the public involvement in the decision process, a better foundation for decisions can thereby be achieved, including allowing local knowledge to enter the decision process (Woltjer, 2009). Participatory decision making could even prevent delays that emerge because of, e.g. societal protest and keep expenditures down. All in all the process of public decision making have to deal with two difficulties: the difficulty of instituting a structured public debate that can serve to legitimize the decisions made,

and the difficulty of providing appropriate instruments for evaluating investment projects that are transparent for all (Damart and Roy, 2009). All though all these promising aspects of participatory decision making, it is important to state that participation is only relevant, if it somehow enhances the quality of decision making. In other words there are situations where participation is not beneficial. Enthusiasm on behalf of researchers for new approaches for participation needs to be qualified by awareness of the difficulties in applying techniques in different contexts and their potential for both failure and manipulation (de Montis et al., 2006).

There has been a powerful argument for shifting from the sole focus on outcome towards the quality of decision processes. In particular, multi-criteria assessment has become a more widely used method, offering the potential integrated assessment of local, national and international policies and as a means for combining different perspectives associated with sustainability goals (de Montis et al., 2006). Maybe one should not measure the success of decision support in looking at which alternatives are implemented, but more in the process of the decision analysis and if the decision process have led to any reshaping of the alternatives or even introducing new ones.

The only objective of a collaborative action involving community stakeholders should be to justify a decision that deviates from the optimum approximation (given by the CBA), which served as the starting point for public debate (Damart and Roy, 2009). In this respect the CBA is relevant because it forms the starting point of the debate and by involving stakeholders, the DMs can make a decision to select not the most optimal solution, from a socio-economic view point, but maybe another solution and at the same time demonstrate responsibility. In order to do this the stakeholders has to acknowledge the CBA as a valid starting point by admitting that, in spite of its imperfection[4], it produces a good approximation of an objective optimum. One approach for involving stakeholders is to use MCDA. An MCDA process can be seen as frame for dialogue, not to reach consensus but to recognize other parties' point of view and achieve a better understanding of the problem at hand.

The involvement of members of the public and/or stakeholders improves decisions in different aspects (Woltjer, 2009):

- Improve the use of knowledge and information and mobilise innovative ideas from participants.
- Help to save time and lead to a shorter decision making process over the long term.
- Help to save development costs at later stages of the decision making process.
- Will improve public support and acceptance, and will help to construct consensus and partnership in anticipation of possible conflicts.

[4] See chapter 4

Expert knowledge (from technical and scientific specialists) is needed to inform policy making, to give insights into fields where DMs lack sufficient knowledge. However, in many cases there is a huge amount of information available which could be difficult to handle. It might become challenging to differentiate between crucial or important issues and unimportant or even wrong or erroneous arguments. Qualitative approaches can help to filter and structure information or arguments. An aspect of representation is how information is delivered through the interaction of science and policy. The role of expert and 'objective' scientific information to inform policy on uncertainty and complex decision problems has been raised (de Montis et al., 2006). But not only interaction from expert knowledge, technical, and scientific specialists can be actors in the decision making process. Also, as mentioned earlier, the public is an important actor together with government agencies and appraisal specialists. All these different actors interact with the DMs and with each other e.g. government agencies might ask for scientific evaluation of a transport project regarding a specific impact (Figure 2-5)

Participation will depend on the project to be appraised. With highly technical complicated projects, the aim of participatory process should be towards improving the use of knowledge and new ideas. Whereas in situations dominated by process complexity, aim should be towards support and acceptance (Woltjer, 2009).

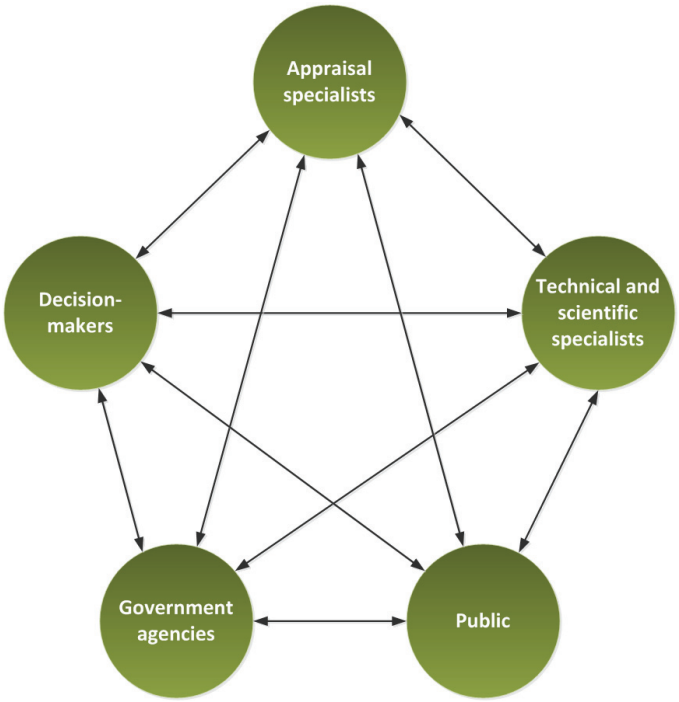


Figure 2-5 Different potential actors in the decision making process (adapted from Lawrence, 2007)

The timing of when to involve stakeholders must also be considered (Figure 2-6). At the low end of participation the stakeholders are invited relatively late in the decision making process. Not much is left at this point for any potential change of course, even if it were highly desired by the stakeholders. One purpose of participation at this stage could be to relate and integrate existing information. In contrast, at the high end of participation, the stakeholders have the opportunity to shape the process from the very beginning. Thereby they provide support in exploring innovative solutions (van den Belt, 2004).

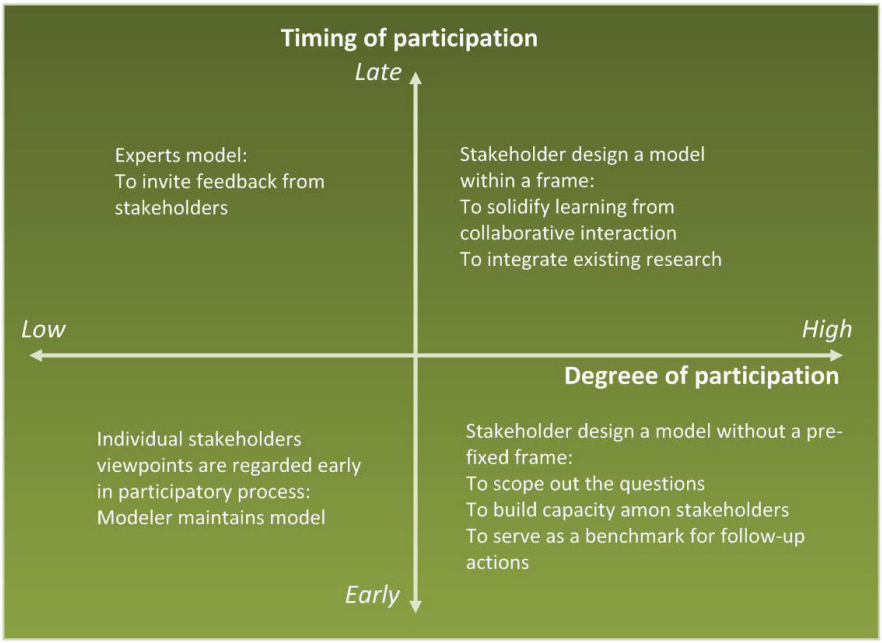


Figure 2-6 Participatory modelling (van den Belt, 2004)

In the papers presented in this thesis the concept of decision conferencing (DC) has been used for assessing knowledge and preferences from stakeholders and experts. DC is just one out of several methods for involving stakeholders in the decision process.

2.4 Appraisal in transport planning

This section introduces how appraisal is conducted with regards to transport planning. Governments have to insure that public funds are spent on activities that provide the greatest benefit to society, and that they are spent in the most efficient way. When decisions are taken at the appraisal stage, they can, and mostly do, affect the whole life cycle of new policies, programmes and projects.

During the 1970s a growing concern about 'non-material' costs and benefits arose. When evaluations of environmental costs took place, they often seemed to be trivial in comparison with the material benefits obtained from the projects; this was especially true for road appraisal and airport construction (Pearce and Nash, 1981). The political climate suggested that the environment should have figured more, not less, prominently. In other words the CBA seemed to be out of step with political pressures and changing values. This inconsistency was enough to cause a major shift in attitude towards its worth. Alternative techniques was then emerging, ranging from disaggregated CBAs, where aggregation was only permitted for some costs and benefits, the others being regarded as non-additive, to virtually physical statements of impact in which matrices of effects were presented to the DMs for judgemental assessment. Apart from planning balance-sheets and EIAs, also goals and achievement matrices as well as other multiple criteria techniques emerged (Ibid.).

Several different appraisal methods are used today to appraise transport projects (Bristow and Nellthorp, 2000), which sets focus on the following main objectives:

- Economic efficiency, measured by the result of a social cost benefit analysis (CBA).
- Minimising environmental damage, by using environmental impact assessment (EIA) methods.
- Other objectives concerned e.g. with equity, with accessibility, with long-term cash-flow implications for the state or with achievement of regional development policies.

In Europe there is a general level of consensus about which direct impacts that should be included in a CBA. But when it comes to impacts where no market values are available, such as time and accident cost, values differ (Bristow and Nellthorp, 2000; OECD, 2011). This introduces a challenge for the planners when appraising cross-border transport projects.

Because of the fast growth of mobility in the society an increasing need for expanding transport infrastructure is present. This is in conflict with the overall objective of sustaining a healthy environment and fostering liveable cities. In the light of these conflicts it is not surprising that transport infrastructure planning is so cumbersome. Some current trends add to this complexity:

- *The huge interests involved*, which are more and more conflicting. Infrastructure affects central interests – such as property, health, safety and well-being – in a both negative and positive way. Opposing impacts relate to: the physical presence of infrastructure, which necessitates space; the construction of infrastructure that results in disturbance of the surrounding area and impacts because of the use (traffic) – this relates to impact on air quality, noise, nature, safety etc. On the positive side infrastructure has: a function in connecting

diverse functions such as housing, working, recreation etc.; and it enables car mobility, which is seen as an important prerequisite for individual well-being and development by many in modern society.

- *The growing scarcity of space.* The conflict between the interests mentioned before and the complexity of the planning task is intensified because of the competing claims for space by the various spatial functions in densely populated areas.
- *The changing roles of government and other parties.* Traditionally the national government plays a leading role in the planning of transport infrastructure. However, this role of national government is subject to changes. This is related to the controversial character of national infrastructure expressed clearly in the so-called 'Not in my backyard' (NIMBY) phenomenon. The collective interest of new infrastructure on national or regional scale seems to conflict with the individual interests of people. As a consequence, national governments have invested much in open, participative planning processes. It cannot assume a solitary position anymore. This relates to developments as: the requirement of stronger cooperation within national government with the other departments; the requirement of more participation by the public in the planning process.

These trends underline the complexity of conducting appraisal of transport projects. Taking multiple impacts into account, when choosing among transport infrastructure investments is a two-fold problem. On the one hand, the limited resources available must be used in the best way possible. Thus, the rational appraisal of transport projects should include all the multiple impacts of the projects, even those that are difficult to evaluate. On the other hand, when choosing among alternative transport projects, the DMs reveal their priority they assigned to the different stakes, and these priorities must be perceived legitimate. Consequently, DMs must spend the limited resources with special care and make the most acceptable decisions possible (Damart and Roy, 2009).

In De Brucker and Verbeke (2007) they discuss two paradigms that can be adopted when assessing transport policy measures and large infrastructure projects, the neo-classical and the neo-institutional[5] approach. The neo-classical approach assumes that the individual engages in a comparison between marginal benefits (as measures by e.g. the willingness to pay) and marginal costs (as measures by opportunity cost). In principle, full rationality and complete information prevail. This is the classical cost-benefit approach. The neo-institutional approach can be considered as 'decision procedures', that is, procedures or sets of rules that enables a group or society to transform individual preferences into collective preferences. Institutions can be formal or informal, where the former include laws and regulations, the latter include social customs and norms. Because of the above changing roles of government and other actors in transport appraisal it seems like the neo-institutional approach is becoming

[5] It should be emphasised that an institution, as a set of rules, is not an organisation but rules based on values and experience

more and more relevant. As Baarsma and Lambooy (2005, p. 473) states "In principle, if the policy goal is already clearly defined, it is possible to complement the monetary approach with the new institutional one as a basis for political decision making".

Any appraisal requires criteria for choosing between alternatives. Different criteria may entail trade-offs, such as between cost and quality or performance. Cost-benefit analysis uses money values as weights, because they express people's willingness to pay (WTP) or willingness to accept compensation (WTA)[6]. This produces the important characteristic that benefit and cost can be directly compared, and specific actions can be compared with doing nothing.

Appraisal is influenced by the asset ownership and regulatory regime. Road and rail projects will generally be seen as in the public domain, and therefore that social appraisal has legitimacy (Mackie, 2010). In France public legislators have tried to introduce an institutional framework designed to facilitate the participation of stakeholders in the transportation decision making process. This illustrates the difficulty of striking the right balance between the expert knowledge produced by CBA methods and the perspectives of various stakeholders (Damart and Roy, 2009). Furthermore, the even rising concern regarding sustainability issues (Table 2-1) puts even more pressure on the appraisal of transport projects.

Table 2-1 Sustainability issues (adapted from Joumard and Gudmundsson, 2010)

Economic	Social	Environmental
Affordability		Pollution prevention
Resources efficiency	Equity	Climate protection
Cost internalisation	Human health	Biodiversity
Trade and business activity	Education	Precautionary
Employment	Community	Action
Productivity	Quality of life	Avoidance of irreversibility
	Public participation	Habitat preservation
Tax burden		Aesthetics

The aim of e.g. sustainable development puts special demands on valuation and appraisal methods.

As a result of above mentioned complex issues with transport appraisal regarding e.g. conflict of interests, multiple impacts, neo-institutional approach, participation of stakeholders, and sustainability, additional approaches for appraisal has been suggested. MCDA has been proposed to overcome some of these issues (OECD, 2011). CBA and MCDA can be defined as below.

[6] The concepts of WTP and WTA are elaborated further in section 4.2

Cost-benefit analysis, a microeconomic framework which (Bristow and Nellthorp, 2000):

- Seeks to include all significant impacts to all sections of the society.
- Uses monetary values to express measured impacts as a total money amount, based on consumers' preferences.
- Seeks to avoid double counting of the benefits in different economic markets.

Multi-criteria analysis frameworks (Ibid.):

- Are objectives-led.
- Reflect the achievement of these objectives in a corresponding set of criteria.
- Allows weights to be attached to the criteria, leading to a total score for each project and hence a ranking of projects in the pool.

In most real-world policy situations there are many possible alternatives, many uncertainties, many stakeholders and many consequences of interest (Walker, 2000). In spite of this, public DMs have a responsibility to develop and implement policies that have the best chance of contributing to the economic development, health, safety and well-being of their constituencies[7]. DMs are confronted with the difficult task of evaluating potential outcomes and choosing policies to achieve the desired outcomes in the presence of increasing complexity. Optimization has been replaced with satisficing, where satisficing is defined as finding an acceptable or satisfactory solution to a problem instead of an optimal solution (Walker, 2000). This calls for new approaches to appraising transport projects, which goes beyond what the classical CBA and MCDA seems suitable for. In the Netherlands they have introduced MCDA in the transport planning system, the appraisal method used here is a nominally a multi-criteria (MCDA) priority model, nested within which is a CBA. However, some concern needs to be addressed; In CBA, the weights (that is, the monetary valuation) are determined on the basis of the best available evidence. In MCDA, weights can reflect evidence, expert opinion, or maybe even policy preferences. The MCDA approach is useful when evidence is poor or absent. However, defining weights on the basis of policy preferences introduces subjectivity into the analysis and in the extreme defies its purpose. Therefore it is essential to be able to cope with this introduced subjectivity in order to provide justifiable decision support. This can be dealt with by the use of sensitivity analysis and robustness measures.

2.5 Importance of robustness in decision making

In any quantitative analysis the output depends upon the input. The significance of this dependence will vary from input to input and one is generally more confident in some

[7] Note here that DMs often also consider their possibilities for being re-elected in their constituency

inputs than in others. Because of this, sensitivity analysis is always recommended when appraising projects.

In order to evaluate different transport projects, their impacts need to be assessed. The magnitude of impacts e.g. travel time, environmental emissions, noise etc. depends on models which seek to model the outcome of the evaluated projects. However, these models, like any other model, are not precise. Furthermore, the impacts also have to be projected into the future by the use of expected economic growth factors, growth in travel demand and other socioeconomic factors, which also incorporate uncertainty into the assessments. These uncertainties have for long time been subjective for several studies and publications, and there is common recognition of the importance of this. An example of this is the UNITE project (Salling, 2012), which analyses and discusses forecast accuracy based on extensive empirical study, and uses this for recommendations and guidelines for model use. However, the extension of the appraisal approach to also include non-monetised impacts by the use of MCDA and participative processes, calls for additional methods for dealing with uncertainty.

Sensitivity analysis and robustness measures are based on the uncertainties which lie in the input to the appraisal model. Therefore it is beneficial to define uncertainty here. Many definitions for uncertainty are available, however, in the context of practical MCDA applications the one presented by Stewart (2005, p. 446) is appropriate:

> *Uncertainty implies that in a certain situation a person does not possess the information which quantitatively and qualitatively is appropriated to describe, prescribe or predict deterministically and numerically a system, its behaviour or other characteristics.*

The definition states that when e.g. asking people at a DC about the performance of a certain alternative with respect to a specific criterion, the persons might not possess the knowledge or information in order to answer this question without uncertainty. Again there could be several different sources for this uncertainty. Stewart (2005) describes two different sources for uncertainty: internal uncertainty, relating to the process of problem structuring and analysis, and external uncertainty, regarding the nature of the environment and thereby the consequences of a particular course of action, which might be outside the control of the DMs. However, both internal and external uncertainties can be treated in much the same manner by conducting sensitivity analysis and assessing the robustness. Although complex mathematical models have been elaborated to deal with these uncertainties the approach chosen in this thesis is more pragmatic. This approach is motivated by the need for real-world decision analysis, which is transparent and that in any MCDA approach it is vital that the process is fully understandable for all participants.

These arguments are likewise important for conducting a public appraisal, including transport appraisal where sensitivity analysis could be extremely useful in the stage where the transport projects are discussed with stakeholders and public opinion, since

new values, suggestions and amendments can be incorporated into the model and a solution be known in seconds.

From above it seems reasonable to state the question on how to apply robustness as a mean to improve the appraisal? This question will be dealt with in chapter 7. However, first a more thoroughly presentation and discussion of existing appraisal methods are presented along with examination of how these methods can be combined in a comprehensive appraisal approach.

3 ENVIRONMENTAL IMPACT ASSESSMENT[8]

The purpose of this chapter is to present environment impact assessment (EIA) and the way it influences the treatment of my research question. Here it can especially be noted that EIA aims to help and manage conflicts by involving different stakeholders representing various preferences. The chapter will deal with research questions A and B concerning how to treat different stakeholder views in the appraisal, and how to obtain decisions that are both rational and legitimate. Furthermore, the purpose of this chapter is to illuminate the strengths and weaknesses of main methods used today in transport appraisal.

3.1 INTRODUCTION

EIA is systematic procedure for collecting information about the environmental impacts of a project or policy. Tracing and measuring environmental impacts can be done with EIA. EIA measures the various environmental changes arising from the project, leaving the impact measured in physical units which will vary from one impact to another (Pearce et al., 2006). EIA is not a comprehensive evaluation procedure, since it ignores non-environmental impacts and costs. However, EIA is an essential part of any evaluation procedure and also function as an important input to CBA. In CBA other impacts are covered and an attempt is made to put money values on the environmental impacts. Most EIAs try to assess the significance of the environmental impacts, some give the impacts a score, and even some EIAs goes even further and also assess a weight to the impacts (Ibid.). Unlike CBA, EIA does not have any aggregated decision rule, but is often used for minimising the environmental impacts or address if some environmental threshold limits are exceeded. In general:

- EIA is an essential input to any decision making procedure.
- Impacts may be scored and weighted, or they become inputs to CBA.
- EIA would generally look for ways to minimise environmental impacts without changing (significantly) the benefits or costs of the assessed project.

The procedural origins of EIA are rooted in rational planning theory, developed in the mid-1950s, and were widely discussed and propagated in the late 1960s and early 1970s (Fischer, 2003).

This should be seen as increasing concerns in developed economies about the impact of human activities on the environment and on human health, which led to the development of the concept of EIA in the 1960s. Later in that decade it was adopted as a legally based decision-support instrument to assess the environmental implications of proposed development.

[8] EIA is here described before CBA, even though CBA has been used priory to EIA, which was first introduced in the early 1970s. The reason for this order is to be found in the proposed modeling framework presented in this thesis.

In France the act of 10 July 1976 sat a new agenda for appraisal of transport infrastructure. It required environmental studies to be completed before making any major transport infrastructure decision and this was completely new at that time. However, implementation is somewhat limited and, in practice, economic studies predominate (Damart and Roy, 2009).

An overlaying appraisal method is Strategic environmental assessment (SEA) which is a decision making support instrument for the formulation of sustainable spatial and sector policies, plans and programmes, aiming to ensure an appropriate consideration of the environment (Fischer, 2003). Where EIA is conducted for single projects, SEA is for larger programmes, which can contain several projects. The goal is to evaluate possible synergies between individual policies and projects and to evaluate alternatives in a more comprehensive manner. A SEA is more likely than an EIA to consider issues such as: is the policy or project needed at all; and, if it is, which alternative options are available? What is the preferred option? And SEA leads to a strategy for action rather than represent an end as the EIA does. This makes the SEA more pro-active than EIA which tends to be reactive (Pearce et al., 2006) and SEA have a broader focus and lower level of detail than the EIA.

The International Association for Impact Assessment (IAIA), one of the leading global authorities on best practice in EIA have defined EIA as (Senécal et al., 1999):

> *The process of identifying, predicting, evaluating and mitigating the biophysical, social, and other relevant effects of development proposals prior to major decisions being taken and commitments made.*

As with regard to decision making EIA is a mean to aid decision making through focus about potential environmental consequences of examined projects. Thus the underlying purpose of EIA is to allow proponents, managers and DMs to enhance the benefits and to minimise the environmental costs of development projects (Senécal et al., 1999), in more specific terms EIA seeks to

- To ensure that environmental considerations are explicitly addressed and incorporated into the development decision making process;
- To anticipate and avoid, minimise or offset the adverse significant biophysical, social and other relevant effects of development proposals;
- To protect the productivity and capacity of natural systems and the ecological processes which maintain their functions; and
- To promote development that is sustainable and optimizes resource use and management opportunities.

The intention of the EIA-procedure is to influence decision making by necessitating that the information set forth in EIA reports are taken into consideration. The EIA-procedure also provides a mechanism for ensuring the participation of potentially affected persons in the decision making procedure. (COWI, 2009)

3.2 Theory

IAIA and the UK Institute of Environmental Assessment (IEA) have set up a set of basic principles which apply to all stages of EIA; they also apply to Strategic Environmental Assessment (SEA) of policies, plans and programs (see Table 3-1).

Table 3-1 Basic principles for EIA. EIA should comply with these principles (Senécal et al., 1999)

Environmental impact assessment should be	
Purposive - the process should inform decision making and result in appropriate levels of environmental protection and community well-being.	**Adaptive** - the process should be adjusted to the realities, issues and circumstances of the proposals under review without compromising the integrity of the process, and be iterative, incorporating lessons learned throughout the proposal's life cycle.
Rigorous - the process should apply "best practicable" science, employing methodologies and techniques appropriate to address the problems being investigated.	**Participative** - the process should provide appropriate opportunities to inform and involve the interested and affected publics, and their inputs and concerns should be addressed explicitly in the documentation and decision making.
Practical - the process should result in information and outputs which assist with problem solving and are acceptable to and able to be implemented by proponents.	**Interdisciplinary** - the process should ensure that the appropriate techniques and experts in the relevant bio-physical and socio-economic disciplines are employed, including use of traditional knowledge as relevant.
Relevant - the process should provide sufficient, reliable and usable information for development planning and decision making.	**Credible** - the process should be carried out with professionalism, rigor, fairness, objectivity, impartiality and balance, and be subject to independent checks and verification.
Cost-effective - the process should achieve the objectives of EIA within the limits of available information, time, resources and methodology.	**Integrated** - the process should address the interrelationships of social, economic and biophysical aspects.
Efficient - the process should impose the minimum cost burdens in terms of time and finance on proponents and participants consistent with meeting accepted requirements and objectives of EIA.	**Transparent** - the process should have clear, easily understood requirements for EIA content; ensure public access to information; identify the factors that are to be taken into account in decision making; and acknowledge limitations and difficulties.
Focused - the process should concentrate on significant environmental effects and key issues; i.e., the matters that need to be taken into account in making decisions.	**Systematic** - the process should result in full consideration of all relevant information on the affected environment, of proposed alternatives and their impacts, and of the measures necessary to monitor and investigate residual effects.

The principles in Table 3-1 set out a high standard for conducting EIA and all of the principles would certainly ensure that the EIA would be highly useful in the decision process. Of special interest regarding this thesis are the principles of practicality,

participative and transparency. The practical principle ensures that the process ought to be able to be implemented and acceptable by all stakeholders. One way of achieving this is to address and handle the uncertainties, which undoubtedly will be present. The participative principle states, that it should be possible for all stakeholders to provide input to the decision process. This calls for procedures to handle various stakeholder inputs, which could be conflicting with each other and be challenging to incorporate in the decision process. The transparent principle states that clear guidelines should characterise EIA, and full public access to all information. Furthermore, the factors involved in the decision process should be identified, but at the same time limitations and difficulties should be acknowledged.

3.3 The stages of EIA

Figure 3-1 summarises the main EIA procedures that will be followed in the assessment of environmental impacts.

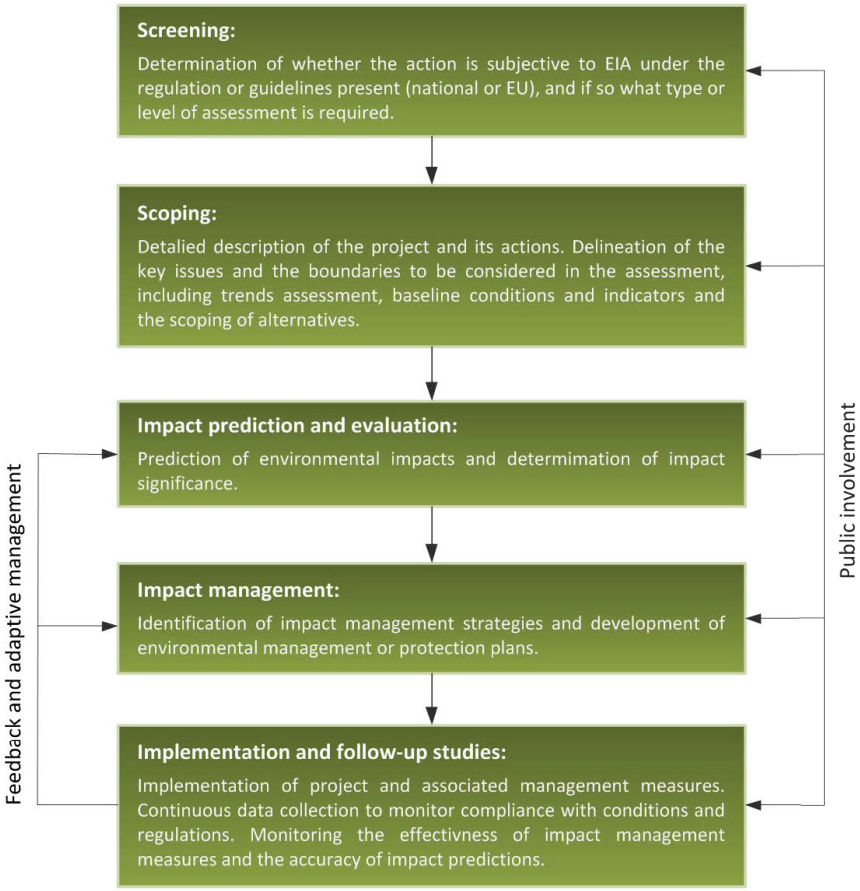


Figure 3-1 Generic environmental impact assessment process (Noble, 2011)

Screening: The purpose of screening is to ensure that only the necessary assessments are carried out and simply refers to narrowing down the application of EIA to the projects that requires assessment, either because of perceived (significant) environmental impacts or specific regulations.

Scoping: After the screening, scoping is an essential first step in the assessment of a component. In this step of the EIA it is important to bring attention to those impacts that will make a difference to decisions about the project. In this step the important issues and parameters that should be addressed are determined. In other words, scoping is about limiting the amount of information to be gathered, in order to get to a manageable level. This implies that what is being studied in various EIAs can vary quit significant from project to project. Although typical components are included such as surface and ground water quality, habitant conditions, components of the atmospheric environment along with physical geography parameters as vegetation, landforms, soil and geological properties. What to include in the EIA is also significantly influenced by the public perception and values (Noble, 2011). The latter can happen as a consequence of the participation which EIA impose.

Impact prediction and evaluation: The likely impacts of a project should be considered for all environmental components and the prediction is fundamental to EIA and should include assessment of (Morris and Therivel, 2001):

- Direct/primary impacts – direct result of a development.
- Indirect /secondary impacts – lead-offs from the direct impacts, not necessarily in the same area as where the direct impacts occurs.
- Cumulative impacts – accrue over time and space from a number of developments.

In order to perform an impact prediction it requires (1) sound understanding of the nature of the proposed examining; (2) knowledge of the outcomes of similar projects; (3) knowledge of past, existing or approved projects which cause interactive or cumulative impacts with the project being assessed; and (4) information about environmental receptors and how they might respond to change. When all the relevant impacts have been predicted it is necessary to assess the impact significance. The impact significance is the product of an impact's characteristics – magnitude and extent in space and time. Often this step is poorly addressed in EIA, mainly because this step is one of the most difficult ones in the EIA process (Morris and Therivel, 2001). Predicting the direct impacts are usually relatively easy, but the indirect and cumulative impacts can be much harder to predict in an accurate and precise way. As it is with all models, impact prediction is not an exact science and the uncertainties should be clearly stated in the EIA.

Impact management aim to avoid, minimise, remedy or compensate for predicted adverse impacts of the project. It is important that these impacts are managed to the point of public acceptability (Noble, 2011). This can be done by changing the project so

the adverse impacts do not occur, reduced to an acceptable level or mitigating so they become less significant. Mitigation is usually done by project redesign or adjustments to the project operation. Compensation can also be used to overcome or offset impact. The compensation can be 'in kind', like creation of new wetland habitat to replace habitat lost due to the project. In addition to mitigation, opportunities for environmental enhancement should be sought in EIA. The enhancement could be improvement of current environmental conditions and features (Morris and Therivel, 2001).

Implementation and follow-up is an essential step for those projects that proceeded to implementation, in order to ensure that predictions were accurate, that there were no unanticipated impacts, and that the environmental management and mitigation plans developed for the project were implemented and were effective (Noble, 2011). Guidance of the procedure of following up on the project contains four main components:

- Monitoring – establishing a baseline and impact monitoring after the implementation of the project.
- Evaluation – determining conformance with standards, prediction or expectation as well as the environmental performance of the project.
- Management – making decisions and taking appropriate action in response to unanticipated impacts and other issues arising from monitoring and evaluation activities.
- Communication – in order to get feedback on project implementation, impact management, and the EIA process it is important to inform stakeholders about the results of the follow-up studies.

One of the main aims of EIA is also to help analyse and manage conflicts. Here participation is essential, since this is the only way to discover preferences and interests of the different stakeholders involved in the decision (Colorni et al., 1999). As Figure 3-1 also indicates public involvement is an important part of the EIA process. Actually, engagement of the public is required by most countries in some form at least. Public involvement can be defined as the engagement of individuals or groups that are either positively or negatively affected by a proposed project. The form of public involvement can vary from passive participation or information reception (media releases or information bulletins), to participation through consultation (public hearings and open-houses), to interactive participation (workshops, negotiation or even co-management). Different levels of public involvement may be relevant to different stages throughout the EIA process, and different members of the public may need to be involved in different stages in different capacities (André et al., 2006; Noble, 2011). The purpose of public involvement is among other things to (1) access a wide range of information; (2) identify socially acceptable solutions; (3) ensure more balanced decision making; (4) minimising conflict and potentially costly delays; (5) reduce the possibilities of legal challenges; and (5) promote social learning.

The public involvement can be interpreted very differently from project to project or/and from country to country. In (COWI, 2009) they found that different countries allow for public participation in different stages of the EIA process. Half of the Member States (in EU) are allowing participation to take place in the scoping procedure. Some Member States that have not set the scoping phase as the relevant phase for introducing participation have chosen to allow participation already in the screening phase. Others have simply decided to allow participation when making the EIA report available to the public. The last can hardly be said to be an involvement of the public, but rather information to the public.

3.4 Pros and cons of EIA

As with any other method there are both pros and cons for the EIA.

Pros:

- EIA provides an opportunity to learn from experience of similar projects and avoids the (often high) costs of subsequently mitigating unforeseen negative and damaging impacts.
- EIA Improves long-term viability of many projects.
- The EIA procedure ensures that environmental considerations are taken into account in decision making processes.
- Transparency in the environmental decision making.
- Provides systematic methods of impact assessment.
- Estimates the cost/benefit trade-off of alternative actions.
- Facilitates the public participation.

By conducting EIAs an experience is build up around these analyses and how different project types can affect the environment. This experience can be beneficial for future assessments, and thereby help avoiding e.g. similar mistakes to be applied again. By applying an EIA the planners are forced to evaluate on the long-term consequences of a project. This is not always the case when an EIA is not applied. Since the traditional CBA uses discounting of the impacts in the future, these impacts do not count as much as impacts in the near future[9]. Not all environmental impacts can be measures in monetary terms and thereby enter a CBA and might therefore be left out of the decision making process. The EIA ensures that all relevant environmental impacts are considered and treated in a formalised way. By conducting an EIA and documenting the analysis which lay out all the environmental impacts the transparency is improved. Furthermore, having a specific framework for examining the impacts, as an EIA is, a systematic approach is achieved. This ensures that all projects are treated the same way with same analysis for appraisal. Included in the EIA is also an estimation of how to reduce some of the unwanted environmental impacts of a project, and thereby

[9] See next chapter for a discussion of this topic

provides basis for trade-off between the different alternative proposals. Lastly, the EIA facilitates public participation as central part of the methodology. As the EIA estimates physically changes a project might have on the environment, it is fairly easy for the public to discuss and comment on the project proposals.

Cons:

- Where environment, social and economic aspects are addressed, they are not always addressed in an integrated way (EIA reports tend to present as separate chapters).
- Which impacts are to be considered as significant? And thereby included in the EIA.
- Time-consuming.
- Costly.
- Little public participation in actual implementation.

One important aspect of the EIA is how it is used in the decision making context. Even if all the sustainability aspects (environment, social and economic) are taken into consideration in an EIA, they are most of the time not addressed in an integrated way, leaving the DMs up to the task of integrating the aspects themselves. Also there are some doubts about when to include what. When is an impact to be considered significant or not? There exist no clear guidelines for this. An EIA process can be very time-consuming and very costly because of the extent it necessarily must have. And this is with no guaranty that the project will get approved, so for many planners and politician the EIA process can be perceived as a hinder for progress. Even though listed as a pro, actual public participation can be very little in practice.

Even though there exists an EU directive describing the process of conducting EIA there is still room for some very different interpretations of the directive. The COWI report (2009) states some issues with conducting EIA; wide variation between Member States (in EU) in the criteria for "screening", processing of trans-boundary EIAs, the lack of requirements for quality control and results of EIA are poorly reflected in development decisions. The latter issue raises questions about the role of the EIA in the decision process and is also addressed in Jay et al. (2007). One issue regarding this is that EIA reports tend to be presented as separate chapters and not as an integrated part of the overall appraisal.

The conclusion of this discussion on the different pros and cons of EIA is then that EIA improve the decision process significantly on environmental issues. However, there are some unclear areas of the approach and its role in the decision making context. This has led to increasing questioning about the nature of EIA and a recognition that its fundamentally rationalistic approach is out of step with the realities of decision making. This has begun to draw attention to decision making contexts themselves, and suggests

that EIA should be more closely adapted to the processes that it seeks to influence (Jay et al., 2007). Here is where MCDA can come into play.

From above it is clear the EIA provide an approach for assessing the environmental aspects of transport projects in a clear and transparent way. However, environment is far from the only criteria when decisions on transport projects are made. The economy is maybe the far most important factor in choosing among several alternatives for fulfilling specific objectives and in this question the CBA is very useful. However, EIA is often an important input to the CBA and thereby the two methods are connected. Furthermore, in many countries both CBA and EIA are required for investments in infrastructure. In the following chapter the CBA will be treated.

4 Cost-benefit Analysis

The purpose of this chapter is to describe and discuss the role of cost-benefit analysis (CBA) in transport project appraisal with emphasis on how this methodology can help illuminating the research questions relating to decisions that are both rational and legitimate. Furthermore, this chapter will set focus upon the strengths and weaknesses of CBA when applied in transport appraisal.

4.1 Introduction

The fundamental of CBA is that in order to approve funding for any transport project the benefits of the project should exceed its costs. In principle CBA is equally applicable to private and public projects, however since the CBA focuses on social welfare the method is most used for public decision making and therefore also sometimes referred to as the social cost-benefit analysis (Saitua, 2007). The use of CBA is more widespread in the transport sector than in other sectors, even though the method could be used for all kinds of public project appraisal[10]. In the transport sector CBA is applied to infrastructure projects as well as other policy measures e.g. comparing the impacts of congestion charges.

In appraising e.g. a new road the financial revenue will be zero if no charges are made for using the road, so the financial costs of building and maintaining it will always exceed financial revenue. However, the social benefits of the road might be very large, if congestion, travel time, and fuel costs are reduced to the users. Though, the road might produce externalities e.g. noise. A financial analysis would deem the road unprofitable, since revenue was less than costs. A CBA would deem the road to be economically feasible if the social benefits of journeys along the road exceed the opportunity cost of resources used to build it. CBA also seeks to include externalities in the appraisal from e.g. noise in judging whether the road was socially justified (Willis, 2005). CBA sets out to do for government what the market does for business: add up the benefits of a public policy and compare them to the costs (Ackerman and Heinzerling, 2004). However, this imposes some issues, as some of the costs and benefits entering a social CBA are inherently very uncertain as they relate to non-market goods and services, to outcomes very far into the future and with very complex cause-effect relations (Staehr, 2006). The treatment of uncertainty will be further explored in chapter 7.

The argument for using CBA as an appraisal tool is that it is a model based on rationality. Not all political decisions are made on the basis of rational thinking, but CBA's insistence on all gains and losses of utility or well-being being counted means that it forces a wider view of DMs. Furthermore, CBA offers a rule for deciding if anything at all should be chosen, in contrast to e.g. EIA and MCDA. Time can also be

[10] There is an extensive academic literature on CBA, some of which may not use the term cost-benefit analysis but instead refers to benefit-cost analysis, policy evaluation, or project appraisal.

accounted for by using a discount factor, although what the 'correct' discount factor should be is still being discussed and is often a controversial subject (Thomopoulos et al., 2009).

The CBA procedure allows to identify the best projects for the society if – based on an exhaustive (and completely unbiased) study of the costs and benefits of every investment project – only those projects with the best net social benefits emerge (Damart and Roy, 2009, p. 206). The requirements of exhaustiveness and unbiased are later questioned.

4.2 Theory

In CBA it is assumed that the consumers behave in a rational manner and seeks to maximise utility. The consumer's preferences can be represented by a utility function, which is a function defined over a choice set, such that alternative A is preferred to alternative B if and only if the value of the utility function taken on A is higher than the corresponding value of B (Gissel, 1999). Each individual in the society may, in this way, be assumed to maximise his or hers individual utility. However, the CBA is not limited to decisions that affect only one individual, but the society in general. It is assumed that the characteristics of rationality remain if we extend it to the social context. As the society consists of its individuals, it is natural to see the social change in welfare from a given investment as an aggregate value of the individual utility gains and losses (or benefits and costs).

Pareto criterion is one of the central elements in the welfare theory. According to this criterion, social welfare increases when a Pareto improvement is possible, that is, a situation whereby the utility level of at least one individual is raised without decreasing the utility levels of any other individual (De Brucker and Verbeke, 2007). However, in real life most changes will decrease somebody's utility and the Pareto principle will therefore not be applicable for practical decision problems (Willis, 2005).

The impracticality and problem associated with the Pareto criterion led to the formulation of the idea hypothetical compensation as a practical rule for deciding on policies and projects. All that is required is that gainers can compensate losers to achieve a 'Potential' Pareto improvement. The Kaldor-Hicks criterion states that if the gainers could compensate the losers while continuing to gain themselves, then the project should proceed (Willis, 2005). However, the compensation need not actually to be paid, and interpersonal comparisons of welfare are not being made. So in this way the Kaldor-Hicks criterion is less restrictive than the Pareto criterion.

For both the Pareto criterion and the Kaldor-Hicks criterion to be able to be implemented a valuation of utility gains and losses is needed. For the Pareto criterion there has to be some sort of value judgement of the relative merits of the utility gainers and losers in order for the criterion to work in decision making. As for the Kaldor-Hicks criterion the compensation between the losers and the gainers is purely hypothetical, therefore it is not possible to evaluate the overall change in welfare without a valuation of the relative value of these gains and losses (Gissel, 1999).

The preferences of individuals are to be taken as the source of value. To say that an individuals' well-being, welfare or utility is higher in state A than in state B is to say that he/she prefers A to B. The preferences are measured by a willingness to pay (WTP) for a benefit and a willingness to accept (WTA) compensation for a cost[11]. It is assumed that individuals' preferences can be aggregated so that the change in social welfare is simply the sum of all individuals' utilities. The traditional approach in CBA is to assign equal weights to all individuals – so the change in social welfare is an un-weighted sum of the individual utility gains and losses. So questions concerning equity or distributional aspects are not taking care of in the CBA.

4.2.1 Decision criteria

Through discounting it is possible to compare different projects which have different time patterns of cost and benefits, and maybe different lengths of life. Discounting is necessary because there is a preference for having benefits now rather than at some point in the future. The net present value (NPV) of a project compares the discounted stream of benefits with the discounted stream of costs over every year of the life time of the project. If the NPV is greater than 0, then the project is socio-economic feasible. When a selection among several different projects has to be made, the project with the highest NPV is implemented. However, if the problem is to rank projects in order of preference and to select the optimal combination of projects so that their total cost exhaust the budget one have to be careful to rely only on the NPV.

The B/C ratio (BCR) can help identify which set of projects produces most benefits relative to costs –a measure of value for money in other words. Where government finance is constrained, which most often the case, the BCR decision rule approve projects sequentially from those with the highest BCR until all investment funds are exhausted. The use of BCRs will be improved if a common approach is taken across all transport projects when calculating the BCRs. A standardisation is necessary of what is considered benefits and put above the line and what is considered costs to be put below the line. Then the DMs can compare the BCR for a project with those arising in the transport sector generally. The procedure of assessing transport projects have in several countries been standardised with a national framework for conducting social-economic appraisal (Odgaard et al., 2005).

Another decision criterion is the internal rate of return (IRR). The IRR is a rate that equates the discounted value of the benefit and cost stream – a rate that gives an NPV of zero. The rule for accepting a project then becomes, accept if IRR is greater than the discount rate. If IRR is below discount rate then the project is rejected (Pearce and Nash, 1981).

[11] Kahneman and Tversky's prospect theory revealed that people value gains and losses asymmetrically and therefore there exists a gap between WTP and WTA measures (Kahneman and Tversky, 1979). Choosing between WTP and WTA, in a given context, remains a contentious issue, although in large part the choice is determined by whether people affected by a project have a property right to the current or the new level of the good in question.

Summing up on which decision criteria to use, the correct criterion for reducing benefits and costs to a unique value is the net present value. The rule is to adopt any project with a positive NPV and to rank projects by their BCRs.

4.3 The stages of CBA

The CBA has in many countries a formal described procedure for conducting the analysis. However, the overall procedure of conducting a CBA is more or less the same. For later comparison between CBA and MCDA in this thesis the presented procedure of conducting CBA is to some extend adjusted, but the overall principles is intact (see Figure 5-2 for an overview of the stages).

Establish decision context is the first step and involves determining the aims and the stakeholders. **Identify objectives and criteria** contains which goal(s) are to be reached by the project (e.g., minimizing the environmental impact of transport in a specific area). The third step **identify project alternatives** is where the different alternatives for achieving the objectives are set out. The number of alternatives may vary between 2 (e.g. should a certain project be undertaken or not) to any number. Given the complexity of transport projects it is not always possible to identify all of the alternatives at this stage and more alternatives could be added later in the process. **Identifying project impacts** comprises the identification of all impacts resulting from the implementation of the project (see next section for an overview of typical impacts for transport projects). In the next step it is to be decided **which impacts are economically relevant?** This is a clarification of which impacts are to be included and which are to be excluded. Strictly speaking all impacts (benefits and costs) should be included, however, there has to be determined a cause-and-effect between some physical outcome of the project and the welfare of society. Some impacts have to be excluded in order to avoid double counting of the impacts. The next step, **physical quantification of relevant impacts**, involves the physical quantification of the impacts found relevant in the previous step. Not only the quantification, but also at what time they occur in the evaluation period has to be determined, along with possible uncertainties in the estimation of the impacts. **Monetary valuation of relevant impacts** has in many countries been assessed and formal requirements exist of using these valuations in the CBA. Different methods for valuation of impacts are discussed later (see section 4.3.2). **Discounting of costs and benefits flows** is possible once all impacts have been expressed in monetary terms. This will convert them all into present values to take into account the time value of money. After all impacts have been converted into present values it is possible to **examine the results**. The examination of the results can be based on the previously mentioned decision criteria. The final step involves **sensitivity analysis** in order to examine to which parameters the final result is sensitive. These parameters can for example include the discount rate, the physical quantities and qualities of the impacts or the project life span and the number of chosen criteria.

4.3.1 Assessing the impacts

Tracing and measuring impacts is the necessary forerunner to valuing those impacts. Individuals derive utility from intra-marginal journeys over and above the price they pay for the journey. In CBA this "consumer surplus" is counted as a benefit in addition to the price they are willing to pay for a journey (Willis, 2005).

The evaluation of impacts can be described in three steps (Damart and Roy, 2009). First step consists of identifying the advantages and disadvantages of the project evaluated. For each of the impacts a physical measurement scale must be defined and a procedure for measuring the quantity of the impact must be conceived and applied to each of the competing projects. Second step consists of a procedure for assigning a monetary value to each physical unit identified in the first step. Third step the future trend of the impacts have to be determined.

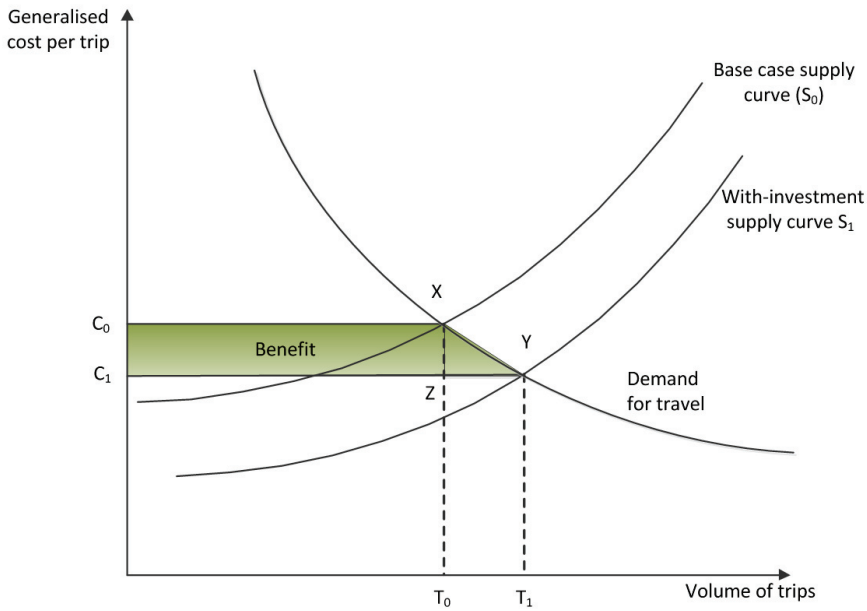


Figure 4-1 Calculation of user benefits

The coloured area in Figure 4-1 shows how the obtained benefits (difference of willingness to pay before and after) are made up of a contribution from existing trip makers (the rectangle area C_0-X-Z-C_1) and a contribution from newcomers (the triangle X-Y-Z). The latter obtain only in average half of the benefit when compared to the benefit of an existing trip maker. This implies that the demand curve is linear, however the case is often that the curve is convex and thereby this approximation of the newcomers benefit is overestimated. Nevertheless, for small changes in generalized cost per trip, this inaccuracy is not significant.

Typical impacts to be included in CBA when appraising a transport projects can be seen on Figure 4-2. In theory all impacts should be covered in a CBA, however this is not the case in practice. This can be because of difficulties in quantifying, valuating the impact or because the impact is considered to be of only minor importance (Odgaard et al., 2005).

Type	Contents
Infrastructure costs	• Construction costs • System operating cost • Maintenance
User benefits	• Passenger transport time savings • Vehicle operating costs • Benefits to goods traffic
Externalities	• Safety • Noise • Air pollution - local/regional • Climate change

Figure 4-2 Typical types of impacts to be included in CBA

Time savings are an important impact in the CBA for transport projects. Often this impact is the largest impact and therefore in many cases essential for the project to be feasible. In The UK 80% of the monetised benefits are often from time savings (Mackie et al., 2001). The valuation of time has been subjected for extensive research for decades and there is a consensus about valuating the time saved during the course of work. This time savings will normally aim to reflect the value of the work being done. But, in the case of non-working time there is no market and valuation is even more complex. For the case of transferring travel time to leisure activities (non-working), the individual value of the time saving will depend on factors such as individual preferences, culture, and marginal utility of an additional unit income and journey purpose (Navrud et al., 2006). Here methods are used to explore individuals' willingness to pay to save time. Revealed preference (RP) and stated preference (SP) methods are often used. RP studies look at behaviour when e.g. individuals are confronted with a choice of mode or route that involves trading time against money. The most notable advantage of the RP approach is that valuation estimates are derived from real economic choices made by individuals in real markets. SP methods offer respondents with a hypothetical travel scenario and choices between paired options, which offer the desired trade-offs. SP methods can, in principle, be used to value any specific non-market good (Fujiwara and Campbell, 2011).

4.3.2 Valuation of impacts in CBA

Once physical impacts have been identified and measured they must be expressed relative to some baseline, usually the 'do nothing' situation, and valuated in monetary terms.

Some costs and benefits of policy decisions can be readily valued because they impact directly on markets, but some cannot, and then other measures are necessary. It will never be possible in practice to value all impacts, but one should aim to extend valuation to as many as possible. Valuation is implicit in most policy decisions, and it is preferable to make it explicit where possible to improve quality and transparency.

Different valuation principles in use for evaluation of infrastructure projects can generally be divided in two groups, where:

- Prices exist in the market.
- Prices can be assigned from quasi-market observation.

For especially infrastructure costs the prices are most likely to be found in existing markets and are thereby easily handled in the CBA. However, for user benefits and externalities there might not be a market, and then quasi-market observations are necessary in order to put a monetary value on the impacts.

In the appraisal of transport projects there are many costs and benefits which do not have a direct market price (e.g. time saving, pain and grief from accidents and environmental impacts). For some of these impacts, values can be assigned or inferred from observed or stated human behaviour. The principal methods normally used are (Danzanvilliers et al., 1999):

- Revealed preference
    - Hedonic pricing
    - Travel cost method
- Stated preference
    - Contingent valuation method
    - Choice modelling method

Revealed preference (RP) refers to the valuation of non-market impacts by observing actual behaviour and purchases made in actual markets (Pearce et al., 2006). RP method implicitly assumes, that individuals know the value of the good in question and are aware of the effect on their utility that an increase or decrease in the good in question would bring. The RP approach has e.g. been used to try and discover what people are willing to pay to save time. The most popular case is that of choice of travel mode, where people may have a choice between two modes, one which is faster and more expensive than the other. If a model gives an estimate of the probability with which a person chooses one mode of transport rather than the other as a function of journey time, money cost and any other relevant quality difference, the relative weight attached to time and money can be used to estimate that persons value of time (Danzanvilliers et al., 1999; Leleur, 2000).

Hedonic pricing, which correlates the environmental good or bad with some actual market item such as houses, so that variations in the price of houses from one locality to another can be correlated with the presence

or absence of some desirable or undesirable environmental feature e.g. a view. How much people are willing to pay is then supposed to reflect their preference for the environmental good in question, or their aversion to the bad. The starting point for the method is the observation that the price of a large number of market goods is a function of several characteristics. The Hedonic pricing method uses statistical techniques to isolate the implicit price of each of these characteristics then.

Travel cost method, is based on the assumption that detailed sample survey of travellers to a site determines how they value the characteristics of the site, and the time spent travelling to the site (e.g. visiting a nature area).

Stated preference (SP), approaches are survey-based and obtain people's intended future behaviour in constructed markets. By means of an appropriately designed questionnaire, a hypothetical market is described where the good in question can be traded.

Contingent valuation method is perhaps the dominant stated preference method or survey-based technique. By means of an appropriately designed questionnaire, a hypothetical market is described where the good in question can be traded. A random sample of people are asked to express their maximum willingness to pay (or willingness to accept) for a hypothetical change in the level of provision of the good. Respondents are assumed to behave as though they were in a real market.

Choice Modelling is a survey-based methodology for modelling preferences for goods, where goods are described in terms of their attributes and of the levels that these attributes take. Respondents are presented with various alternative descriptions of a good, differentiated by their attributes and levels. By including price/cost as one of the attributes of the good, willingness to pay can be deduced.

Once all the impacts have been assigned a monetary value it is possible to convert them into present values by discounting. Typically when assessing transport projects the time horizon for how far into the future the CBA shall cover, is set to the expected physical life time of the project. However, if the CBA concerns policies, it is not so straightforward to set the time horizon, because it can be difficult to estimate how long the effects of policies last. Another argument for setting the time horizon is that it should be set according to either the uncertainty of future estimates, or to the extent to which discounting makes future gains and losses insignificant[12]. The latter case arises because any discounting of impacts over long distant future quickly reduces the magnitude to insignificant values (Pearce et al., 2006).

[12] Regarding the uncertainty, it is very difficult to estimate impacts that will happen 30 or 40 years out in the future, the estimates cannot be very accurate.

The discount rate should reflect the alternative return on investment of the project costs if the project drives out other investments, and it should reflect consumers' time preference if the project drives out consumption. In European countries the discount rate varies, but is mainly set to about 3-6% (Odgaard et al., 2005). Setting a positive discount factor implies that a lower weight is put on the future than on the present. E.g. France lowered its discount rate for transport project appraisal from 8% to 4% in 2006 (and to 3% for flows occurring 30 years or more into the future). The declining discount rate regime means that costs and benefits occurring after 30 years are now more significant – a time declining discount rate. This approach is also used in the UK (DfT, 2011). The argument for using a time declining discount rate, is that people do not behave as if their own discount rate are a constant (Pearce et al., 2006). The setting of the discount rate is one of the most debated issues in CBA. Discounting has a theoretical rationale in the underlying welfare economics of CBA, but with consequences that may seem to find morally unacceptable. This stems from the fact that current activities imposing large costs on future generations may appear insignificant in CBA, and similarly actions now that will benefit future generations may not be commenced in light of a CBA (Ibid.).

4.4 Pros and cons of CBA

As with any other appraisal methodology there are both pros and cons of CBA.

Pros:

- Measures gains and losses in the same unit. CBA provides an absolute monetary measure of social profitability.
- Value on cost and benefits are consistent which implies that social profitability can be compared across different projects and sectors.
- Possible to assess impacts over time by using a discount factor.

When all gains (benefits) and loses (costs) are measured in the same unit (money), it possible to assign an absolute evaluation of the proposed project in a unit which is easy to relate to. Furthermore, a monetary value on all costs and benefits provides a base for comparing different alternatives and even projects from different sectors. This makes CBA a useful tool for allocating the limited funds a society poses to maximise the welfare of the same. By the use of an appropriated discount factor, it is possible to assess benefits and costs which may appear in the future in the present value.

Cons[13]:

- Pricing the priceless.
- Distributional issues.
- Measurement problems, concrete measures of benefits and costs are necessary, but not accurate.

[13] See also Mackie and Preston (1998) they address twenty-one sources of error and bias in transport project appraisal. Some of these are overlapping with the mentioned cons her.

- Discounting.
- Built-in biases.

Putting monetary values on impacts which do not have a market, can be seen as rather controversial e.g. when a power plant pollutes the air, our gains from the cheap power thus obtained simply cannot be compared with the pristine view of the Grand Canyon we sacrifice (Frank, 2008). Also, CBA requires that all benefits and costs are monetised, but this is not always possible to fulfil. Particularly impacts of policy measures, strategic impacts and long term environmental impacts can be very difficult to assign a monetary value. Even in the case where an estimate of the monetary value can be assessed, the cost of such analysis might be considerable and the analysis might not be justified (Gissel, 1999). During the 1970s a growing concern about non-material costs and benefits was expressed. If these often environmental costs were applied to the CBA, they often seemed to be trivial in comparison with the material benefits obtained from the projects, especially for transport projects. CBA seemed to be out of step with political pressures and changing values for more attention towards the environment. Yet CBA should have reflected those changing values, proceeding as it does on the basis of eliciting values through simulation of wants and desires in 'shadow' markets. This apparent inconsistency was enough to cause a major shift in attitude towards the worth of CBA and emerging in a range of alternative techniques for appraisal of projects (Pearce and Nash, 1981).

As no actual compensation between consumers takes place, the distributional effects of a project are ignored. In a separate analysis these equity effects needs to be taken into account, even though it will often be difficult to distinguish between the socio-economic groups affected by the project. Moreover, it seems legitimate to ask whether society, constructed as the sum of the individual comprising it, could meaningfully be said to be 'better off' if the loser actually loses (Ibid.).

As with any other evaluation methodology there are difficulties in getting precise estimates of the expected impacts. This inaccuracy can have a big influence on the final outcome of the analysis.

CBA requires that the present generation values impacts, which might have large consequences in the future. This means that they are valuing such impacts on behalf of the future generations. Furthermore, the value of resources (e.g. environmental) may rise in the future, because they get more limited. Another aspect of the time horizon is that it can be very difficult to foresee the future developments in society (traffic behaviour, preferences, social norms etc.) and economic developments.

Some of these sources of error and bias can be difficult to deal with methodologically, and some are outside the scope of the appraisal process, e.g. prior political commitment can be difficult to address in the appraisal process. But, for other sources it can be possible to introduce new methods or procedures for appraisal in order to eliminate errors and bias. The treatment of non-quantifiable impacts can, as suggested, be managed by an MCDA.

Even though all of these mentioned issues of technical character would be solved, there have been raised doubts if it would be enough to reverse the opposed attitude that several planning actors have towards the use of the instrument itself. Beukers et al. (2012) found, through interviews with different types of participants in the CBA process (e.g. regional politicians, academics, planners and those who apply the CBA), several different issues with the CBA process. In general they found that the issues were a result of clashing values and approaches resulting in e.g. that the CBA process being perceived as a black box; and strategic behaviour in relation to CBA input and output. As discussed in the section 1.1 there seems to be a mismatch between the results of the CBA and the actual decisions made, although a perfect match is not to be expected, some relation is expected between the CBA results and the decision made. A reason for this unexpected relation could be the mistrust of the DMs towards the CBA.

As suggested in (OECD, 2011) one way of improving the CBA process is to put more emphasis on presentation and discussion of the obtained results from the CBA, rather than refining the technical aspects of the CBA. As later described in paper 3 there have been made efforts in dealing with the inaccuracy of transport models as concerns estimating costs and time savings. Not only point estimates, but intervals illustrating the results of a CBA. Such information is very useful for the DMs when they have to make their decisions.

From the discussion on the different pros and cons of CBA the conclusions must be that the appraisal of transport projects using CBA is useful for supporting decisions, but it is not sufficient, as it provides no, or very imprecise, information on a range of effects that DMs care about. Thus while economic evaluation has established itself as a useful tool in this area, its limitations must be recognised, and its results weighted with other evidence before decisions can take place. Most likely CBA will to some degree maintain its position as an important part of the decision making in transport projects appraisal (especially in the light of the current global declining economic development). As discussed in this chapter there seems to be room for improvement for some of the technical issues and the process of the CBA in order to gain more of the advantages that the CBA offer in principle. Traditionally CBA has been an important factor in deciding on which transport projects should be implemented, but recently CBA weigh less in decisions than it used to (Ibid.). One approach, that could be suggested, is to embed the CBA in the more participative oriented MCDA approach, and thereby provide opportunity for better communication between the different types of participants of the CBA. In the following chapter the MCDA will be presented and discussed.

5 Multi-criteria Decision Analysis

The purpose of this chapter is to introduce multi-criteria decision analysis (MCDA) as an appraisal methodology for transport projects. This relates to research question A and B concerning how to obtain decisions that are both rational and legitimate and how to involve stakeholders in the appraisal. Especially the public acceptance will be treated in this chapter; lastly, also a comparison between CBA and MCDA will be presented together with the strengths and weaknesses of MCDA.

5.1 Introduction

Multi-criteria decision analysis (MCDA) is a range of evaluation techniques that seek to help DMs in situations when judgment depends on more than one criterion. MCDA can be seen as the main judgement-oriented alternative to CBA (Howarth et al., 2001). In MCDA, the performance of project alternatives is confronted by the objectives of the DMs, stakeholders, or social welfare at large. Overall MCDA can be described in a very simple way: MCDA takes a set of alternatives, defines a set of criteria for decision making and ways of measuring them (measures are often not in the same units), and derives weights for the criteria. Alternatives have different scores on the different criteria. Performance scores of alternatives on the different criteria are aggregated into a total score (Sijtsma, 2006).

An MCDA problem can be expressed in a matrix format, with alternatives as columns and criteria as rows and then the performance values (scores) in the cells (Yoon and Hwang, 1995). This performance matrix is seen as the core of MCDA. The basic principle underlying every MCDA approach is that explicit introduction of several criteria, each representing a particular dimension of the problem to be taken into account, is a better path for robust decision making, than optimising a single dimensional objective function (as CBA). The MCDA facilitates learning about the problem and the alternative courses of action, by enabling people to think about their values and preferences from several points of view (Bana E Costa et al., 1997).

MCDA affects all phases of the planning process. Initially, an extensive analysis is required to ascertain objectives, their significance and likely conflicts. Where several objectives exist, as assumed here, the DMs preferences play a decisive role and must be investigated in detail. The MCDA procedure discussed in this chapter supports processes of goal setting and decision-planning.

As stated above MCDA is not one single method, but a range of different evaluation techniques, which all seeks to assist DMs in situations when judgement depends on more than one criterion. As Belton and Stewart formulates it (2002, p. 2):

> *Thus, we use the expression MCDA as an umbrella term to describe a collection of formal approaches which seek to take explicit account of multiple criteria in helping individuals or groups explore decisions that matter.*

Each of the different MCDA methods has advantages and disadvantages. To mention some of the different methods: The lexicographic model consists in applying the criteria one-by-one to the set of alternatives, eliminating those alternatives that do not fulfil a limit established by the applied criterion. Ideal point approach is to find or generate alternatives as close as possible to the ideal point: the alternative that makes discrepancies disappears. Another variant is the anti-ideal point, here defensible alternatives should be as far from the anti-ideal point. Outranking methods are another family of MCDA methods. These are based on the concept partial comparability. This is that the preference between two alternatives can be modelled by means of binary relationships: indifference, strict preference, weak preference and incomparability. These models use pair-wise comparisons of alternatives in order to establish whether one alternative is at least as good as the other, according to most of the criteria. Other methods again are based on the linear aggregation of criterion scores.

MCDA aims to be both a tool and a decision process for informing DMs on how to regulate to achieve public policy goals. Here I assume: a DM chooses one (or a subset) of a set of alternatives evaluated on the basis of two or more criteria or attributes. The feasible set of alternatives may be either small or finite or large and perhaps infinite. Uncertainty may be involved. More conceptually, it is assumed that a DM acts to maximize a utility or value function that depends on the criteria or attributes. MCDA methods are intended to help a DM think about the problem as part of the decision making process. Actually, there has been an argument for shifting away from the sole focus on outcome, towards the quality of decision processes (Spash et al., 2006).

For many investment decisions, the DMs wish to pursue several objectives, rather than a single objective as the CBA described in previous chapter. Such a decision making problem is typical in strategic investment decision making as, e.g. when choosing among several different transport investments. The following chapter describes and discusses models and procedures developed to satisfy several objective measures simultaneously.

5.2 Theory

MCDA is explained in several textbooks (Belton and Stewart, 2002; Goodwin and Wright, 2009; Leleur, 2000; Lootsma, 1999; von Winterfeldt and Edwards, 1993). However, MCDA seems to be less of a unified technique than CBA. Although despite this lack of a uniform approach and sophistication, most practitioners often make use of rough MCDA implementations, which are far from the complexity of the largest part of the MCDA literature (Beinat, 2001).

Because of the above it would be fruitful to start with some overall lying principles for MCDA; to be able to assess the different alternative scores on different criteria and aggregate them, it is necessary to construct a model to represent preferences and judgements. Such a preference model contains two components (Belton and Stewart, 2002):

1. Preference in terms of each individual criterion, the relative importance of achieving different levels of performance for each criterion.
2. An aggregation model, allowing inter-criteria comparisons, in order to aggregate preferences across criteria.

Within the value measurement approach, the first component of preference modelling (measuring the relative importance of achieving different performance levels for each identified criterion) is achieved by constructing marginal (or partial) value functions, that is $v_i(a)$, one for each criterion. A fundamental property of the partial value function must be that alternative a is preferred to b in terms of criterion i if and only if $v_i(a) > v_i(b)$; likewise, indifference between a and b in terms of this criterion exits if and only if $v_i(a) = v_i(b)$. However, the partial value functions will in addition need to model the strength of preference in some sense. When criterion i is associated with an attribute $z_i(a)$, then $v_i(a)$ is of necessity a non-decreasing (but not necessarily linear) function of this attribute. The marginal value function can then be described as a function of this attribute directly, say $v_i(z_i)$, without any reference to any specific alternative a (Belton and Stewart, 2002).

A central feature of value measurement is that the properties required of the partial value function and the form of aggregation used is critically interrelated. The most frequent used aggregation form is the additive:

$$V(a) = \sum_{i=1}^{m} w_i v_i(a)$$

Where, $V(a)$ expresses the overall performance value for alternative a, calculated as the weighted sum of the performance on m criteria, where the criterion i is weighted with the weight w_i.

The reason for the additive model being the most frequently used, is that it is the form easily explained and understood by DMs from a wide variety of backgrounds, while not placing any substantially greater restriction on the preference structures than more complicated aggregation formulae (Ibid.)

Within MCDA there exist two main strategies; compensatory and non-compensatory. Compensatory strategies (or methods) allow bad performance by an alternative on one or more attributes to be compensated for by a good performance on others. In this case a bad performance on one or more attributes doesn't necessarily rule out the alternative. It can be argued that DMs using this strategy are using all available information when making their decisions. The Multi-attribute value theory (MAVT) and Multi-attribute utility theory (MAUT) belongs to compensatory strategies. Non-compensatory strategies operate on rules that do not allow a bad rating on one attribute to be balanced out by a good rating on another. A bad rating usually leads to the alternative being rejected. Satisficing is non-compensatory, given that an alternative is rejected as soon as it fails to reach an acceptable level on any of the attributes (French et al., 2009).

Three broad groups of MCDA techniques can be identified (Stewart, 1992):

- Goal programming and reference point approaches.
- Descriptive methods and outranking methods.
- Value/Utility function approaches.

With a large number of alternatives, goal programming is useful and can be used to narrow the search for the 'best' solution quite considerably. If the alternatives have been reduced to a small number, descriptive and outranking methods can be used with advantages. They are often used when compensation between performance areas is difficult and when it is doubtful whether a 'true' order of alternatives exists which corresponds with the preferences of the DMs. Value function approaches are very judgement oriented and present a full ranking of alternatives. This method is the one of the most used MCDA approaches. The most widespread value function approach is the additive scoring. Additive scoring means that the performance of an alternative is determined by adding the scores (standardised and weighted) on different criteria. Additive scoring approaches are generally regarded as very effective and easy to understand, and maybe therefore also recommended for use if the rationale for the final choice has to be defended in the public (Stewart, 1992). However, it should be noted here that it is very important to ensure that the definition of criteria, and the scoring method used are properly justified and understood by those providing the inputs in order for the model to provide valid output.

This PhD thesis will only treat a restricted number of methods for MCDA. Because human beings usually express their preference in terms which reveal degrees of intensity (indifference, weak, definite, strong, or very strong preference), I limit myself to cardinal methods on the assumption that the preference information which is obtained in the elicitation procedures establishes ratio or difference information[14]. Two different variants of the pairwise comparison method will be treated in the following sections: additive AHP were their relative performance can be expressed as a difference of grades and multiplicative AHP- REMBRANDT were the performance is measured as a ratio of subjective values (Multiplicative AHP - REMBRANDT)

5.2.1 Pairwise comparisons – AHP and REMBRANDT

Analytic Hierarchy Process (AHP)

The method was developed by Saaty (1977) in the early 1970s to structure and analyse complex decisions. On important application of the method is the support of decision making involving multiple objectives.

The AHP splits the decision process into partial problems in order to structure and simplify it. A hierarchy containing multiple objectives is constructed, such that the main

[14] E.g. when examining different alternatives for a specific transport project, the difference between the alternatives will be the basis for selecting the preferred alternative

objective is broken down into sub-objectives. At the lowest level of the hierarchy, the alternatives are included.

Using AHP, both quantitative and qualitative criteria can be considered. In each case, the relative importance (weightings) of the different criteria, and the relative performance of alternatives, are determined with respect to each element of the higher level by using pairwise comparisons. Then, a total value is calculated for sub-targets to determine their relative importance for the whole hierarchy, and, ultimately, to assess the overall performance value for the alternative projects.

The AHP is carried out using the following steps (Saaty, 2008):

1. Structure the decision hierarchy from the top with the goal of the decision, then the objectives from a broad perspective, through the intermediate levels (criteria on which subsequent elements depend) to the lowest level (which usually is a set of the alternatives).

2. Construct a set of pairwise comparison matrices. Each element in an upper level is used to compare the elements in the level immediately below with respect to it.

3. Use the priorities obtained from the comparisons to weigh the priorities in the level immediately below. Do this for every element. Then for each element in the level below add its weighed values and obtain its performance value. Continue this process of weighing and adding until the overall performance values of the alternatives in the bottom most level are obtained.

4. Examinations of the consistency of the performance assessments.

With regard to the pairwise comparisons, it is assumed that the DM is able to determine values a_{pq} for all pairs p and q from the set A (target criteria or alternatives) on a relational scale. Reciprocity should apply for the estimated values. That is, the comparative value of p relative to q must equal the reciprocal of the comparison between q and p. A comparative value a_{pq} should never be infinite. An infinite relative importance would mean the criteria or alternatives regarded were not comparable, and a renewed problem analysis would be required.

For the pairwise comparisons, the nine-point scale suggested by Saaty (1977) is used (Figure 5-1).

Scale value	Definition	Intrepretation
1	Indifference	Both compared elements have the same importance
3	Weak preference	Experience and estimation suggest a slightly greater preference of one element in comparison with the other element
5	Definte preference	Experience and estimation suggest a definte preference of one element in comparison with the other element
7	Strong preference	The strong preference of on element in comparison with the other element has been shown clearly
9	Very strong	The maximum difference of preference between two elements
2,4,6,8	Intermediate values	

Figure 5-1 Saaty's nine-point scale for pairwise comparisons

This scale has the advantage of converting verbal comparisons into numerical values, so that measurability on a relational scale is possible.

The results of pairwise comparisons related to an element may be shown in the form of a C x C matrix with C elements being compared. The values along the main diagonal of this pairwise comparison matrix are always 1. To determine a set of relative priorities amongst n alternatives, only n-1 judgements are in principle needed. By asking for a complete set of pairwise comparisons (n(n-1)/2 in total), more information than necessary is provided, hopefully reducing the impact of a poor, or erroneous response. Furthermore this introduces the possibility of assessing the consistency of judgements. The first step in synthesising these judgements is to reduce the pairwise comparison matrix to a comparison vector, i.e. a set of performance values representing the relative performance of each alternative. The values in the pairwise comparison matrix are interpreted as ratios of these underlying scores. E.g. the value of 3 in a comparison matrix (*p* compared to *q*) indicates that the assessment of *p*'s performance is 3 times as large as the corresponding performance of *q*.

The aim in AHP is to find the set of performance values $v_1,\ldots v_n$, such that the matrix values a_{pq} are approximated as closely as possible by corresponding ratios v_p/v_q. The standard AHP method of doing this is to extract the eigenvector corresponding to the

maximum eigenvalue of the pairwise comparison matrix. A further development of the AHP method uses the geometric mean method to find these values.

An examination of the consistency is possible if the eigenvector method has been utilised. If all the assessments are totally consistent, the maximum eigenvalue is C. Where there are inconsistencies, however, a higher eigenvalue λ_{max} arises. The difference between λ_{max} and C increases with increasing inconsistency, so it provides an indication of the consistency of the estimates. A consistency index has been formulated by Saaty:

$$CI = \frac{Principal\ eigenvalue - size\ of\ matrix}{size\ of\ matrix - 1} = \frac{\lambda \max - C}{C - 1}$$

The consistency index is compared to a value derived by generating random reciprocal matrices of the same size, to give a consistency ratio (CR) which means to have same interpretation no matter what the size of the matrix. The random indexes from random matrices are as follows for $3 \leq n \leq 10$ (Belton and Stewart, 2002, p. 156)

Table 5-1 Random index for consistency ration calculation

Size of matrix	3	4	5	6	7	8	9	10
Comparative value	0.52	0.89	1.11	1.25	1.35	1.40	1.45	1.49

Then the CR can be calculated as the consistency index divided by the respective random index. A CR of 0.1 or less is generally stated to be acceptable. Accordingly, pairwise comparison matrices with a CR ≤ 0.1 are regarded as being sufficiently consistent, while matrices with CR > 0.1 require an examination and revision of the pairwise comparisons.

The AHP method has some pros and cons, and these will be discussed in the following. The pros of the AHP method are that conducting pairwise comparisons is simple and straight forward and this means that the DMs can focus, in turn, on each small part of the problem. Only two attributes or options have to be considered at any one time. Verbal comparisons are also likely to be preferred by DMs who have difficulty in expressing their judgements numerically; The AHP requires more comparisons to be made than needed to establish a set of performance values (PVs). In this way it is possible to check for inconsistency in the judgments; The practical nature of the AHP method has led to its application in highly diverse areas including education, politics and sociology (Yoon and Hwang, 1995) this proves its versatility (se Vaidya and Kumar (2006) for an overview). However, there has also been extensive debate about the AHP method and its cons. First the conversion to a numerical scale from the verbal scale has been criticised (Belton and Stewart, 2002; Goodwin and Wright, 2009). DMs using the verbal method of comparison will have their judgements automatically converted to the numeric scale, but correspondence between the two scales is based on untested assumptions; The 1-9 scale is bounded to create inconsistencies if the DMs using the verbal scale wish to incorporate very extreme rations in the decision model; New

alternatives can reverse the rank of existing alternatives. This last criticism has attracted much attention, however, empirical evidence has shown that this rarely happens in practice (Belton and Stewart, 2002).

REMBRANDT

A group in the Netherlands, led by F.A. Lootsma, has developed a system which uses **R**atio **E**stimation in **M**agnitudes or deci-**B**ells to **R**ate **A**lternatives which are **N**on-**D**omina**T**ed (REMBRANDT). The REMBRANDT system has been designed to address three criticized features of AHP. The first issue addressed by Lootsma is the numerical scale for verbal comparative judgment. Saaty presented a verbal scale for the ratio of relative value between two objects. Lootsma feels that relative advantage is more naturally concave, and presents a number of cases where a more nearly logarithmic scale would be appropriate, such as planning horizons, loudness of sounds, and brightness of light. Therefore, Lootsma presents a geometric scale where the gradations of DM judgment are reflected by the scale as follows (Lootsma, 1992; Olson et al., 1995):

1/16	Strict preference for object 2 over base object
¼	Weak preference for object 2 over the base object
1	Indifference
4	Weak preference for the base object over object 2
16	Strict preference for the base object over object 2

The comparative judgements are converted into values on a geometric scale, characterised by scale parameter γ (van den Honert, 1998):

$$r_{jk} = exp(\gamma \delta_{jk})$$

Lootsma considers two alternative scales γ to express preferences. For calculating the weight of criteria, $\gamma = \ln \sqrt{2} \approx 0.347$ is used. In REMBRANDT, only one hierarchical level (no matter how many criteria) is used, superior to the level of alternatives. For calculating the scores of alternatives on each criterion, γ = In 2 ≈0.693 is used. The difference in levels of value δ_{jk} is graded as in Table 5-2, which compares Saaty's ratio scale with the REMBRANDT scale.

Table 5-2 AHP scale and corresponding REMBRANDT scale (Olson et al., 1995, p. 524)

Verbal description	Saaty's ratio w_j/w_k	REMBRANDT δ_{jk}
very strong preference for object k	1/9	-8
strong preference for object k	1/7	-6
definite preference for object k	1/5	-4
weak preference for object k	1/3	-2
indifference	1	0
weak preference for object j	3	2
definite preference for object j	5	4
strong preference for object j	7	6
very strong preference for object j	9	8

The second improvement concerns the calculation of PVs. The arithmetic mean used in AHP has been shown to be subject to rank reversal of alternatives. The geometric mean is proposed since this is not subject to rank reversal, nor is logarithmic regression (Olson et al., 1995).

Logarithmic regression has been proposed (Lootsma, 1992), minimising $\sum_{j<k} \left(\ln r_{jk} - \ln v_j + \ln v_k\right)^2$ where r_{jk} are the ratio comparisons made by the DM for base object j and compared object k, and the weight for $j(w_y)$ is represented by In v_j. Ratio r_{jk} is the ratio of w_j/w_k. The analysis is to calculate these weights. Since $r_{jk} = w_j/w_k$, error is represented by $r_{jk} - w_j/w_k$ the ratio comparisons made by the DM are observations, and regression minimizing the squared error yields the set of weights w_i which best fit the DM stated preferences. Solving this is complicated further by the fact that the resulting data set is singular. However, a series of normal equations can be solved to yield the desired weights (Olson et al., 1995).

The third improvement is aggregation of scores. REMBRANDT uses one hierarchical level (allowing 25 criteria), with the alternative level (allowing 25 alternatives) subordinate to it. This lowest level is normalised multiplicatively, so that the product of components equals 1 for each of the k factors over which the alternatives are compared. Therefore, each alternative has an estimated relative performance w_k for each of the k factors. The weights, O(j), are normalized additively, so that they add to 1. The aggregation rule for each alternative j is (Lootsma, 1999):

$$w_j = \prod_{i=1,k} w_i^{o(i)}$$

The REMBRANDT method shares some of the same pros as the AHP method. The use of pairwise comparison breaks down the decision problem to comparing two alternatives at a time. The use of verbal comparison is as well as in the AHP a pro. However, then also some of the cons, that follows.

5.2.2 Methods for elicitation of weights

The purpose of a criteria weighting technique is to establish a set of cardinal or ordinal values which indicate the relative importance of each criterion.

Although AHP and REMBRANDT provide a method for assessing criteria weights (CWs) it is possible, if desirable, to use other methods as well. Some of these methods will be treated in the following with a discussion concerning the concept of CWs.

Applying an MCDA requires weights to be attached the different elements of the appraisal. In CBA, the weights (the monetary valuation of the effects) are determined on the basis of the best available evidence. But, in MCDA these weights can reflect evidence, expert opinion, stakeholder opinion or even policy preferences. When defining weights on the basis of policy preference, subjectivity is introduced to the analysis. It is important to be able to communicate the influence that the weights have on the overall appraisal to the DMs, in a transparent and concise way.

One of the most difficult tasks in MCDA is determining the weights of individual criteria. However, this step is necessary in order to be able to compare all alternatives based on the aggregate performance of all criteria (Kao, 2010). Because of the difficulties with assigning weights, several methods have been proposed for reducing the burden of the process. Several of these methods involve asking the DM simple questions about relative importance of the attributes and then using the response to identify weights that are intended to approximate the DM's 'true' weights (Roberts and Goodwin, 2002).

Assessing the impacts does not necessarily differ from those used in CBA. However, which weights should be attached to the various impacts is often debated (OECD, 2011).

Most methods for determining weights fall into one of three categories (Roberts and Goodwin, 2002):

a) *Direct rating:* uses direct numerical ratio judgements of relative attribute importance. One of them is the Swing weighing method.

b) *Point allocation:* the DM has a 'budget' of points to allocate between attributes in a way that reflects their relative importance.

c) *Ranking:* a simple ranking of the attributes after importance is translated into 'surrogate' weights that represent an approximation of the 'true' weights.

The ranking method was introduced because of evidence, that the weights generated are highly influenced by the method used to produce them. There wasn't and still isn't agreement about which method to use, there is no way of directly identifying the 'true' set of weights (Ibid).

Below the direct rating and ranking method for determining weights are elaborated.

Swing weighting takes into account both the importance of the attributes in determining the preference and the particular difference. Using the Swing weight method, the DMs are asked to consider a hypothetical alternative that is described by the worst consequences on all attributes. Then they are asked which attribute they would like to change ("swing") from its worst level to its best level. Once this judgement is made, they are asked which attribute they would like to change next and so on until a complete rank order of these weights is established. With the rank order in place, one can assign an arbitrary weight of 100 to the highest ranked attribute and ask the DM to scale the remaining weights based in their relative importance of those ranges (Keeney and Von Winterfeldt, 2007). The Swing weighing method, which is a ratio based method, can be demanding for the stakeholders and DMs, but it does give direct and precise weights (Sijtsma, 2006).

Several methods for determining weights by a ranking of the attributes have been developed, Rank Order Centroid (ROC), Rank Sum (RS), Rank Reciprocal (RR) and Rank Order Distribution (ROD). All of them rely on the idea that a 'true' set of weights exists. ROC weights are the best surrogate weights to use for the point allocation method and the ROW weights are the best for direct rating methods of assigning raw weights. However, ROD weights tries to overcome some of that critique ROC has received. This critique was concerning that the ROC weights are too extreme, and that the lowest ranked attribute will only have a very marginal influence on the decision (Roberts and Goodwin, 2002). The ROD weights are constructed so they are not as extreme as the ROC weights.

Table 5-3 Allocation of ROD weights, dependent on the number of criteria (Roberts and Goodwin, 2002)

ROD Weight for rank no.	Number of criteria									
	10	9	8	7	6	5	4	3	2	1
1	0.1867	0.2058	0.2292	0.2590	0.2966	0.3471	0.4180	0.5232	0.6932	1
2	0.1667	0.1808	0.1977	0.2174	0.2410	0.2686	0.2986	0.3240	0.3068	
3	0.1466	0.1565	0.1672	0.1781	0.1884	0.1955	0.1912	0.1528		
4	0.1271	0.1332	0.1375	0.1406	0.1387	0.1269	0.0922			
5	0.1081	0.1095	0.1084	0.1038	0.0908	0.0619				
6	0.0893	0.0867	0.0805	0.0679	0.0445					
7	0.0709	0.0644	0.0531	0.0334						
8	0.0527	0.0425	0.0263							
9	0.0349	0.0211								
10	0.0173									

In order to prevent strategic behaviour by stakeholders Renn (2006) argues that the actual measurement of impacts should be left to a group of independent experts and the weighting to an unbiased jury of uncommitted citizens.

5.3 The stages of MCDA

Establish decision context involves outlining the aims as well as identifying DMs and other stakeholders. It is important to know when and how stakeholders contribute to the analysis, as well as in what form and how this implementation constitutes the design of the social and technical system for conducting the MCDA.

Identify objectives and criteria will be the next step. Objectives can be created either using a top-down approach, in the case of having a larger project where objectives need to be set by a central body, or a bottom-up approach, where various stakeholders participate in generating objectives (Gamper et al., 2006). The criteria should be identified so that it allows measuring the strength of the options in fulfilling the objectives in this step. Criteria should be developed through participation of the stakeholders, in order to make sure all interests are represented and can then be regarded when conducting the analysis.

The third step **identify project alternatives** is where the different alternatives for achieving the objectives are set out. The number of alternatives may vary between 2 (e.g. should a certain project be undertaken or not) to any discrete number. Given the complexity of transport projects it is not always possible to identify all of the alternatives at this stage and more alternatives could be added later in the process.

Next step for MCDA is **scoring options (alternatives) against criteria.** The creation of the performance matrix requires the determination of the relationships between alternatives and their impacts on criteria. Alternatives still cannot be compared as a

unit of preference since the values do not as yet reflect preferences as such, so the next step will be **weighting of criteria.** The assessment and the elicitation of weights does not necessarily have to be done by the same group of people. The assessment can be carried out by experts on the examined area and then the weights can be set by the DMs[15].

Creation of adjusted performance matrix will follow as the next step. Once the scores are re-calculated with their weights they can be counted together and a final sum of each option can be given. These scores can be built into the adjusted impact matrix.

Next step is **examining results**. The interpretation of MCDA results can show the differences in the rankings of the alternatives.

The final step is **robustness analysis and measures** which uses sensitivity analysis to discover which parameters the final outcome is sensitive to, and thereby measures the robustness of the different alternatives.

The stages of the MCDA are, for comparisons sake, illustrated on Figure 5-2 along with the stages of CBA. From the figure it is possible to see that the two methods share many of the stages in the process of conducting CBA and MCDA – they share so to say some of the same 'language'.

[15] The weights can even be set by different groups of stakeholders and thereby the DMs can consider different sets of weights, when the decision is made.

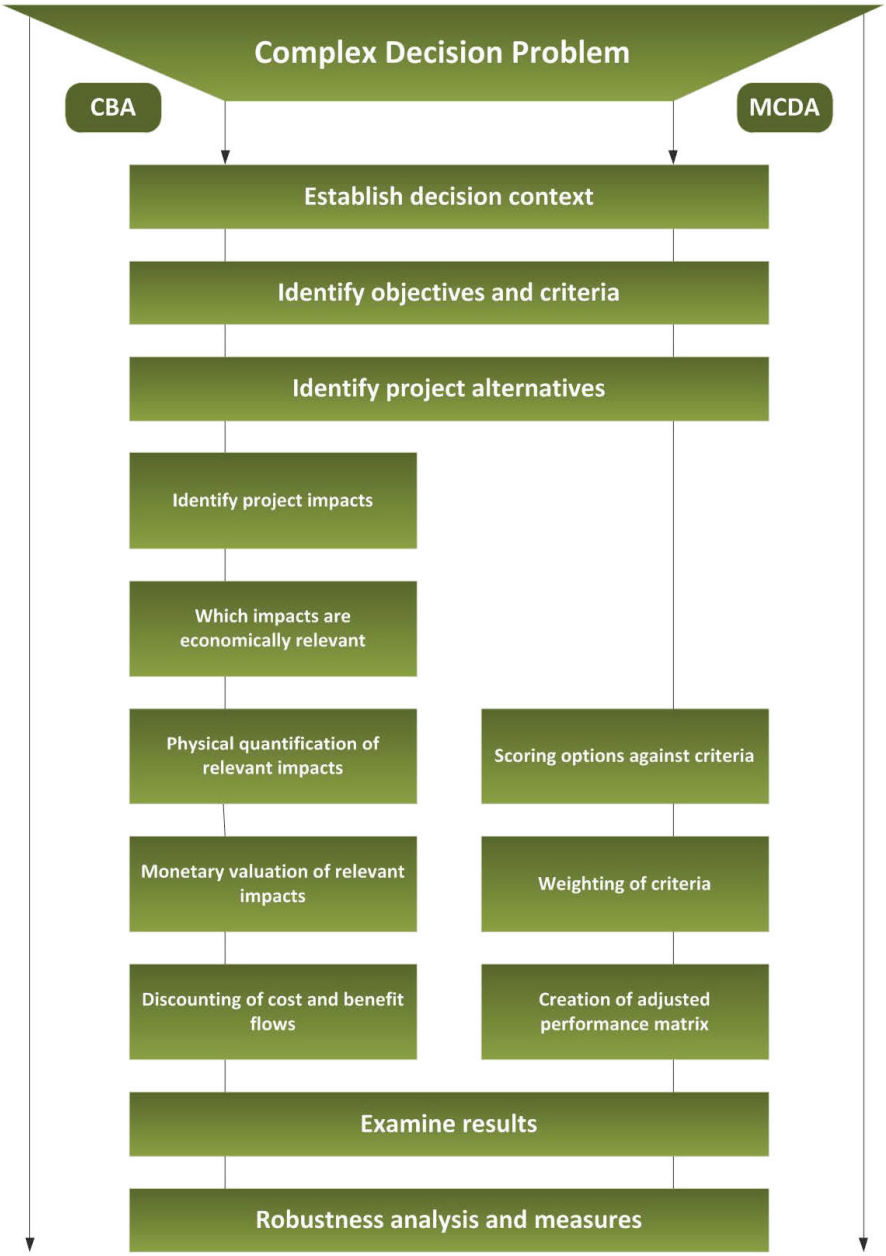


Figure 5-2 The methodological steps of CBA and MCDA in comparison (adapted from Gamper et al., 2006, p. 294)

5.4 Pros and cons for MCDA

As for EIA and CBA, there also exist pros and cons for MCDA:

Pros:

- MCDA provides a transparent and informative decision process.
- Makes it possible to weight objectives and criteria according to preferences.
- Possible to consider a wide range of different criteria (environmental, economic and social).
- Helps to give insight on conflicts and provides ways of reaching compromises.
- Facilitates participation.

MCDA can help provide more insight in the nature of conflicts and into ways of landing political compromises in case of divergent preferences and increasing the transparency of the decision process. The use of MCDA offers the potential to weight objectives and criteria, including changes in weighting over time, hence assisting in the achievement of the 'important' strategic societal goals (Hickman et al., 2012). MCDA makes it possible to consider a large number of data, relations and objectives which are generally present in a specific real-world policy problem, so that the problem at hand can be studied in a multi-dimensional way. MCDA provides a picture from each viewpoint. By combining a tight focus on decision or policy options while at the same time 'opening up' the practical implications of different real-world perspectives, MCDA tries to avoid a serious but often neglected problem suffered by economic, decision and risk assessment techniques, as well as by many more qualitative deliberative and participatory approaches. This problem concerns the way in which such methods attempt to derive a single definitive picture of option performance irrespective of the divergent uncertainties, interests, priorities, and values associated with different expert and socio-political perspectives. Where they are used in this manner to 'close down' policy debates, such methods can undermine underlying principles equally of 'rationality' or 'inclusion' (Stirling, 2008).

Cons:

- Uses subjective qualitative assessments.
- May be perceived as 'Black-box'.
- No consensus on which impacts to include.
- Difficulties in evaluation of unrelated projects.

CBA is carried out from the perspective of the whole society. The perspective is usually a national one. MCDA is less inclusive about 'standing'. Those who are consulted about the performance values or the determination of criteria weights are often experts, focus groups or members of a special interest group. The probability of a result that is biased in favour of a proposal can thus be very high. Also the complexity of some of the MCDA

methods can result in the process being perceived as a black-box and thereby mistrust to the outcome of the MCDA. There exist no single or overriding principles on which the criteria of a project proposal are determined. E.g. the criteria can be selected by the stakeholders in a participative process. This could introduce bias in the MCDA process. An important aspect of public decision making is the opportunity to consider the relative merits of alternative projects. Even though the DMs favour a specific project, the public and opponents will often wish to know what alternatives are feasible and available, or whether the resources could be better used elsewhere. MCDA is incapable of comparisons between unrelated projects, because their impacts or attributes can be very different. In order to use MCDA across different types of projects a common set of criteria would have to be used.

These cons imply that not in all evaluation situations the MCDA is an optimal appraisal method if used as the only appraisal method. However, this is also rarely the case in real-world appraisal of transport projects, here a mixture of different appraisal methods are often applied. As described earlier the CBA seeks to take all relevant impacts into account. However, the monetisation is often problematic and some impacts are presented as very uncertain in the analysis. In these cases the CBA will not fully capture all the externalities and intangible impacts. The CBA can give some information about the efficiency of the assessed project, but it cannot decide whether the project is justifiable from a societal point of view (Macharis et al., 2009). Today not only the economic impacts of a project are important, but also ecological, spatial and social impacts. Including such impacts can be seen as a search for more sustainable solutions. The MCDA can contribute to providing a more comprehensive appraisal using the output from CBA, EIA and additional input. By using the CBA, EIA and the MCDA in combination, the method can give complementary insight in the appraisal. In an MCDA problem, there is no solution optimising all the criteria at the same time and therefore the DM has to find compromising solutions. In the absence of a unique 'correct' policy or alternative as the product, the focus is on the quality of the process. Therefore the use of MCDA in combination with EIA and CBA might improve some of the difficulties these two methods have standing alone in a decision making context.

The use of MCDA within the environmental sector has been a prevalent methodology for some decades now, especially with the rising concern for sustainability. This has also spread to the transport sector and several publications deals with MCDA within this sector (Browne and Ryan, 2011; De Brucker et al., 2011; Haezendonck, 2007; Macharis et al., 2009; Tsamboulas and Mikroudis, 2000). Some countries even have a more or less formal manual for conduction transport appraisal with MCDA e.g. UK and the Netherlands (Department for Communities and Local Government, 2009; Geurs et al., 2009; Janssen, 2001). However, the framework for appraisal in transport planning is still mainly based on CBA, even when the DMs increasingly favour MCDA techniques (OECD, 2011). The conclusions from the sections describing EIA, CBA and MCDA provide some evidence for a possible combination of these methods in a comprehensive appraisal framework. This will in the following chapter be further elaborated.

6 COMPREHENSIVE APPRAISAL

The purpose of this chapter is to examine how a comprehensive appraisal of transport projects can be applied, by the use of already existing and used evaluation techniques.

6.1 INTRODUCTION

CBA is a comprehensible decision making guide based on the concept of economic efficiency, as noted earlier in this thesis. But it is not the only one. If decisions are not made with the help of CBA then they must either be made on an ad hoc basis, i.e. with no particular systemic guidance, or using some other guidance procedure. On the assumption that purely ad hoc decision making is rejected on the grounds that it is not rational, the alternatives to CBA are the other systematic procedures that might be used. The main 'contenders' for rational decision making are (Howarth et al., 2001):

- Environmental Impact Assessment (EIA)
- Multi-criteria Decision Analysis (MCDA)
- Cost-benefit Analysis (CBA)

Table 6-1 presents a comparison of the various evaluation techniques and it shows that they are often complementary. EIA is essential for CBA: all valuation techniques require the physical impact data to work with. Even if techniques appear to be different, they might be profitably combined. E.g. a CBA might be combined with some scoring and weighting procedure (MCDA) for those impacts that are not easily monetised. The interconnections between the techniques indicate that a debate about which technique is better can easily be misleading. The techniques tend to build on each other.

Table 6-1 Relationships between evaluation techniques (adapted from Howarth et al., 2001, p. 177)

	CBA	MCDA
EIA	EIA essential as an input to CBA	EIA needed to provide the impact data for the MCDA. Environmental impacts would be one objective.

As suggested in Table 6-1, EIA provides necessary information for both CBA and MCDA, this implies that an integration of EIA with CBA and MCDA would be beneficial. However, it would be useful to examine a combination of CBA and MCDA in more detail.

One could raise the question whether CBA and MCDA are competing or complementary approaches? From chapter3 and 4 the following can be derived:

- Both are frameworks for assessing options facing DMs.

- Both try to construct a common metric for comparing options.
- Both are sensitive to assumptions; however, different assumptions.
- Are they aids to decision making or a way of decision making?
- MCDA can be seen as incorporating a CBA; CBA is one criterion for assessing options.
- MCDA can define a narrow set of options then subjected to CBA.
- MCDA can sit alongside CBA: one handles measurable costs and benefits, the other handles qualitative cost and benefits.
- MCDA is inferior to the 'objective' CBA approach.

CBA and MCDA share the purpose of assessing alternatives, so DMs can make decisions. They are just two different approaches for fulfilling this purpose. In order to aggregate the different assessments of the impacts/criteria to be able to compare different alternatives, both methods construct a common metric. CBA uses monetary metric and MCDA a score. As with any method, both CBA and MCDA are sensitive to assumptions (see chapter 4 and 5). An open question is if they are aids to decision making or ways of decision making? From the previous chapters CBA seems to be intended to be a way of decision making, however, as also shown previously, this assumption is often not the case when comparing CBA results with actual decisions made. Actually, often the CBA is more a decision aid, or used as screening of projects early in the decision process. Letting CBA be one of many criteria in an MCDA is an approach for combining CBA and MCDA[16]. MCDA can incorporate cost-benefit estimates such as NPV as one element of the multi-dimensional considerations. Using MCDA as a screening technique before submitting them to CBA is an option for a sequential combination of the two methods. Using CBA and MCDA parallel could also be an approach. Here monetary impacts are taken care of by CBA and the qualitative non-monetary impacts can be treated by MCDA. Lutz and Munasinghe (1994) suggested using CBA where it is possible to evaluate environmental management options, and recommended the use of MCDA to analysis consequences that cannot be measured in monetary terms. However, MCDA will not be as objective as the CBA, because input from stakeholders or DMs is needed in order to make an assessment of a project. Thereby some subjectivity is introduced.

Within environmental issues the counter critique of CBA has been that because society uses a monetary metric, so must environmentalist (de Montis et al., 2006). Surely the treasury department of government are often the strongest and lies behind many decisions, especially in a light of the current economic regression, but there are also other branches of government and civil society in which decision processes operate. Other alternatives or supplements to CBA can clearly be most rewarding in terms of gaining insight into value conflict, while aiding and improving policy. However, one should be aware that these alternatives to the traditional CBA lack their own problems and can be difficult to implement successfully. In more general terms, more work is

[16] This approach is used in paper 3 and 4 later in this thesis.

required which breaks down the boundaries between disciplines and more researchers are required who are prepared to venture out of the comfort zones of mono-disciplinarily into challenging areas of inter-disciplinarily (Ibid).

The appraisal framework would need to go beyond the standard cost benefit analysis, as this tends to be heavily weighted to the short term, with strong assumptions on discount rates and in CBA only a limited range of quantifiable measures are used. The multi-criteria analysis, combined with a strong input from a variety of stakeholders as part of a participatory process, could provide a more robust approach. Proceeding to MCDA without CBA could be seen as dangerous, as many MCDAs do not account for public preferences, but uses expert judgement; There are some problems with incorporating time in an MCDA; MCDAs often risk double counting of objective, if one is not very careful in the process of applying MCDA to transport appraisal. (Howarth et al., 2001). From this discussion and previous chapters in this thesis, MCDA seems to be a rational way of presenting alternatives in transport project appraisal. However, this doesn't disqualify CBA, which must be part of the appraisal. In comprehensive appraisal the aim is not simply to combine the CBA, EIA and MCDA approaches, but to integrate them in such a way as to harness and clearly reflect their fundamental co-dependencies.

Different approaches for assessment of transport projects exists in various countries. Many primarily use CBA for evaluating new incentives, even more use EIA along with the CBA, fewer uses MCDA (Goodwin and Persson, 2001; Grant-Muller et al., 2001; Odgaard et al., 2005). By using EIA and CBA as input to an MCDA process, the appraisal makes use of information which for most cases is already available, but just not used in the decision making as an integrated decision-support.

The process of transport appraisal using MCDA and comprehensive appraisal is illustrated on Figure 6-1. First step is to **identify the problem** – what is to be solved? Then a **problem structuring phase** follows, which is one of divergent thinking, opening up the issues, surfacing and capturing the complexity. This is the beginning of managing and understanding how the DMs might move forward (Belton and Stewart, 2002). The phase involves assessment of stakeholders; alternatives; uncertainties; key issues; external environment; constraints, goals and values. For a thorough review of how to structure problems see Jeppesen (2009). After the problem structuring phase the **model building phase** follows. This involves the setup of the model to be used for assessing the projects and include a more detailed specification of alternatives, eliciting values and defining which criteria to include (see Barfod (2012)). Thereafter follows a phase where the **model is used to inform and challenge thinking**. This phase includes a synthesis of the information; challenging intuition; creation of new alternatives; robustness analysis and robustness measures. The output from these three phases form the input to the decision making process. The feedback loops between all of the phases indicate that the process is not necessarily sequential, but can also be an iterative process, where e.g. new knowledge in a phase can lead to a revision

in one of the previous phases. It should be noted here that participation from stakeholders is expected in all three phases.

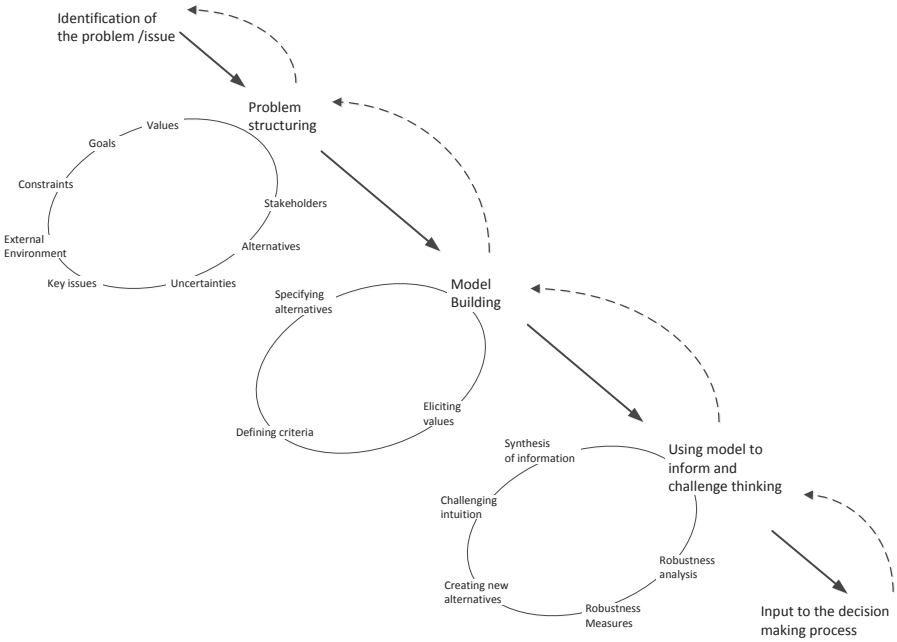


Figure 6-1 The process of transport appraisal using MCDA and comprehensive appraisal (adapted from Belton and Stewart, 2002)

If the interests of the stakeholders are not taken into account, the appraisal will be ignored by the policymakers or be attacked by the stakeholders (Walker, 2000). An analysis of the stakeholders' point of view can show if a certain project or measure will possibly be adopted or rejected by these stakeholders (Macharis et al., 2009). It can be useful for the DMs to test the acceptance of the proposed project or measure before it is proposed for implementation. Therefore it is essential that the comprehensive appraisal also includes a participative approach. By actively incorporating the stakeholders in the decision process it is possible to explore their preferences and take them into account when analysing the decision problem at stake. Furthermore, when actively engaging the stakeholder it is possible to also inform them better about the preceding EIA and CBA. Thereby some of the difficulties found about the stakeholders' ability to relate to the findings of especially CBA may be solved. One approach for such involvement of stakeholders is to use DCs. The concept of DC will be further described in the following example of comprehensive appraisal (section 6.3).

6.2 Applicability in different decision problems

Decision problems come in a wide variety of different types and they may be more or less structured.

If the decision problem can be characterized as a well-structured problem, with clear certainty about the existing knowledge about the problem and a clear agreement about the norms and values in connection to the decision problem e.g. with road maintenance work, then the existing well-founded appraisal methodologies like CBA can be sufficient to appraise the problem. But when there is sensitivity to key assumptions, significant uncertainty and divergent framings of the overall objectives, there are powerful arguments for other appraisal approaches. What can be envisaged is a continuum from well-structured problems to more unstructured ones, where consensus on the relevant parameters are missing, with an increasing need for dialogue and the inclusion of different knowledge (Owens et al., 2004; van den Belt, 2004). So if uncertainty about knowledge base increase or consent on norms and values differ it is necessary to use other means of appraisal, e.g. with traffic safety issues, there are a clear consent on norms and values; all can agree that saving life is important, however there are no clear certainty of how this is best achieved (Figure 6-2). Then the usage of EIA and/or MCDA can add value to the decision process by providing a more comprehensive appraisal and thereby seek to cover the uncertainties at hand on knowledge base and/or the diversity in norms and values.

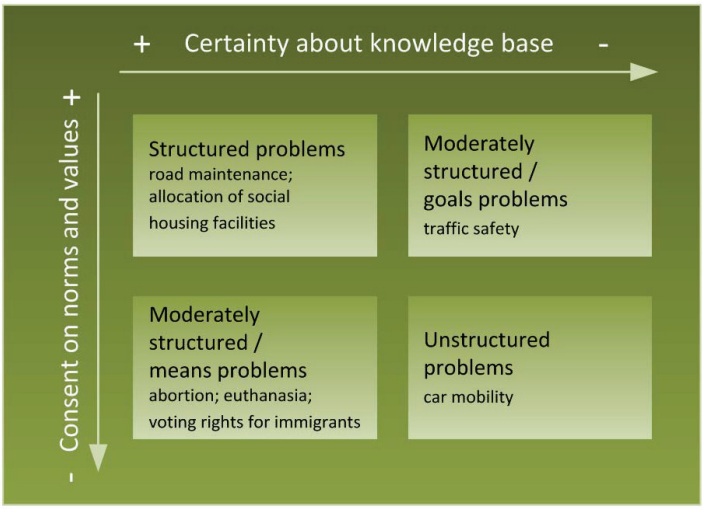


Figure 6-2 Types of appraisal problems (Hoppe, 2002, p. 309)

In other words a comprehensive appraisal is most useful in cases where the correlation between initial model results and holistic judgement is low (moderately and unstructured problems). When the correlation is high (structured problems), the resulting comprehensive appraisal model provides a good description of judgements

that is already so well-understood by the people making them, that the model is of little use in helping them to construct a new understanding of the problem (Phillips, 1984).

In a situation with low correlation or moderately structured problems EIA, CBA and MCDA seem to provide the DMs with decision support that accommodates their need for rational and legitimation in the decision making process. The three methods seem to cover different import aspects of public awareness in decision making.

The example presenting the EM-model in the next section illustrates how such approach can be implemented in practice. The case study is an example of how to involve stakeholders in the decision process. In this particular case one group of experts provided input to the DMs, however this could easily be extended to taking multiple stakeholder groups into account, and thereby examining a wider range of preferences. Furthermore, as presented in the next chapter, robustness analysis can with advantages be implemented to examine other preferences not represented at the DCs conducted in relation to the appraisal. This extension will seek to answer research question C of this thesis: how to apply robustness as a mean to improve the appraisal.

6.3 Example of comprehensive appraisal

In the following an example of an approach to comprehensive appraisal is presented. The example can function as a background for the more specialised models and findings that follow in Part 3 and 4, while at the same time the example represents application of the theoretical issues dealt with in Part 2. It is a book chapter (Jensen et al., 2012) in the book 'Rethinking Transport in the Øresund Region' which has been elaborated in the Interreg IVA project 'EcoMobility'.

The EcoMobility Modelling Framework for Sustainable Transport Planning

Introduction

Addressing the sustainability issues in contemporary decision making is a topic of growing concern. The concept of sustainability necessitates the revision of traditional decision making processes, where the generally acknowledged cost-benefit analysis (CBA) is used for a systematic quantification and comparison of the various benefits and costs generated by a project (Banister and Berechman, 2000). Decision making based on CBA is often found to be inadequate to incorporate and assess multiple, often conflicting objectives, criteria or attributes like environmental or social issues which are usually intrinsically difficult to quantify (Beukers et al., 2012; Thomopoulos et al., 2009). Therefore, it is necessary to broaden the decision making process beyond the consideration of merely economic factors. The ultimate implementation of such a decision making framework under the multiple criteria will require multi-disciplinary and multi-participatory approaches, especially when there is need for assessing a decision problem from different perspectives, e.g. a sustainability perspective (Banister, 2008). These issues are of great concern for the EcoMobility project, which

seeks to reduce the environmental impact of transport while increasing economic growth and accessibility in the Øresund region.

This subsection introduces the EcoMobility (EM) modelling framework developed within the decision modelling group of DTU Transport for the evaluation of complex transport infrastructure decision problems. The EM modelling framework consists of two parts, namely a decision conference and an Excel-based software model (entitled the EM-model). The latter employs the use of two multi-criteria decision analysis (MCDA) techniques namely REMBRANDT (ratio estimation in magnitudes or deci-bells to rate alternatives which are non-dominated), which is based on pair wise comparisons, and SMARTER (simple multi-attribute rating technique exploiting ranks), which is based on criteria rankings. The concept of a decision conference (DC) is introduced in order to formalise and operationalise group processes that enable the assessment of non-quantifiable impacts/criteria within a decision support context.

For illustration, the application of the model is presented by a case study in the Øresund region considering the alternatives for a new fixed link between Helsingør (Elsinore) in Denmark and Helsingborg in Sweden (referred to as the HH-connection), where a sustainable solution for the connection was found. The case study has proved the decision support system to be a valid and useful tool for making decisions under complex circumstances of multiple objectives, conflicting interests and involvement of different stakeholders.

The subsection is organised as follows: after this short introduction the overall modelling framework is introduced more specifically in terms of small introductory paragraphs to the various methodological approaches of decision conference, REMBRANDT and SMARTER. Hereafter, the case study is presented and used to demonstrate the applicability of the EM modelling framework for sustainable decision making. The subsection ends with a discussion on how the model can support and validate the research findings about sustainable transport in the Øresund region as well as the main conclusions and perspectives for the future research and application tasks.

The EcoMobility modelling framework

The main scope of the EM modelling framework is to assist the DMs in assessing complex decision problems, which usually involve multiple and often conflicting objectives. The main aim is to allow for stakeholders involvement in order to obtain informed and transparent decision support. The EM modelling framework is illustrated in Figure 1, where the decision conference and the EM-model are separated into two independent items aggregating into a set of resulting total scores for a given set of alternatives within a project scope.

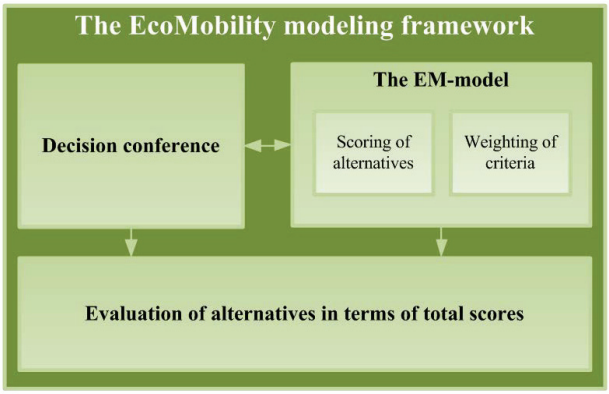


Figure 6-3 The EcoMobility Modelling framework

From Figure 6-3 it is given that the decision conference and the EM-model are interacting in terms of producing a set of total scores i.e. prioritisation between a set of pre-defined project alternatives. The following paragraphs summarise in short the approach of respectively the decision conference and the EM-model containing of the two subsequent techniques REMBRANDT for scoring of alternatives and SMARTER for weighting of criteria.

Decision conference

The overall purpose of a decision conference (DC) is to combine technical and social aspects for solving different types of complex decision problems. It enables a structured debate between the participants that are either involved in and/or affected by the decision problem. The debates, evolving between the participants representing different perspectives on the problem, are able to enrich the basis on which the decisions have to be made. The goal of a DC is to develop a common understanding of the decision problem between the participants, to create a sense of common purpose and achieve a group commitment (Phillips, 2007). A decision achieved in consensus at a decision conference has a higher possibility for being implemented than e.g. the results from a complex decision analysis that only involves one DM who later has to justify his decision to an organisation or to the general public. But the final decision is still in the hands of policy-makers and based on many other factors (Zurita, 2006).

The concept of a DC consists of three main components: group processes, decision analysis and information technology (the EM-model). The group processes are assisted by an impartial facilitator who is guiding the participants though the DC. The main task of the facilitator is to ensure that all the participants get a chance to express their opinions (Franco and Montibeller, 2010). The decision analysis is supported by the software model which is constructed and run by a decision analyst who on-the-spot collects the relevant data and judgments of the participants (Phillips and Bana, 2007). In principle, the model represents the collective view of the group at any point of time

along the process allowing the participants to understand each step such that no black-box process/solutions should occur and only confident results be achieved.

The five-step procedure contained within the DC is formulated in Figure 6-4 in order to lead the participants through the decision process and motivate them to produce the input needed for the assessment. The arrows pointing back from step 5 indicate that it is possible to go back in the process and redo the assessments made in step 3 and 4 if a shared understanding has not been achieved (Barfod, 2012).

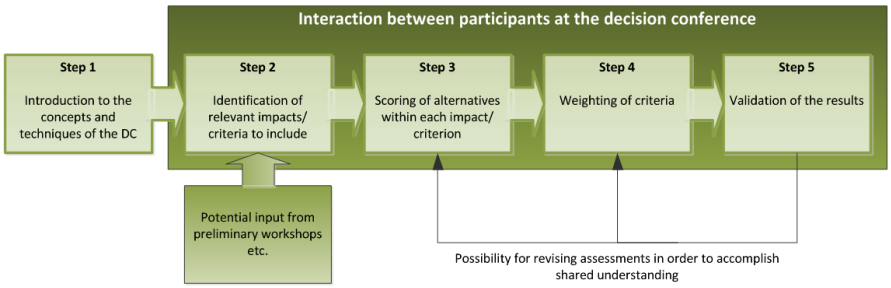


Figure 6-4 The five steps in the decision conference (adapted from Barfod, 2012)

In detail the steps are as follows (Ibid):

1. *Introduction to the concepts and techniques of the DC*. It is very important for the quality of the decision conference that the facilitator starts by introducing the concepts and methods being used in simple terms. When the participants know the basic characteristics of the model, they will be more comfortable with the later process.

2. *The identification of relevant impacts for the assessment*. Before any ranking can be carried out, the alternatives have to be characterized by a number of decision criteria, e.g. technical, economic, environmental, etc. In this respect it can be very useful to conduct a workshop already in the initial planning phase, where issues regarding the project can be discussed and criteria with influence on the decision making can be developed. It is up to the participants to structure and reduce the criteria into a number of relevant criteria which all contribute to the differentiation between the alternatives. However, it is up to the facilitator to ensure that no criterion overlaps with other criteria in order to avoid double counting.

3. *The scoring of the alternatives within each impact/criterion*. The relative scores of alternatives are determined through the pair wise comparison mechanism, where the alternatives are compared two by two under each criterion and assigned with numerical values based on preference intensity.

4. *The weighting of the criteria*. The criteria are weighted using rank order distribution weights (ROD) embedded within the SMARTER approach. It is considered to be very difficult to make the participants agree directly upon a weight set of criteria. Instead the model contains the possibility to examine all the different weight sets provided by each participant individually.

5. *Validation of the results*. The participants will after this process be able to make their choice based on a broader basis of knowledge, as they now are aware of the other participants' viewpoints and can take them into account.

Consequently, DCs are conducted "here-and-now" preferably without any formal presentations. Each DC may proceed differently and have different outlines/outcomes; however, the four basic stages that typify most of them can be outlined as follows (Ibid):

1. Broad exploration of the issues.
2. Construction of the model based on each participant's judgments about the issues.
3. Exploration of the model.
4. Summary of key issues, conclusions and agreement about the way forward.

At the beginning of the DC the data are introduced and the issues are tackled in line with the different opinions actively sought to encourage the debate. Based on that, input is given to the model and participants examine the results and their realism and consistency. New perceptions about the issues might lead to successive revisions. This iterative and interactive process proceeds until a shared understanding of the issues is obtained. Finally, the commitment of the participants to the way forward is agreed upon.

The EcoMobility-Model

As presented the third component within the settings of a decision conference are the application and introduction of a software system capable of arranging and ultimately assessing the inputs brought forward by the various participants of the DC. The EM-model is capable of handling up to eight alternatives with up to ten criteria. During the interaction with the participants the alternatives are compared two by two assigning them with a score within each criterion, the criteria are assigned with weights according to their relative importance and the results are found by aggregating the preference information. Figure 6-5 presents the flow of the model embedding the five steps presented in the DC procedure.

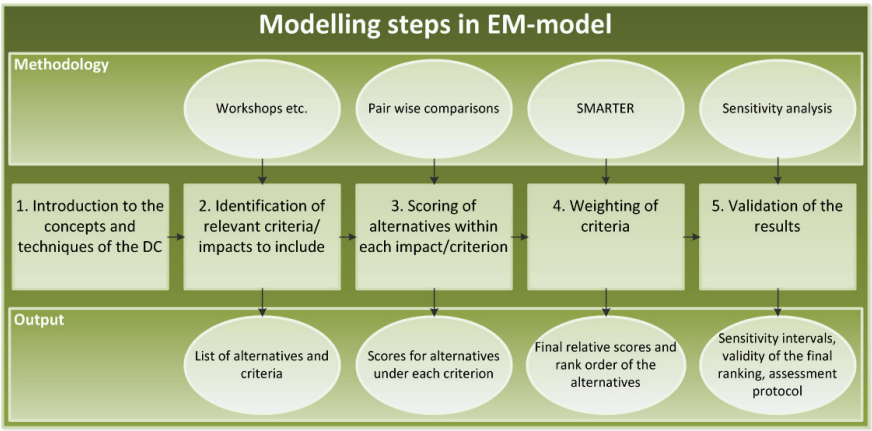


Figure 6-5 The modelling steps of the EM-model

The process-related steps to be followed in the EM-model in order to conduct the assessment are shown in Figure 3. After the introduction to the DC, the information in relation to the alternatives that are formulated to remedy the problem and the criteria that are developed relevant to the assessment of alternatives is firstly fed into the model. Secondly, the alternatives and criteria to include in the assessment are listed. Thirdly, the EM-model makes use of the REMBRANDT approach (see section 2.2.1) for scoring the alternatives and measuring the contribution of each alternative to a specific criterion. The relative score given to each alternative is determined by comparing all the alternatives pair wisely under each of the criteria. Fourthly, the EM-model requires the determination of the criteria weights which currently are performed by the use of the SMARTER approach. Fifthly, the information is aggregated into single value measures resulting in the total scores making it possible to define a prioritised list of the alternatives. These total scores indicate the degree to which the alternatives contribute to the problem solution. Finally, the EM-model performs sensitivity analyses where it tests whether the final ranking would be different if the weighs of the criteria are changed.

The REMBRANDT approach

The EM-model involves the use of a structured hierarchical technique named REMBRANDT by (Lootsma, 1992) which is designed to evaluate a finite number of alternatives under a finite number of conflicting criteria by a single stakeholder or a group of stakeholders.

In order to assess the project alternatives (make a prioritised list of preferred alternatives) the REMBRANDT approach for pair wise comparisons has been applied. The approach is a multiplicative version of the Analytic Hierarchy Process (AHP) developed by Saaty in (1977) which attempts to overcome some of the theoretical difficulties associated with the original AHP (Barfod, 2012; Belton and Stewart, 2002).

The applicability of REMBRANDT is based on three parts: decomposition, comparative judgment and synthesis of priorities.

The decomposition part requires decomposing the decision problem into a hierarchy that reflects the essential elements of the decision problem dealt with: an overall objective or goal at the top level, the criteria (sub-objectives) that define the alternatives at the middle level and finally, the competing alternatives at the bottom level of the hierarchy. The principal structure of such a hierarchy is presented in Figure 6-6.

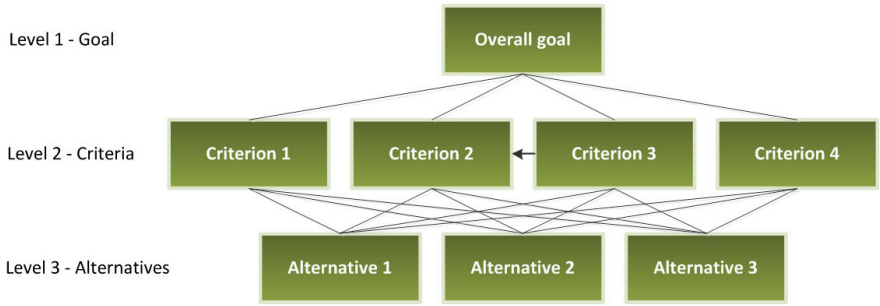


Figure 6-6 A decision hierarchy

The comparative judgment principle requires pair wise comparisons between the decomposed elements within a given level of the hierarchical structure with respect to the next higher level. Thus, the pair wise comparisons have to be made between the alternatives to determine their impacts under each criterion and between the criteria to determine their relative importance to the overall goal (Figure 6-6).

Finally, the synthesis principle requires aggregating of the results derived at the various levels of the hierarchy in order to construct a set of priorities for the elements at the lowest level of the hierarchy allowing a rank ordering of the alternatives.

When the decision problem in hand is to be assessed using the REMBRANDT approach it is beneficial to have a group (participants of the DC) to make the assessment. A finite number of pre-selected alternatives A_1, A_2,..., A_n (Level 3), thus, are pair wisely compared against a set of predefined criteria (Level 2). During the process participants are presented with each pair of alternatives A_j and A_k under a specific criterion and asked to express their preference for one alternative over another. The strength of such a procedure lies in terms of the preference information which is collected in terms of verbal statements as denoted in Table 6-2 that in turn are corresponding with a numerical value to be entered in the EM-model and processed using the mathematical principles behind the REMBRANDT approach.

Table 6-2. The REMBRANDT intensity scale for comparing two alternatives A_j and A_k (Lootsma, 1999)

Verbal description	Numerical value
Very strong preference for alternative A_k	-8
Strong preference for alternative A_k	-6
Definite preference for alternative A_k	-4
Weak preference for alternative A_k	-2
Indifference	0
Weak preference for alternative A_j	+2
Definite preference for alternative A_j	+4
Strong preference for alternative A_j	+6
Very strong preference for alternative A_j	+8
For the compromise between the neighbouring values	-7, -5, -3, -1, +1, +3, +5, +7

All the information about the pair wise comparisons conducted and the participants' arguments regarding them during the decision making process must be documented in an assessment protocol. This can be valuable to justify the decision and also it can be useful if the process is going to be repeated after some time.

When the relative scores of the alternatives under the criteria are determined through the pair wise comparisons, the criteria should be weighted in order to synthesize all the scores. The criteria can be weighted using different techniques, such as pair wise comparisons or the SMARTER (simple multi-attribute rating technique exploiting ranks) technique with the ROD (rank order distributions) weights.

The SMARTER approach

In order to simplify the process of eliciting criteria weights the SMARTER approach has been proposed by Edwards and Barron (1994). Using SMARTER the participants of the DC place the criteria into an importance order: for example 'Criterion 1 (C1) is more important than Criterion 2 (C2) which is more important than Criterion 3 (C3) which is more important than Criterion 4 (C4), and so on, C1 ≥ C2 ≥ C3 ≥ C4... The SMARTER approach then assigns surrogate weights to the criteria based on this ranking. A number of methods that enable the ranking to be translated into surrogate weights have been developed. These are among others Rank Order Centroid (ROC), Rank Sum (RS), RR Rank Reciprocal (RR) and Rank Order Distribution (ROD) weights. Roberts and Goodwin (2002) have examined these methods in details and found that ROD weights seem to provide the best approximation of the participants' preferences.

ROD is a weight approximation method that assumes that valid weights can be elicited through direct rating. In the direct rating method the most important criterion is assigned a weight of 100 and the importance of the other criteria is then assessed relative to this benchmark. The ROD weights for between 2 and 10 criteria are shown in Table 6-3.

Table 6-3 Rank Order Distribution (ROD) weights (Roberts and Goodwin, 2002)

	Criteria								
Rank	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	0.6932	0.5232	0.4180	0.3471	0.2966	0.2590	0.2292	0.2058	0.1867
2	0.3068	0.3240	0.2986	0.2686	0.2410	0.2174	0.1977	0.1808	0.1667
3		0.1528	0.1912	0.1955	0.1884	0.1781	0.1672	0.1565	0.1466
4			0.0922	0.1269	0.1387	0.1406	0.1375	0.1332	0.1271
5				0.0619	0.0908	0.1038	0.1084	0.1095	0.1081
6					0.0445	0.0679	0.0805	0.0867	0.0893
7						0.0334	0.0531	0.0644	0.0709
8							0.0263	0.0425	0.0527
9								0.0211	0.0349
10									0.0173

The use of ROD weights goes some way to reduce the value problem of having criteria with very low weights in the assessment. However, it can be argued that the inclusion of criteria with very low weights, e.g. 0.02, does not contribute in any way to the overall result and, therefore, should be omitted from the analysis, see Barfod (2012) for a discussion of this.

It should be noted, that the four decimals that are shown for the ROD weights in Table 6-3 express a much higher accuracy in the weights than should be expected in practice. Normally, the participants assign weights with not more than two decimals as this seems to be the limit to what can be comprehended by the human mind without difficulties. Thus, the weights in Table 6-3 should be presented with only two decimals to the participants if this technique is used in the decision process.

Evaluation of alternatives: Results

After deriving the separate scores for the alternatives within all the assigned criteria and furthermore determined weights for the criteria – it is possible to produce a result based upon the participants at the decision conference.

It should be noted, that if the participants in the DC do not agree with this result it is possible to go back in the process and, hence, revise the assessments or perhaps test the various weight settings applied.

On this basis decisions made in consensus at a DC seem to have a fairly higher probability for being implemented than results from another more standardised evaluation approach only involving one DM who then later has to justify the choices or causes of action during the calculations. Moreover, decisions made by group effort have better terms for working in practice both in terms of operability (since the group somehow is committed) but also in light of project ownership.

The following section enhances the perspective of the EM modelling framework in terms of a presented case study with respect to a second fixed link between Denmark and Sweden, the so-called HH-Connection.

The case study: HH-connection

Eleven years have passed since the Øresund fixed link connecting Copenhagen in Denmark with Malmö in Sweden opened to traffic (see Figure 6-7). The fixed link between Zealand and Scandinavia has led to a strong increase in traffic across Øresund as a whole. In 2009, an average of 19,500 vehicles and 184 trains crossed the link per day, corresponding to 141% and 125% increase respectively compared to the first full year of operations in 2001 (Øresundsbro Konsortiet, 2010). Moreover, the fixed link across Fehmarn Belt between Denmark and Germany which is expected to open in 2018, will increase these numbers due to more travellers from central Europe through Denmark to the rest of Scandinavia (Sweden and Norway). Especially, the number of freight trains through Denmark is expected to grow significantly, turning the Øresund fixed link into a bottleneck, because the existing capacity is already close to the limit.

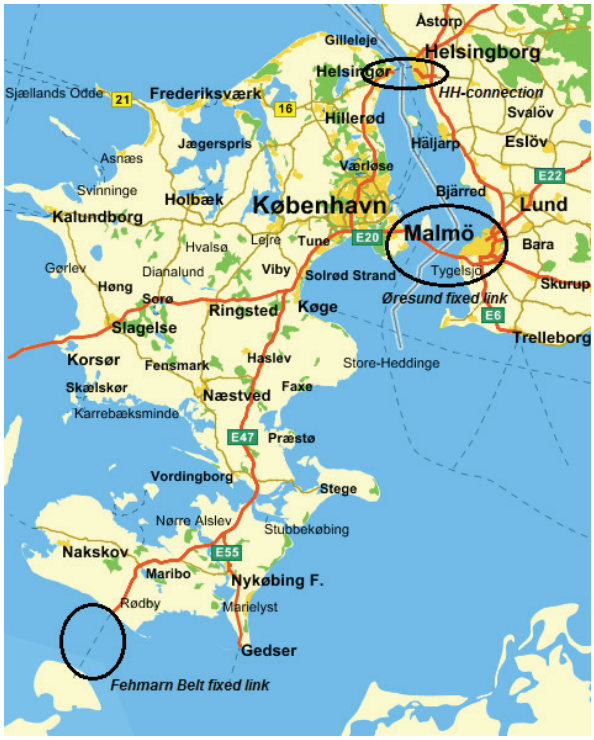


Figure 6-7 The proposed new fixed link (HH-connection), the Øresund fixed link and the forthcoming Fehmarn Belt fixed link (from map.krak.dk)

The proposal of a fixed link between Helsingør (Elsinore) in Denmark and Helsingborg in Sweden – referred to as the HH-connection – has been considered since the 1980s, however, the opening of the Øresund fixed link postponed the planning and implementation. The case is now re-actualised in order to cope with the increasing traffic across the Øresund and the planned Fehmarn fixed link. A new northern fixed link would reduce the travel time between Zealand and Scandinavia, relieving the Øresund fixed link for some of the car and railway traffic. In the autumn of 2011 three tunnel alternatives are considered as main alternatives for the HH-connection. The alternatives are listed in Table 6-4 with indication of type of construction, type of traffic and total costs (in mill. DKK).

Table 6-4 The three proposed alternatives for the HH-connection (Larsen and Skougaard, 2010)

HH-connection	Description	Cost (mill. DKK)
Alternative 1	Tunnel for rail (2 tracks), passenger trains only.	9,500
Alternative 2	Tunnel for rail (2 tracks), passenger trains only + tunnel for vehicles (2 × 2 lanes).	24,500
Alternative 3	Tunnel for rail (2 tracks), passenger trains + tunnel for vehicles (2 × 2 lanes) + tunnel for rail (single track), goods trains.	32,500

The alignments of the different alternatives depending on the type of traffic are shown in Figure 6-8.

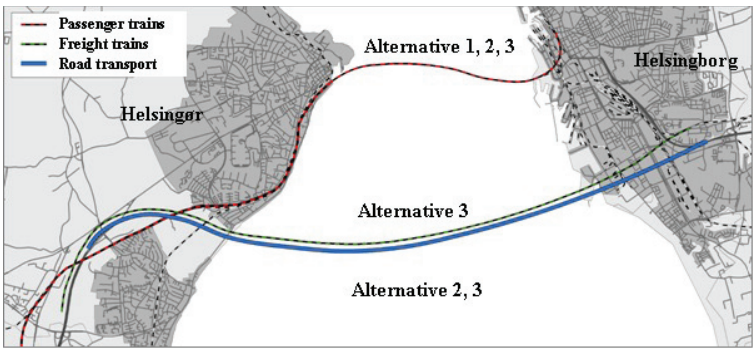


Figure 6-8 The alignment of the alternatives

The overall goal of this case study is to find not only the socio-economically most sound, but also the most sustainable alternative for both passenger and freight transport. Based on this, the alternatives have been assessed against a set of decision criteria as follows:

- *C1: Socio-economic robustness*. The criterion embraces the overall economic performance of the alternative. The main indicator is the certainty value (CV) calculated based on the results stemming from the cost-benefit analysis (CBA). Risk analysis (on construction costs and time savings) is applied to the CBA and the resulting CV describes the certainty for obtaining a benefit-cost rate (BCR) above 1 (Salling and Leleur, 2011b). A high CV is, therefore, preferable.

- *C2: Improvement for passenger cars and public transport*. The criterion emphasises the accessibility for both cars and public transportation. This is represented by the increased mobility potential that the commuters obtain (they can cover more geographic space using the same time as previously).

- *C3: Impact on towns and land-use*. The criterion first and foremost emphasises the visual environment in the towns of Helsingør and Helsingborg. The form of the land-based facilities and their geographical placement will for this reason be in focus. Moreover, the housing prices in the towns will most likely be affected as well, especially, for the houses close to Øresund and the land facilities.

- *C4: Impact on regional economics*. The criterion considers the alternatives potential for contributing to the economic development in the Øresund region. In order to obtain economic development in the northern part of the Øresund region the area should become more attractive both to housings and businesses. If more businesses are located in the area it becomes more attractive to live there which also creates the basis for more shopping opportunities. In this way, the potential for gathering businesses and creating a specialised business life increased. A new connection should moreover make it more attractive for students to live in the area as it will become easy to cross the sound for which reason the universities in Helsingborg, Lund and Copenhagen will be within a reasonable travel distance/time.

- *C5: Impact on flexibility in logistics*. The criterion covers the impact on the efficiency, punctuality, security, co-modality and risk in the logistic chains. Relocation of warehouse facilities reflects the benefits that arise when companies reduce e.g. their number of warehouses because the new infrastructure makes it possible to serve customers from fewer warehouses. Moreover, a new connection can help to expand the companies' clientele, and at best, it can result in that some companies can close down a production area, thereby, saving money.

- *C6: Contribution to the EU green corridors*. The criterion emphasises the alternative's potential for promoting the green transport corridors. According to the Danish Transport Authority (2011) Green Corridors are a European concept denoting long-distance freight transport corridors in which advanced technology and co-modality are used to achieve energy efficiency and reduce environmental impact. Launched in the Freight Transport Logistics Action Plan (2007), Green Corridors support the EU's agenda towards decarbonising transport while emphasising the need for efficient logistics. The existing Øresund fixed link is at the current time a part of the EU east-west green corridor. However, a north-south green corridor is also going to pass through Denmark and Sweden, and the only way to relieve the pressure on the Øresund

fixed link in order not to exceed the capacity limit is to place a new connection between Helsingør and Helsingborg.

Application of the EM modelling framework to the HH-connection study

A decision conference was conducted concerning the HH-connection decision problem, where the purpose was to locate the most attractive alternative amongst the three presented in Table 3. Four experts joined the group from DTU Transport on 6th October, 2011 in order conduct the decision conference and ultimately create a decision support base with regard to the prioritisation of the alternatives. The conference was guided by a facilitator supported by a model analyst using the EM-model to perform on-the-spot modelling of the information obtained from the participants.

After the introduction and the information regarding the alternatives and criteria, the first task for the participants was to rate the alternatives under each defined criterion. The ratings were done using pair wise comparisons, where the participants had to state their preferences for one alternative over another in a comparison under the criteria one by one. The verbal information was then converted into numerical values based on the intensity scale from 0 to 8 (see Table 6-2) and filled into the comparison matrices in the model. An example of such a matrix is shown in Table 6-5, where the results of the pair wise comparisons of the alternatives for the socio-economic robustness criterion are presented as they were agreed upon by the four participants.

Table 6-5 The comparison matrix for the socio-economic robustness criterion

Criterion 1:	Socio-economic robustness			
	Alt1	Alt2	Alt3	Score
Alt1	0	-6	-4	0.10
Alt2	6	0	3	8.00
Alt3	4	-3	0	1.26

The idea of incorporating an economic criterion in the MCDA is based on the SIMDEC (risk simulation and multi-criteria decision analysis in combination for decision support) approach (Leleur, 2012). In SIMDEC, the MCDA includes the results of the risk analysis on the possibility of not obtaining socio-economic feasibility examined for each of the project alternatives as one criterion. This criterion concerns how each alternative is affected by the uncertainties underlying the two major impacts of every large transport infrastructure project: construction costs and travel time related benefits (Salling, 2008).

Afterwards, the participants were asked to rank the criteria, first, individually then as a group. The criteria were then assigned with the surrogate ROD weights based on the rankings, see Table 6-3. The group's joint ranking of criteria and the assigned weights are shown in Table 6-6 below. Note that the ROD weights are only shown with two decimals here.

Table 6-6 The DC participants' joint ranking of the criteria

Criteria	Rank after importance	Weight
C1: Socio-economic robustness	2	0.24
C2: Improvement for passenger cars and public transport	3	0.19
C3: Impact on towns and land-use	6	0.04
C4: Impact on regional economics	1	0.30
C5: Impact on flexibility in logistics	5	0.09
C6: Contribution to the EU green corridors	4	0.14

All the information fed into the model was aggregated to obtain the resulting total scores for each alternative. The results are shown in Figure 6-9 with alternative 3 as the most attractive, while alternative 1 only achieved a very low score.

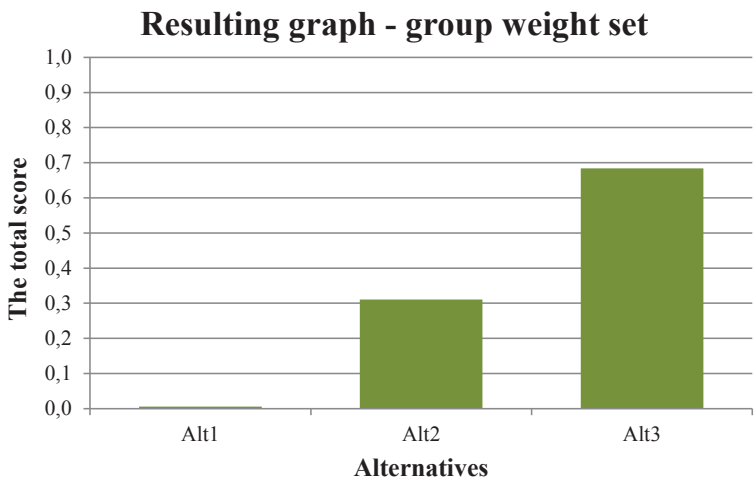


Figure 6-9 The resulting graph showing the total scores for each alternative

It is worth noticing that similar results were produced applying the four individual weight sets even though the participants had very different perspectives on the weighting of criteria. Three of the participants had the socio-economic robustness criterion as the most important criterion, but one participant had the criterion ranked as the least important criterion. The argument for the low priority of this criterion was that all of the alternatives were socio-economically feasible and, therefore, this criterion was not given a high priority. A conducted sensitivity analysis revealed that the weights of all the participants were within the interval for the socio-economic robustness criterion and, thereby, it made no difference for the overall ranking of

alternatives which weight profile was applied. This also applies to the other criteria weights. This means that the participants at the DC achieved a common vision on the decision problem with alternative 3 as the most preferred for the HH-connection.

Discussion and conclusion

The modelling framework described in this subsection attempts to encompass a wider set of criteria in transport planning than a traditional cost-benefit analysis. Sustainable transport planning in the Øresund region necessitates a broader decision support tool that is capable of taking into account and assessing the multiple and often conflicting criteria and objectives which are difficult to measure in monetary terms. Thus, the use of pair wise comparisons (REMBRANDT) makes it easier for the participants to assess how the alternatives perform under each criterion and assign them with values enabling a ranking of alternatives. By selecting appropriate criteria it becomes possible to express sustainability in operational terms for actual decision support.

The case example shows that it is possible to take into account a wide range of criteria such as environment, economic growth and accessibility in the same decision support model. The EM-model is not only a multi-disciplinary, but also multi-participatory decision support model. Several stakeholders can be included in the assessment as recommended by e.g. Musso et al. (2007) and Macharis (2007) and in the case study four different participants in the DC gave input to the EM-model. The setup of the DC should obviously depend on the involved participants, but more importantly be based upon the decision problem to be investigated. Therefore, a major obstacle is the ability to convey appropriate methodological approaches that provide a theoretically approved course of action while at the same time maintain its transparency and applicability.

It is planned that the EcoMobility modelling framework should also be used for the localisation problem of an Urban Consolidation Centre (UCC) in Copenhagen. An UCC initiative for Copenhagen has been proposed aiming to remedy the urban freight transport in the city centre of Copenhagen. Hence, the localisation of such an UCC is a multi-disciplinary and multi-participatory decision problem, since technical, social, economic and environmental criteria are all important for the choice of localisation. The technical criteria consist of logistics, level of service to be provided, volume and the trip effectiveness. The social criteria are mobility, accessibility, quality of life and traffic safety. The economic criteria are related to the direct influence on benefits and costs, while the environmental criteria cover congestion, air pollution, noise and energy consumption. These numerous and often conflicting criteria together with a range of different stakeholders and their different preferences make this localisation problem an obvious task for the future application of the EM modelling framework.

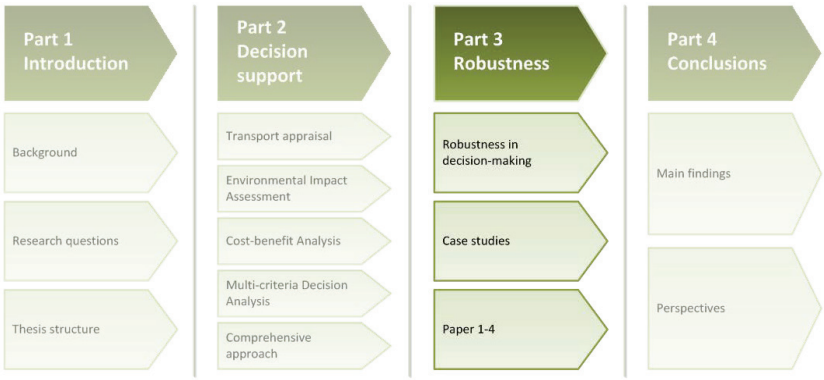
A key concern within the EcoMobility project is to identify the effective means for sustainable transport planning in the Øresund region. This complex challenge can be met as concerns assessment methodology with the multi-faceted EM-model, which involves risk analysis on the socio-economic part of the decision problem, MCDA to

embrace various and often conflicting criteria and sensitivity analysis for taking into account the interests and preferences of different stakeholders. To optimise the use of the EM-model customised decision conferences become essential where the engagement of stakeholders and their different preferences provide a common platform for understanding a decision problem and for seeking out the most attractive decision alternative.

Overall, the EM modelling framework provides a theoretically sound, and at the same time, practical and effective decision support tool for planning sustainable transport in the Øresund region.

Part III

Robustness measures in decision making



7 ROBUSTNESS IN DECISION MAKING

The purpose of this chapter is to introduce how to assess robustness of an appraisal of an alternative. This chapter will present the concepts of robustness analysis in connection with sensitivity analysis along with an overview of various approaches for doing this. The chapter ends with a formulation of robustness measures to be used for assessing the robustness.

7.1 INTRODUCTION OF ROBUSTNESS MEASURES

As stated earlier, sensitivity (SA) and robustness analysis (RA) should be conducted in order to examine whether preliminary conclusions are robust or if they are sensitive to changes in aspects of the model. The two terms SA and RA are sometimes used interchangeably, but it should be noted here that RA means an inversion of perspective in relation to SA. Thus rather than knowing how the parameter values in RA could vary without changing the conclusion, the purpose of SA is to know how the conclusion changes (or doesn't) when the parameter values change (Dias and Clímaco, 1999). At the same time RA can be viewed more broadly as a term where not only different types of SAs are considered but also aspects linked to the methodologies and processes in a comprehensive assessment.

The significance of missing information, exploration of the effect of a DM's uncertainty about values and priorities or to offer a different perspective on the problem, changes can be made investigated by making changes. However, there may be no practical or psychological motivation for changing value; the exploration may be driven simply by a wish to test the robustness of the results. SA can determine whether the main results of a ranking system change substantially when assumptions are varied over a reasonable range of possibilities. In this respect the robustness of the ranking can be assessed by studying its sensitivity to three types of technical uncertainties: a) the indicator set, b) the aggregation rule and c) the set of weights (Munda, 2008).

French, Maule and Papamichail (2009) argue that SA can support the decision process in eight ways: exploration of consequence models, exploration of the relationship between science and the consequence model, support the elicitation of judgmental inputs, development of efficient computational algorithms, design of surveys and experiments to gather relevant data, guidance for inferences, forecasts and decisions, communication and consensus building, and development of understanding.

All of the above listed reasons for SA are relevant for public decision, but especially communication and consensus building are important. In order to get both public and political support for the decisions it is essential that the foundation for the decision making process can be made transparent and thereby easy to communicate. Furthermore the existence of several different preferences calls for a methodology to examine how the decisions are sensitive to differences in preferences.

There are several other reasons for conducting SA especially in an MCDA framework. One reason is because of the nature of an MCDA process, which inherently contains various levels of uncertainty because of qualitative and subjective choices of different parameters. In fact MCDA has been criticised for not being an exact procedure and therefore it is very important to test how robust the results are. Secondly, the SA of MCDA procedure enables the data and the decision making problem to be explored at greater depths. This provides insight into the nature of the decision problem; it unravels its complexities, and may even provide recommendations for future analysis. Furthermore, a SA may be carried out in order to deal with the uncertainty in estimation of some of the input figures. The DMs may not be able to derive at a set of criteria weights and may provide a range of weightings or the decision might involve various groups of stakeholders with different preferences towards the weights which can be analysed. Furthermore, the impacts of various options under different criteria may fall within a statistically estimated range that can be incorporated into the analysis (Proctor and Qureshi, 2005).

As MCDA includes a comprehensive process involving a rich interaction between human judgment, data analysis and mathematical/computational processes the need for SA is clear. Errors and unintentional biases can enter at any of these stages, and it is the process as a whole that needs to be robust. Perhaps some of the key points at which such errors and biases may intrude would be the following (Stewart, 2008): external uncertainties, internal uncertainties, choice of model and identification of criteria and alternatives.

Belton and Stewart (2002) divide SA into three different perspectives:

- *Technical perspective*: The objective examination of the effect on the output of a model changes in input parameters of the model. A technical SA will determine which, if any, of the input parameters have a critical influence on the overall evaluation e.g. where a small change in a criterion weight or an alternative's score can affect the overall preference order.
- *Individual perspective*: The function of an individual SA is to test the individual's intuition and understanding of the problem. Do they feel comfortable with the results from the model? If not, why not? Has an important criterion been overlooked?
- *Group perspective*: Within the group context, the SA is a tool for examining other perspectives on the problem, often by a different set of CWs. This makes it possible to test the overall evaluation for different perspectives e.g. economical or environmental perspectives.

In addition to the technical perspective are the uncertainty different models introduce. Applying the same decision problem with the same input data, doesn't necessarily give the same result when applied to different models or methodologies of appraisal. Therefor the choice of model for appraisal introduces uncertainty which is also necessary to deal with. The technical perspective does not consider the input from the

DMs or stakeholders, but merely performs an analysis of the input parameters influence on the overall evaluation. The individual perspective seeks to test if the outcome of the model is in accordance with the intuition and preferences of an individual. If the outcome of the model is not understood and acceptable for the individual a criterion might be missing in the model or some of the performance scores are not assessed in line with the preferences of the individual. Lastly the group perspective involves analysing the sensitivity of introducing other perspectives/preferences to the model than those initially incorporated in the assessment. In public decision making this could be of high importance in order to gain public acceptance of the appraisal. If the DMs can test the outcome of the model in the view of different perspectives by the use of sensitivity analysis and thereby provide a robustness measure for the alternatives, they will accommodate the most likely following discussions of the recommendations from the appraisal. Furthermore, such robustness measures can inform the DMs of the sources for possible divergences of results because of different perspectives applied. A more extensive treatment of these sources can reveal if any acceptable changes in the e.g. assessments of the alternatives or criteria can be made, and thereby provide a more robust appraisal result.

Often there is an informal approach to SA in application of MCDA. However, in the literature there are numerous examples of formal approaches to SA (Hyde, 2006). The formal sensitivity analysis methods can be classified in a variety of ways. Frey and Patil (2002) classify SA methods as:

- Mathematical: Methods that assess the sensitivity of a model output to the range of variation of an input.
- Statistical: Running simulations in which inputs are drawn from probability distributions and the effect of variance in input on the output distribution is assessed.
- Graphical: Representation of sensitivity is giving in the form of graphs, charts or surfaces. Generally used to give a visual indication of how an output is affected by variation in inputs.

Often the statistical is combined with the graphical in order to communicate the results from e.g. simulations over input parameters. In the following an overview of different measures for robustness is provided.

7.2 Overview over different measures for robustness

This section describes a range of different sensitivity analysis methods. The sensitivity analyses are divided into two main groups of methods deterministic and stochastic

methods. While the deterministic methods are traditionally used, the stochastic sensitivity analyses have more recently been introduced[17].

Deterministic sensitivity analysis methods

- Mareschal (1988) proposes an SA method where stability intervals for the weights of different criteria are defined. The intervals consists of values that the weight of one criterion can take without altering the results given by the initial set of criteria weights (CWs) while all other CWs are kept constant. In order to keep the modified set of CWs normalised, all of the CWs need to be adjusted to ensure that only the importance of the CW being assessed relative to the other criteria is modified. Three types of stability are assessed: 1) Full stability if defined as the absence of any modification in the whole structure. 2) Partial stability concerns the stability of only a part of the ranking alternatives. 3) Subset stability – if the DM(s) only wants to eliminate the worst alternatives and to obtain a subset of good alternatives, with no additional information about these alternatives, the subset stability is defined as the stability of this set.
- Triantaphyllou and Sanchez (1997) present a methodology for performing an SA on the weights of the decision criteria and the performance values of the alternatives expressed in terms of the decision criteria. The method seeks to identify the critical criterion, defined as the criterion which CW has to be altered with the smallest value to reverse the initial ranking of the alternatives. Likewise analysis is performed for the performance values. SAs are proposed for three MCDA methods, weighted sum model, weighted product model, and analytic hierarchy model.

The majority of the deterministic SA methods (also others than those described here) seek to determine what the smallest change in the input parameters are for an alternative to outrank the initial highest ranked alternative (Hyde, 2006).

Stochastic sensitivity analysis methods

- Qureshi et al. (1999) proposes an SA where both the uncertainty of the effects and the criteria weights are examined by Monte Carlo Simulations (MCS) of the initial values assessed by different stakeholder groups by varying them with -+ 30%. Furthermore they examine the final ranking with three different evaluation models.

- Gervásio and Silva (2012) develop an approach based on PROMETHEE and AHP to address conflicting perspectives and uncertainties. Criteria are assessed using a probabilistic approach by MCS and probabilistic distributions fitted to

[17] The described methods here are not an exhaustive list of existing sensitivity analyses, other reviews of methods can be found in Hyde (2006) and Delgado and Sendra (2004).

the results obtained for each criterion. Likewise a probabilistic distribution is fitted to the assessed sets of criteria weights.

- Butler et al. (1997) suggest an approach where MCS is used to vary all of the CWs of an MCDA model simultaneously. The approach is applicable to any MCDA method that utilises CWs in an aggregative scheme. They presents three classes of simulation models 1) Random CWs which requires no weight assessments to explore the entire domain of possible weight combinations. 2) Random rank order CWs where the rank order weights on the measures is maintained, but the weights are otherwise generated at random. 3) Response distribution CWs where the assessed CWs are treated as means of probability distributions of responses and CWs are generated from these distributions. Output statistics of the ranking results are mode, minimum, 25th percentile, 50th percentile, 75th percentile, maximum, mean, and standard deviation. Cumulative ranking distribution figures are also produced.

- Jessop (2002) proposes a methodology which models the effect of uncertainty in the CWs by the use of MCS. Starting with the idea that all criteria are equally important, it is suggested that the initial uncertainty may be modelled with a uniform distribution, with limits ranging from about zero to approximately twice the initial value. The simulated ranks of the alternatives are ordered according to the mean simulated rank.

- Mateos et al. (2006) propose a methodology for aggregating DM preferences when multiple DMs are present. For each DM a class of utility functions and CW intervals are assessed. By the use of MCS an aggregation of the DMs preferences is produced. The resulting ranking of alternatives is presented as boxplots to the DMs and they form the basis for a negotiation process where individual component utilities and weights are tightened. Mateos et al. state that the negotiation process outputs more meaningful information and leads to a consensus strategy. The methodology is not presented as a sensitivity method, but it could be seen as taking the aspect of robustness actively in the decision analysis process and therefore the methodology is listed here as a measure for robustness.

The stochastic SA through simulation is an alternative approach to the deterministic. All of the five stochastic methods described here utilise MCS to simultaneously assess the impact of the uncertainty of the CWs on the decision problem. Qureshi et al. (1999), Gervásio and Silva (2012) and Mateos et al. (2006) also include the uncertainty of the PVs in the sensitivity analysis.

7.3 Formulation of robustness measures

The choice of MCDA model may affect the way that specific robustness measures can be formulated. This relates to the determination of weights and the MCDA method used. This thesis sets focus upon examining MCDA methodologies based on pair-wise comparisons such as AHP and REMBRANDT methods. Based on Barfod (2012) the main

methodology is REMBRANDT. If people are unsure of the weighting method/principles a robustness measure of changes in the weights is more relevant – with regard to the uncertainty in the weighting. But even if people are sure of the method, robustness analysis could be useful to deal with changes in different preferences. Elicitation of weights is usually a time-consuming process and is often controversial, and it is also difficult to arrive at exact weights and to determine consistent intervals between the weights. In such circumstances, ordinal ranking might be a reasonable compromise that makes use of consistent information and often provides output rank order for alternatives similar to the rank order based upon cardinal information (Shepetukha and Olson, 2001). This thesis uses the ordinal ranking method Rank Order Distribution (ROD) for assessing CWs (Roberts and Goodwin, 2002). Focus is on the CWs because of the fact that the impact measurements are usually more objective, or more dependent on expert judgement (Sijtsma, 2006).

Like Gervásio and Silva (2012) represented the results of MCDA by probabilistic distribution functions (PDFs) instead of single values, the results from CBA can also be presented as PDFs (Salling and Leleur, 2011a). The PDFs can be determined by the use of reference class forecasting (Salling and Leleur, 2011b) and thereby a robustness measure of the CBA can be obtained. These PDFs form the input to the MCDA from the CBA part of the decision problem.

Global SA allows all the input factors to vary over their range of uncertainty. The global SA seeks to take a broader range of preferences into account, not only those revealed by a DC. Local methods restrict the analysis to the neighbourhood of a working point where the factors are fixed at their nominal values. The working points are those assessed in the DC(s) conducted to analyse the decision problem. The ranges of variation of input factors are determined by the different inputs assessed by the participants of the DC(s). If no variance or very small variance is assessed, the input range can be extended in order to get a reasonable variance of the input data.

Combination of deterministic and stochastic sensitivity analysis:

Deterministic:

- Weight stability interval (WSI): Determination of critical criteria by examining when a change in the top ranking alternatives occur by changing one CW while all other weights being kept constant.

Stochastic:

- Random ROD CWs: Simulation assigning a random ROD weight to each criterion. Requires no weight assessments.
- Random CWs preserving rank order: Simulation with random generated CW, preserving the rank order of criteria determined by the stakeholders or DMs. Requires an importance ranking of criteria.

The three SAs can be interpreted as different levels of knowledge. With random weights no knowledge exists about the relative importance of the weights and is an extreme case of SA. For Random order weights the rank order of the criteria is maintained, but the weights are otherwise generated at random. Lastly, for weight interval it is assumed that the weights assessed by the stakeholders or DMs are valid, but with some uncertainty, and the simulation can identify the critical criteria where a change of weight can result in a different ranking of the alternatives. This information can be useful, especially within the public decision domain where there is a need to justify the decision to the public. An indication of how critical the criteria are if the weights are changed can defuse a potentially lengthy debate about the setting of the weights (Butler et al., 1997).

As Frey and Patil (2002) divided SA in three different classifications, the proposed approach here involves all three of those; the mathematical, where the WSIs are determined, the statistical, where the stochastic SA examine the change in the resulting ranking of alternatives with different CWs and the graphical where the results of the above SA can be examined and further studied.

The proposed approach is designed to be used actively during a decision conference and thereby as a model help to achieve a decision. As Phillips and Phillips (1993, p. 548) describe it: "Computer models help to take the heat out of disagreements. The model allows participants to try different judgements without commitment, to see the results, and then to change their views" and that "computer-based tools, which are external to the group and not part of it, can provide the facilitator with a powerful means". The output of the probabilistic analysis, by running a MCS, enables the DM to assess the probability that an alternative achieves rank *n* based on all probable criteria input parameters (Gervásio and Simões, 2012).

Robustness analysis can be used as a conflict resolving tool. By highlighting if any changes in preference towards the criteria set will affect the overall resulting ranking of the alternatives. Mapping these potential conflicts of interest or preference among the stakeholders can help the decision process at three different levels:

- *Inform about a conflict*: By conducting a robustness analysis it will be clear if there are any conflict among the stakeholders or DMs. This clarification of conflict can help the DMs to be able to provide a solid argument for their decision to the broad public. Also, and maybe more importantly, the analysis can create a higher degree of consensus among the stakeholders if the potential conflicts have been put forward in the process. Potentially there are no conflict and time and efforts can be used more efficiency to discuss other aspects of the decision problem.

- *Dissolve a conflict*: A robustness analysis can reveal which parameters can potentially change the resulting ranking of alternatives if a reasonable change is made to these parameters. By pinpointing the critical parameters, a more

detailed examination and discussion of these parameters might resolve the conflict.

- *Explicate the conflict*: Even if the conflict cannot be solved, an explication of the critical parameters can provide important information to the DMs. The knowledge of which parameters are critical for reasonable changes, provide the DMs with useful information and maybe request a further analysis of these parameters.

The proposed approach to RA is presented in paper 4, where the approach is applied to two case studies.

The case studies presented later in this thesis are all decision problems where the different stakeholders had agreed upon which criteria to include in the MCDA, or in other words they had all agreed upon the structure of the decision problem. Thereby the difference among the stakeholders can be described as differences in assessment of the different alternatives with respect to each criterion and differences in weights. The latter is described in the classical work of Edwards and Newman (1982), where they have found that resolving conflicts about weights are easier to interpret and discuss than conflicts about the structure of the decision problem, e.g. which criterion to include or not.

In order to use SA actively and directly in the decision process with the stakeholders there is a need for specific properties of the SA. First the SA needs to be a transparent methodology in order to get acceptance and thereby the necessary attention from the stakeholders. Second, the SA has to be able to produce measures for robustness fairly fast, both with respect to the set-up and computation time.

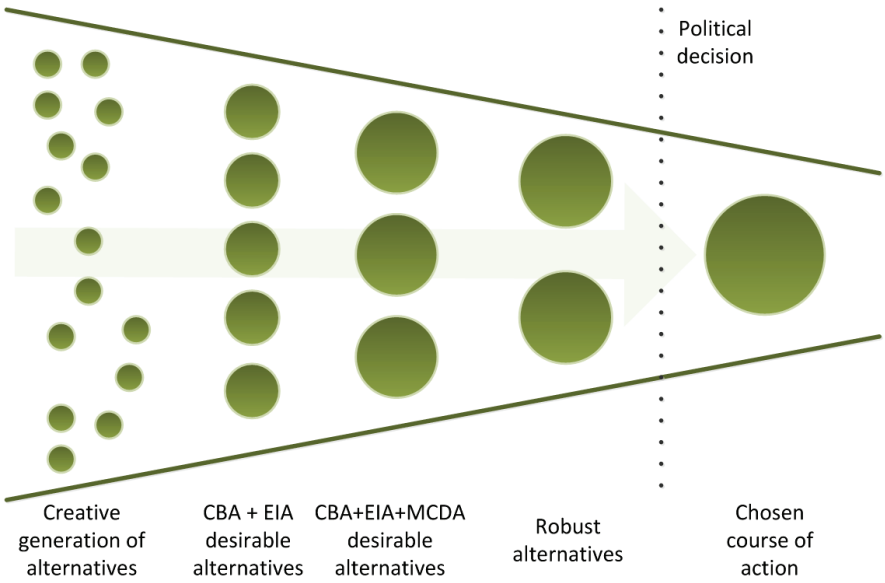


Figure 7-1 Overview of the proposed decision making process with comprehensive appraisal, making use of CBA, EIA and MCDA in combination with robustness analysis

Figure 7-1 presents an overview of the discussed approach for comprehensive appraisal by the use of CBA, EIA, MCDA and robustness measures. The process starts with a creative generation of alternatives to solve the problem at hand. CBA and EIA assessments thereafter reduce the pool of desirable alternatives[18]. Introducing MCDA to apply a comprehensive appraisal may reduce the pool additionally. Finally SA is introduced to examine how robust the different alternatives are for variability in the preference among the stakeholders. Very sensitive alternatives can then be identified and be omitted from the pool of desirable alternatives. However, in the end the final decision is still in the hands of policy-makers and may be based on many other factors.

The papers developed during the PhD study will in the following chapter be presented. They all exemplify different aspects of how to use robustness measures in the decision making process.

[18] It should be noted here that other screening tools can be applied before the CBA and EIA assessments are conducted. Such tool could be a requisite model applied by the use of MCDA. However, for sake of simplicity this step has been omitted from the figure.

8 CASE AND PAPER PRESENTATION

This chapter presents the four papers, which have been written during this PhD study. The papers describe the methodology for assessing the robustness of the project appraisal introduced in Chapter 7. The papers address various topics within the area of uncertainty, appraisal of transport projects, and decision support (see Table 8-2). The papers relate to the theory and issues presented in Chapter 1 to 6. However, first the case studies which have been utilised in the papers are presented in the following.

8.1 DESCRIPTION OF APPLIED CASE STUDIES

The following three sections describe the three case studies that throughout the PhD study have been used to test the research findings: Ostlänken, railway link in Sweden; Malmö, new light rail; HH, new fixed link between the two cities of Helsingør and Helsingborg. The three case studies have been elaborated in connection with two research projects concerning transport evaluation methodologies. As the papers only briefly introduce the case studies, because of limited available space in the papers, they will in the following be described in detail. All of the case studies are comparable as they all anticipate imposing strategic impacts on the surrounding society, however, in different scale. They share another common characteristic by being a larger public transport project with a large public attention in the planning process. Common for all of the case studies is the use of DC in order to provide input to the DMs concerning the various preferences towards the importance of the included criteria. The first two case studies, Ostlänken and Malmö light rail, make use of the composite appraisal methodology COSIMA (COmpoSIte Model for Assessment) in the appraisal, while the third case study, HH connection, makes use of the SIMDEC approach. Table 8-1 provides an overview of the applied cases.

Table 8-1 Overview of case studies

	Ostlänken	*Malmö*	*HH*
No. of alternatives	4	3	3
No. of criteria	9	11	6
Appraisal approach	COSIMA	COSIMA	COSIMA/SIMDEC
Participants at DC	3	10	4
Level of decision power of participants	High	Low/Medium	Limited
Available supplement to CBA	EIA	Appraisal of impacts by municipality	None

8.1.1 Ostlänken

The case study concerning the new railway link Ostlänken examine a smaller section of the project between Norrköping and Linköping. Three corridors for alignment have been identified for the Ostlänken project, red, green and blue corridor (Figure 8-1). For the blue corridor two alternatives exist with either a short or long tunnel going into Norrköping, leaving the total no. of alternatives for the alignment to be four.

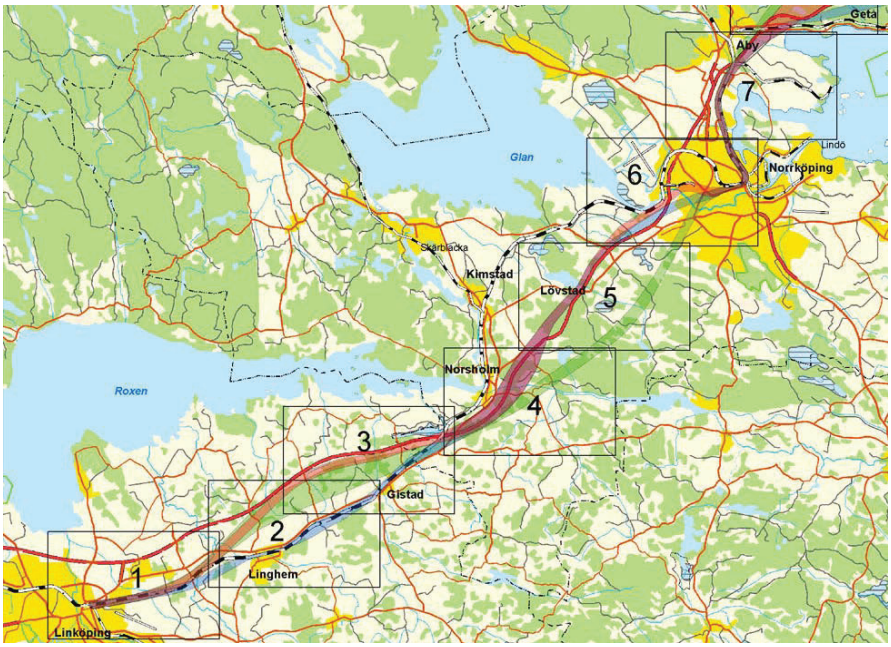


Figure 8-1 Section of the proposed railway used in the case study

The examined section of the project can be divided into two sections, where the three corridors can freely be combined at the northern section with the corridors on the southern section via the passing point at Bäckeby. The northern section is the subject for study in this paper. Red corridor follows the E4 highway. Blue corridor to some extent follows the existing railway line. Green corridor runs mainly in previously untouched areas.

A decision conference with three participants was conducted in order to get input to the decision analysis. One participant represented the National Railway Agency and the other two the municipality of Norrköping. The participants in the DC were directly involved in the case and could be described as stakeholders with a high level of power in the actual decision making.

8.1.2 Malmö light rail

In the city of Malmö, Sweden, it is planned to introduce (actually to re-introduce as there has previously been such a transport system in the city) a new public transport

system in the shape of light rail. The new public transport system is to be planned to support the already existing transport system in the city. The light rail system is to be introduced in stages and in the first stage three different alignments have been identified to be examined further (alignment A, B and C). The three alignments run through parts of the city which are different in respect to the benefits to be obtained. All of the three alternatives start at the centre of Malmö city (Figure 8-2).

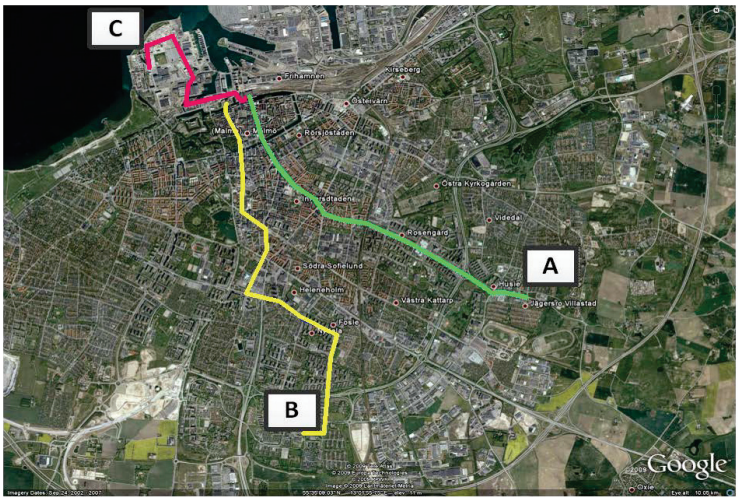


Figure 8-2 The three alternatives for alignment

A priori comprehensive background material has been worked out for the proposed transport system. This material includes CBA and identification of several non-commensurate criteria to be included in the decision analysis. The background material has been worked out by a steering committee in the municipality of Malmö. This committee consists of 10 different stakeholders e.g. local authorities and public transport operators. The committee participated in a DC to analyse the decision problem and the output from this DC forms the data for the SA. The participants could be described as stakeholders that are not necessarily directly involved in the case, but consists of a broad range of stakeholders with a low level of power in the actual decision making.

8.1.3 HH Connection

The proposal of a fixed link between Helsingør (Elsinore) in Denmark and Helsingborg in Sweden – referred to as the HH-connection – has been considered since the 1980s; however, the opening of the Øresund fixed link at Copenhagen in 2000 postponed the planning and implementation. The case is now re-actualised in order to cope with the increasing traffic across the Øresund and the planned Fehmarn fixed link, connecting Denmark and Germany. A new northern fixed link would reduce the travel time between Zealand and Scandinavia, relieving the Øresund fixed link for some of the car

and railway traffic. In the autumn of 2011 three tunnel alternatives were considered as main alternatives for the HH-connection.

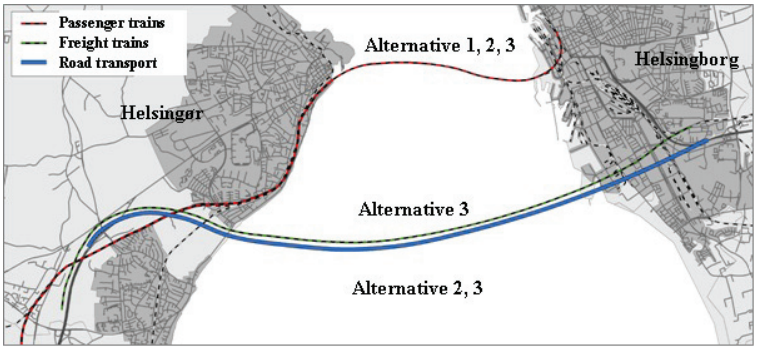


Figure 8-3 The alignment of the three alternatives for a new fixed link

A decision conference was conducted concerning the HH-connection decision problem, where the purpose was to locate the most attractive alternative amongst the three presented in Figure 8-3. Four experts participated in order to conduct the decision conference and ultimately create a decision support base with regard to the prioritisation of the alternatives. The conference was guided by a facilitator supported by a model analyst using an MCDA model to perform the modelling of the information obtained from the participants. The participants could be described as stakeholders that are not necessarily directly involved in the case, but could be of use to the DMs as expert opinion on the performance of the three alternatives. The participants are assumed not to have any power in the actual decision making.

8.2 Paper presentation

In the following the four papers, which have been developed during the study, will be presented. First an introduction to the paper is given, next an overview of the main findings of the paper is presented and finally the paper as published or presented.

In order to provide an overview of the papers and their subject see Table 8-2.

Table 8-2 Overview of the papers written during the PhD study

	Papers developed during the PhD study
1. Paper	**Estimating the Robustness of Composite CBA and MCA Assessments by Variation of Criteria importance Order** The first paper introduces a method of using rank variation concerning the stakeholders' perception of criteria importance for exploring the sensitivity of criteria weights in MCA. The method preserves the CBA information and applies ROD weights based on the stakeholders' ranking the criteria after importance. The proposed framework is applied to the case of choosing a light railway line to be implemented in the Swedish town Malmö. The method has shown that it is possible to conduct a sensitivity analysis by using rank variation and the analysis gives useful information to the decision makers.
2. Paper	**Importance of Criteria Weights in Composite appraisal of Transport Projects – A Robustness Analysis** The second paper presents a methodology for testing the robustness of the assigned weights to the analysis and thereby gives information about how different preferences about the importance of the criteria may result in a different outcomes of the appraisal. The work on the two case studies presented in this paper illustrates that the methodology is particularly useful where groups make decisions, when it is often hard to get agreement about the setting of the weights. The analysis can defuse a potentially lengthy debate about the weights and thereby leave more time for discussing other parts of the analysis among the decision makers.
3. Paper	**Robustness Assessment of Strategic Transport Projects** The third paper examines the criteria weights for estimating the robustness of the different alternatives proposed for a HH connection. The framework is based on a participatory MCDA approach by using decision conferences for identifying the preferences of stakeholders. The focus of this paper concerns a framework which provides a transparent analysis that can be communicated to DMs
4. Paper	**Comparisons of Robustness Measures – as Communicative Means for Involvement of Decision-Makers** The fourth paper addresses how to measure the robustness of decision-making relating to transport projects involving multiple objectives as regards involving views of multiple and diversified stakeholders within the MCDA. It is examined how the choice of MCDA methodology can affect the robustness by examining two different pair-wise comparison methodes: The AHP and the REMBRANDT.

8.4 Overview of paper 1

Robustness model for composite appraisal

Title: Estimating the Robustness of Composite CBA and MCA Assessments by Variation of Criteria Importance Order

Authors: Anders Vestergaard Jensen, Michael Bruhn Barfod and Steen Leleur

Presented at: The 20th International Conference on Multiple Criteria Decision Making, Chengdu, China, 2009.

Published in: Y. Shi et al. (eds.), (2011) *New State of MCDM in the 21st Century – Selected Papers of the 20th International Conference on Multiple Criteria Making 2009*, Lecture Notes in Economics and Mathematical Systems, Vol. 648, pp. 59-67, Springer.

9.4.1 Introduction to paper

The aim of paper 1 is to present a model for assessing the robustness of the COSIMA composite assessment methodology. The COSIMA model is characterized by the principal that the total result of the approach consists of two parts, a CBA and an MCDA[19]. The total result of the COSIMA can be expressed by the Total Rate of Return (TTR). By using this approach the MCDA part has to be additive to the CBA part. The additive MCDA part is expressed by an MCDA percentage. This percentage expresses the MCDA part of the total analysis in relation to the CBA part. The robustness model presented considers changes in the MCDA percentage and the CWs. The CWs are assessed by the use of Rank Order distribution (ROD) technique. The sensitivity of the CWs are examined by using all possible ranking of criteria after importance along with a more restricted approach where for example that criterion XX is the most important criterion, or criterion XX always has to be ranked higher than criterion YY. This latter restricted approach would mean that the outcome of the method is a subset of the total solution space.

The model for assessing the robustness is illustrated by a case study regarding implementation of light rail lines in the city of Malmö, Sweden. A decision conference is used for assessing the performance scores for all of the alternatives, sets of criteria weights, and the MCDA percentage.

[19] It should be noted here that the paper makes use of the notation multi-criteria analysis (MCA) and not MCDA. In this context the meaning of the two terms are identical. However, during the PhD study I made a choice that MCDA better describes the decision process using MCDA and therefor ended up using MCDA instead of MCA.

Furthermore, the paper presents a method for visualising the results from the sensitivity analysis, showing how robust each alternative is to changes in the MCDA percentage, and applying different rank order of the criteria. The results can be illustrated in one single chart for each alternative and are thereby easily communicated to the DMs. The limitation of the model presented is that it is not suited to be used actively during the DC because of calculation time of all the possible ranking of criteria. With 8 criteria there are 40,340 different combinations of ranking and for each of them the COSIMA calculation has to be executed.

9.4.2 Main findings

The paper has presented an approach for illustrating the simulation results containing a huge amount of data. Figure 8-4 illustrates the ranking of alternative B as a function of the MCDA percentage (how the criteria in the MCDA part of the appraisal are weighted compared to the CBA part) and all possible preference order of the criteria. From this figure it is possible to conclude that if the MCDA percentage is below 25% it does not matter which preference orders the criterion is assessed, alternative B will always be the most desirable. For the MCDA percentage set to 60% (vertical dashed line in the figure) alternative B will be ranked as the most desirable alternative in most of the preference orderings of the criteria, however, in some cases also ranked as no. 2 and in some special cases also as no. 3. With a specific preference order of the criteria (horizontal dashed line) alternative B is ranked as no. 1 with an MCDA percentage in the range of 0 to approximately 76%.

How to utilise the information gained at a DC in a constructive manner to inform the DM about the consequences of the decision problem's uncertainty has been treated in this paper. This is done by setting constraints to the simulation of the setting of the criteria weights. These constraints stem from the preferences revealed during the DC by the stakeholders participating in the DC. Then the simulation is conducted with respect to these constraints and the results will reveal how the constraints are affecting the robustness of the decision analysis.

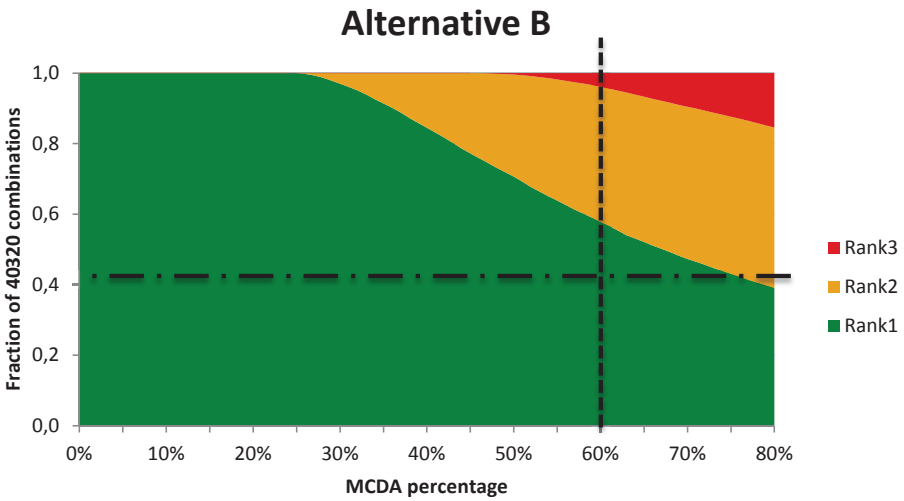


Figure 8-4 Simulation results for alternative B. The vertical dashed line represents one specific MCDA percentage and the horizontal dashed line represents one specific ranking of the criteria.

The paper has presented the COSIMA methodology for assessing a comprehensive appraisal consisting of both monetary impacts (CBA) and impacts which cannot be monetised (MCDA). This approach has formalised the process of assessing these non-monetised impacts, which the municipality of Malmö had already begun. Thereby the paper contributes to answer research question A concerning how to obtain decisions which are both rational and legitimate. Rational, because the methodology is a transparent method when each impact is assessed by its PV and CW. Legitimate, because the method has taken various stakeholders' preferences into account and by robustness analysis has examined the influences of these preferences on the overall ranking of the alternatives. The latter also contributes to answering research question B concerning how to involve stakeholders in the appraisal. The DC has been presented as a mean with respect to this research question, where the combination of technical appraisal methodology, that captures the different perspectives, and a social process, that engage the stakeholders, was a successful process.

8.5 Paper 1 Robustness model for composite appraisal

Estimating the Robustness of Composite CBA and MCA Assessments by Variation of Criteria Importance Order

Anders Vestergaard Jensen, Michael Bruhn Barfod and Steen Leleur

Abstract

This paper discusses the concept of using rank variation concerning the stakeholder prioritising of importance criteria for exploring the sensitivity of criteria weights in multi-criteria analysis (MCA). Thereby the robustness of the MCA-based decision support can be tested. The analysis described is based on the fact that when using MCA as a decision-support tool, questions often arise about the weighting (or prioritising) of the included criteria. This part of the MCA is seen as the most subjective part and could give reasons for discussion among the decision makers or stakeholders. Furthermore, the relative weights can make a large difference in the resulting assessment of alternatives (Hobbs and Meier 2000). Therefore it is highly relevant to introduce a procedure for estimating the importance of criteria weights. This paper proposes a methodology for estimating the robustness of weights used in additive utility models.

Keywords: Multi-criteria analysis; Rank order distribution; Decision-support tool; Stakeholder preference; COSIMA.

1 Introduction

When assessing larger transport infrastructure projects often several non-monetised impacts could be relevant to include in the appraisal (UNIVERSITAET STUTTGART 2009). For many decision makers and stakeholders the task of setting the criteria weights for several criteria can be very difficult (Alfares and Duffuaa 2008). To overcome this, the proposed method uses importance weights based on rankings of the criteria, by the use of Rank Order Distribution (ROD) weights (Roberts and Goodwin 2002). ROD weights are one of many techniques for utilizing the ranking order to establish criteria weights (Alfares and Duffuaa 2008). This reduces the problem to assigning a rank order value for each criterion. A method for combining the MCA with the cost-benefit analysis (CBA) is applied as described by Salling et al. (2007). This methodology, COSIMA, uses an indicator which expresses the trade-off between the CBA and MCA part resulting in a total rate expressing the attractiveness of each alternative. However, it should be mentioned that the proposed procedure for estimating the importance of criteria weights is not limited to the ROD and COSIMA methods described above.

The proposed framework is applied to the case of choosing a light railway line to be implemented in the Swedish town Malmö. The decision concerns which of the 3

identified lines should be built and put into operation first. Preliminary studies have found 8 non-monetised criteria, which are not all included in the conducted CBA. The alternatives are compared to each other with respect to each criterion by using the REMBRANDT methodology. With 8 criteria there are 40320 (8!) possible combinations of ranking the criteria. The proposed method calculates all combinations and produces a set of rank variation graphs for each alternative and for different values of the trade-off indicator. This information is relatively easy to grasp for the decision-makers. The result is compared with the results from a conducted decision conference about the decision problem. During the decision conference the different stakeholder preferences were unveiled by the participants who had to assign weights.

The proposed method also introduces a more constrained approach. In this approach the stakeholders/decision makers have the possibility to set up some constraint to the decision problem. This could for example be that criterion XX cannot assume a rank lower than 3 or criterion XX always have to be ranked higher than criterion YY. This would mean that the outcome of the method is a subset of the total solution space.

The paper finishes up with a discussion and considerations about how to present the results. The question whether to present a single decision criterion, such as the benefit-cost rate or the net present value, or instead to present graphs showing the robustness of the decision analysis is discussed. Furthermore a perspective, for estimating the robustness of weights using other MCA methodologies (and weighting methods) than the proposed framework, is discussed.

2 Case study: Public transport system

For testing the proposed methodology, a case study has been examined. It concerns the introduction of a new public transport system in the Swedish town Malmö, located in the southern part of Sweden. The case study was conducted under the project "An overall evaluation of public transport investments" (Hiselius et al. 2009). The new transport system consists of several light railway lines, which are planned to be introduced in different stages and in the first stage 3 different lines have been chosen to be further examined. The 3 lines run in different parts of the city and therefore differ in which impacts they result in. As a preliminary analysis the municipality of Malmö had a cost benefit analysis (CBA) conducted (Hiselius et al. 2009). The CBA was carried out following the national Swedish standard for transport project appraisal.

Table 1 Results from the cost benefit analysis (Costs and benefits in M SEK)

Line	Costs	Benefits	BCR
A	652	1475	2.26
B	603	1978	3.28
C	296	579	1.96

The result from the CBA indicated that all 3 lines are economically feasible with line B as the most feasible with at benefit-cost-rate (BCR) on 3.28. In addition the municipality also conducted a wider analysis of the 3 lines covering impacts which not are included in the CBA, all these impacts could not be assessed in a monetary way. This wider analysis of the 3 alternatives was performed by a working group formed by the municipality and consisted of 10 different stakeholders. However, the municipality had difficulties in making an overall assessment of the 3 lines taking all the impacts into consideration. Consequently, there was a need for a more comprehensive assessment. This assessment was done using a decision support system called COSIMA, which is a composite model for assessment of transport projects. Using decision models which make use of both CBA and MCA in combinations have been recommended to give complementary insights (Macharis et al. 2008).

2.1 Composite assessment

In order to conduct a composite assessment of the 3 alternatives the COSIMA (COmpoSIte Model for Assessment) framework was used (Salling et al. 2007). As with other composite models COSIMA is an extension of a traditional CBA to a more comprehensive analysis, which seeks to meet the demands from the decision makers. The COSIMA analysis is characterized by the way it includes, in addition to the CBA, missing impacts which have relevance for the examined project. These impacts are often not possible to express by monetary values and thereby they are excluded from the CBA, but by including them in the assessment a better foundation for a decision can be achieved.

As mentioned the COSIMA consists of two parts, a CBA and a MCA and the total result of the COSIMA can be expressed by the Total Rate of Return (TTR). By using this approach the MCA part has to be additive to the CBA part. The additive MCA part is expressed by a MCA percentage. This percentage expresses the MCA part of the total analysis in relation to the CBA part. In a situation where the MCA part is regarded as just as valuable as the CBA part, the MCA percentage would be 50 % (see Fig. 1). The decision makers have to make this value judgment in order to calculate the TRR.

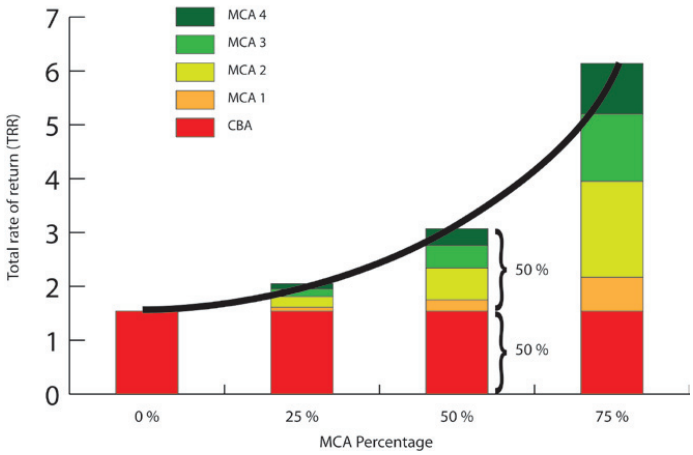


Figure 1 The total rate of return (TRR) as a function of the MCA percentage.

The COSIMA framework uses pair wise comparisons for rating of the alternatives, either the Analytic Hierarchy Process (AHP) or the REMBRANDT technique is used. For setting the criteria weights Rank Order Distribution technique is used (Roberts and Goodwin 2002). The ROD technique provides importance weights based on the rankings of the criteria, thereby the weight approximation problem is reduced to rank the criteria after importance. The ranking of the criteria reveals the decision maker's perception of how important the criterion is relative to other criteria under consideration (Monat 2009).

2.2 Decision conference

In order to capture the stakeholders' preferences a decision conference was conducted with the purpose of finding the most attractive alternative seen from the involved stakeholders' point of view (Phillips 2006). The decision conference was tailored in order to capture all the different perspectives of the decision problem and to give input to the COSIMA model. One of the advantages with involving the stakeholders (and the decision makers) are that many stakeholders tend to reject decision criteria that appear to provide the "right" answer through a single number coming out an economic mode. They want the flexibility of implementing their own value judgments in the decision (DeCorla-Souza et al. 1997). Furthermore, a decision made during a decision conference has a higher probability of acceptance than a decision made based on complex decision process (Goodwin and Wright 2004).

The decision conference was conducted with 10 participants, which all were stakeholders of the project. All the stakeholders were also part of the working group formed by the municipality in order to find the most appropriate light railway line to build first. Prior to the decision conference the CBA was conducted and a simplified multi-criteria analysis was also carried out by the working group. This simplified MCA

was used as foundation for defining the criteria to be implemented in final MCA. However, the criteria set used by the working group had some overlap with the CBA and some of the criteria were in need of a clarification, in order to ensure that all the members had a common perception of each criterion. The set-up of the decision conference was: presentation of the alternatives, presentation of the CBA, defining the criteria, description of the criteria, ranking of the criteria after importance, evaluating the alternatives (pair wise comparisons), setting the MCA percentage and finally a presentation of the results from the decision conference. The group decided on the following criteria to be implemented in the analysis (in preference order): capacity, urban renewal, social integration, compact city, environment, corresponding public transport, impact on car traffic and finally impact potential. All the participants had opportunity to make an individual ranking of the criteria and setting an individual MCA percentage.

After some discussion the group agreed on a ranking of the criteria, the pair wise comparisons and a MCA percentage of 60 %. The resulting total rate of return is illustrated in Fig. 2. The TRR for the 3 alternatives show that alternative B has the highest TRR – based on the preferences the group could agree upon.

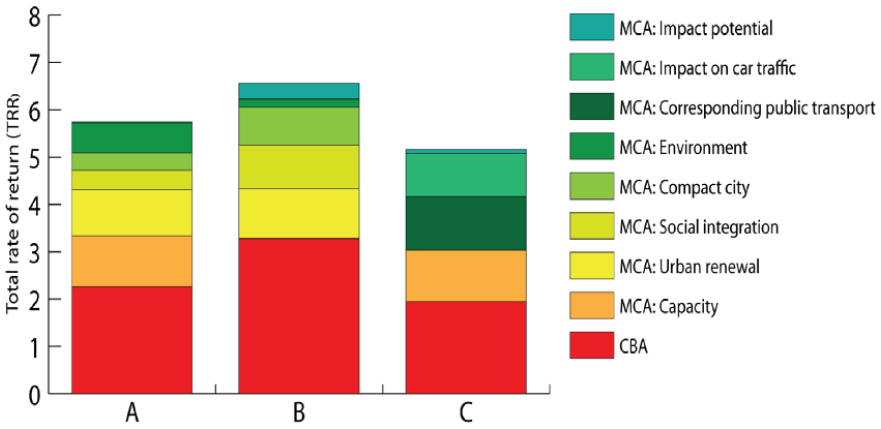


Figure 2 The group's consensus with a MCA percentage on 60 %

To test the result sensitivity for changes in the MCA %, all TRRs was calculated for MCA % in the range from 0 % (only contribution from the CBA) to 80 %. The sensitivity test is illustrated in Fig. 3. The figure shows that with any MCA percentage up to approximately 75 % it makes no difference to the overall result. This result does not differ from the result from the CBA, which also had alternative B as the most attractive alternative. However, this composite analysis gives valuable information to the decision makers anyway. A possible decision about building line B first would be based on a better foundation than just the CBA. The municipality is better prepared for eventual criticism of the decision because they can publish a comprehensive analysis taking a wide range of impacts into account.

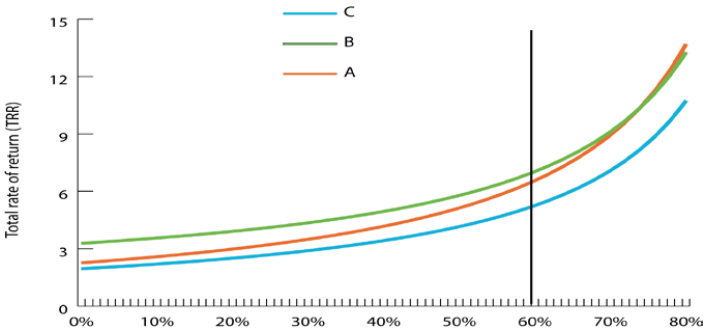


Figure 3 Sensitivity test for changes in the MCA %. The bar indicates the group's consensus on a MCA percentage

The sensitivity analysis showed that many of the disagreements or uncertainties found during the decision conference had no impact on the overall result, line B is the most attractive alternative. However, during the decision conference the participants had difficulties in agreeing on a consensus about the ranking of the criteria after importance. So the individual criteria weights, which also were recorded, were compared to the criteria weights agreed upon in consensus. All the participants had ranked the same criteria as most important and there seemed to be most agreement about the higher ranked criteria and more disagreement with the lower ranked criteria. This disagreement gave reason for also testing the TRRs robustness for variation in the ranking (importance order) of the criteria.

2.4 Robustness of the composite assessment

In order to test the results robustness for different ranking of the criteria all possible combinations of how the 8 criteria can be ranked were calculated. The 8 criteria can be ranked in 40320 different ways and the group consensus on the ranking was only one of all these combinations. The simulation was carried out with the same pair wise comparisons and same CBA information. Furthermore, all combinations were calculated with the MCA % in the range from 0 % to 80 %. A simulation result for alternative B, the alternative with the highest TRR, is illustrated in Fig. 4.

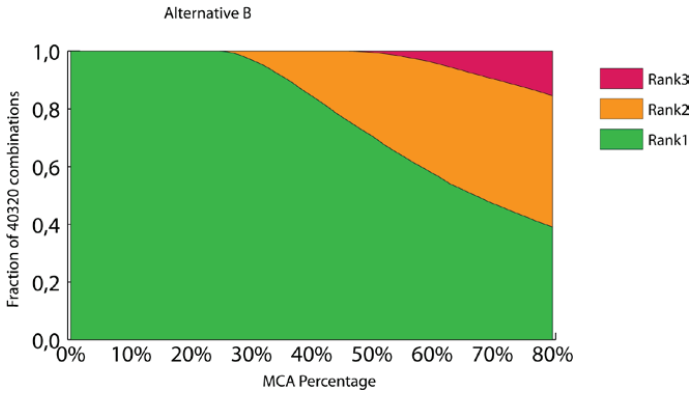


Figure 4 Simulation results for alternative B

The simulation results are illustrated in a manner where the decision makers easily can see how robust the TRR of the alternative is for changes in the ranking, and thereby weighting, of the criteria. For alternative B it can be seen that in the range of a MCA % on 0 % to approximately 25 %, it does not matter how the criteria are ranked after importance – alternative B will be ranked as number 1 among the alternatives in all cases. In the case of a MCA % set to 60, which was chosen by the group, alternative B is ranked as the most attractive alternative in approximately 58 % of all possible rankings. The group did choose a ranking which put alternative B as the highest ranked alternative, but other rankings of the criterion would put either alternative A or C as the most attractive alternative. Similar illustrations could also have been made for the other alternatives.

However, the decision conference revealed a tendency in the ranking of the criteria, and this information could be used in the simulation. All the participants ranked capacity as the most important criterion, so a simulation with this criterion fixed as ranked number 1 could be made. Furthermore, there was also a tendency in which criterion (urban renewal) was ranked as second most important. These tendencies can be simulated and the results can either confirm that the decision makers (in this case the participants of the decision conference) have a robust result that points out a certain alternative as being the most attractive or that the result is sensitive to changes in the ranking of the criteria. Fig. 5 illustrates the results from 3 different simulations at a MCA percentage at 60 %. The first bar shows the simulation results with all combinations, where in 58 % of all the combinations alternative B would be ranked as number 1. The next bar is the result from a simulation with capacity fixed as criterion ranked as number 1. It can be seen that there is a lower fraction of all combinations giving alternative B the highest rank than in the simulation with all combinations. The reason for this result is that alternative B has the lowest score among the alternatives for the capacity criterion and ranking this criterion as number 1 is then a disadvantage for alternative B. The next bar shows the results for the simulation with capacity and urban renewal fixed as ranked as number 1 and 2. This simulation shows the highest

fraction of all possible combinations giving alternative B the highest rank (61 %). A reason for this can again be found in how the alternative B is scoring on the criterion. Alternative B has the highest score on urban renewal, together with alternative A, this gives alternative B an advantage when the criterion is fixed as the next most important criterion.

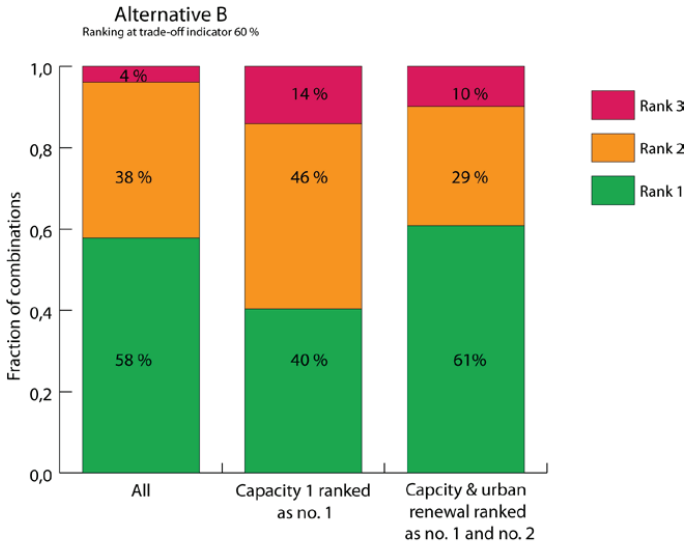


Figure 5 Constrained simulation approach with one or two criterion rankings fixed

This approach provides an opportunity to deal with disagreement among the participant about the ranking of the criteria. If the group have difficulties in ranking one or more criteria, a simulation can be made based on keeping all the agreed upon rankings fixed.

3 Conclusion and perspectives

This paper has introduced a method of using rank variation concerning the stakeholders' perception of criteria importance for exploring the sensitivity of criteria weights in MCA. The method preserves the CBA information and applies ROD weights based on the stakeholders' ranging the criteria after importance. The method has shown that it is possible to conduct a sensitivity analysis by using rank variation and the analysis gives useful information to the decision makers.

The combination of a technical solution (COSIMA model) that captured the different perspectives with a social process (decision conference) that engaged the stakeholders were a successful process where the stakeholders were confident with the final results (see also Phillips and Bana e Costa 2005 for similar conclusions).

A further research task is to consideration if the robustness test should be revealed for the decision makers or not? This might give reason for not trusting the model results among the decision makers. However, if kept away from the decision makers, it will be the analyst's responsibility to ask for more detailed information of for example the pair wise comparisons for getting a more robust result. Furthermore, it can be discussed if the discrete ranking of the alternatives is the most optimal way of presenting the results of the robustness test.

References

Alfares, H. K., Duffua, S. O. (2008) Determining aggregate criteria weights from criteria rankings by a group of decision makers. *International Journal of Information Technology & Decision Making* Vol. 7, No. 4, 769-781

DeCorla-Souza, P., Everett, J., Gardner, B., Culp, M. (1997) Total cost analysis: An alternative to benefit-cost analysis in evaluating transport alternatives. *Transportation* 24, 107-123

Goodwin, P., Wright, G. (2004) *Decision Analysis for Management Judgement*. Wiley

Hiselius, L., Barfod, M.B., Leleur, S., Jeppesen, S.L., Jensen, A.V., Hjalte, K. (2009) *An overall evaluation of public transport investments.* University of Lund – In Swedish.

Hobbs, F.B., Meier, P. (2000) *Energy decisions and the environment – A guide to the use of Multicriteria Methodes*. Kluwer, Boston

Macharis, C., De Witte, A., Ampe, J. (2008) The Multi-Actor, Multi-Criteria Analysis Methodology (MAMCA) for the Evalutaion of Transport Projects: Theory and Practice. *Journal of Advanced Transportation*. Vol. 43. No. 2. 183-202

Monat, J. P. (2009) The Benefits of global scaling in multi-criteria decision analysis. *Judgement and Decision Making* Vol. 4, No. 6. 492-508

Phillips, L.D., Bana e Costa, C. A. (2005) Transparent prioritization, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing. Working paper LSEOR 05.75

Phillips, L. D. (2006) Decision Conferencing. A working paper LSEOR 06.85

Roberts, R., Goodwin, P. (2002) Weight approximations in multi-attribute decision models. J. Multi-Crit. *Decis. Anal*. 11: 291–303

Salling, K. B., Leleur, S., Jensen, A.V. (2007) Modelling decision support and uncertainty for large transport infrastructure projects: The CLG-DSS model of the Øresund Fixed Link. *Decision Support Systems* 43 1539–1547

UNIVERSITAET STUTTGART (2006) HEATCO - Developing Harmonised European Approaches for Transport Costing and Project Assessment. EU Sixth framework Programme 2002-2006. http://heatco.ier.uni-stuttgart.de/. Accessed 24 February 2009

8.6 Overview of paper 2

Involvement of stakeholders

Title: Importance of criteria weights in composite appraisal of transport projects – a robustness analysis

Authors:	Anders Vestergaard Jensen
Presented at:	Uncertainty and Robustness in Decision Making – 25th Mini EURO conference, Coimbra, Portugal, 2010.
Submitted:	For publication in Journal of Science & Technology, May 2012, the paper is currently under review.

9.5.1 Introduction to paper

The aim of paper 2 is to extend the findings of paper 1 and to illustrate sensitivity analysis performed at specific settings of the MCDA percentages. Some of the findings of paper 1 are included in paper 2 in order to put a broader perspective on robustness analysis. Furthermore, this paper presents an additional case study using the decision conference (DC) approach. An environmental impact assessment (EIA) was available prior to the conducted DC, for the added case study, and thereby this paper seeks to use EIA, CBA and MCDA methodologies for appraising the different alternatives.

9.5.2 Main findings

A framework appraisal methodology is presented in paper 2, where the use of robustness analysis is used actively in the decision process. By implementing the robustness analysis in an early stage in the appraisal process, the stakeholders identify and quantify the sources of uncertainty and thereby get a deeper understanding of the problems with the composite assessment. This adds value to the process. This paper also extends the sensitivity analysis presented in paper 1 to examine the sensitivity of a decision problem with a specific setting of the MCDA percentage.

The CWs of individuals should be assessed prior to group CW to obtain the variance in preferences. The high correlation between the group weights and individual CWs shows that the process of discussion the CWs has influenced the setting of preferences of the various participants at the DC, which again illustrates, that it is possible to reach some degree of consensus at a DC with 10 participants.

The paper has also illustrated that SA can be used as a screening tool before a more thoroughly RA might take place. By ruling out alternatives earlier in the decision process, more resources can be concentrated on examining the remaining alternatives. This finding contributes to answering research question C concerning applying robustness to improve the appraisal process.

The paper has demonstrated that EIA is useful as basis for further discussion of other impacts than those implemented in CBA. Although, one interesting finding regarding the use of EIA was that the participants did not fully agree with the assessments of the alternatives' PVs. So the EIA provided the criteria set and a starting point for discussion among the participants on how the alternatives perform on the different criteria. This indicates that sensitivity analysis of the PVs can be useful; nevertheless this has not been treated in this PhD study.

8.7 Paper 2 Involvement of stakeholders

Importance of criteria weights in composite appraisal of transport projects – a robustness analysis

Anders Vestergaard Jensen

Abstract

This paper presents a framework for estimating the robustness of appraisal of transport projects using a composite appraisal methodology which makes use of both cost-benefit analysis (CBA) and multi-criteria analysis (MCA). The setting of criteria weights in MCA is often seen as the most subjective part of the analysis and is frequently a subject for discussion. The methodology examined here makes use of importance order for setting the criteria weights; this means that all the decision makers or stakeholders involved have to do is rank the criteria by their importance. In some cases, the setting of criteria weights can make a big difference in the resulting assessment of alternatives. This paper presents a methodology for testing the robustness of the assigned weights to the analysis and thereby gives information about how different preferences about the importance of the criteria may result in a different outcome of the appraisal. The work on the two case studies presented in this paper illustrates that the methodology is particularly useful where groups make decisions, when it is often hard to get agreement about the setting of the weights. The analysis can defuse a potentially lengthy debate about the weights and thereby leave more time for discussing other parts of the analysis among the decision makers. The methodology may be useful in a wide variety of contexts outside the field of transport projects.

Keywords: Multi Criteria Decision Analysis (MCDA), Decision support, Robustness, Stakeholder preferences, Decision conference

1 Introduction

When transport projects are assessed, the well-established cost-benefit analysis (CBA) takes care of appraising socioeconomic impacts, such as time savings, pollution, accidents, etc. However, it is well known that the CBA does not cover the 'soft' impacts, to which investments in the transport sector also contribute (Thomopoulos et al., 2009). These impacts are often difficult to identify and even harder to appraise, but are important in the appraisal because they may tip the balance in the choice between various options. Often these impacts are described verbally without any formal inclusion in the assessment. COSIMA (Composite Model for Assessment) is a composite appraisal methodology which enables the analysis to take both monetary and non-quantifiable impacts into account by using multi-criteria analysis (MCA). This paper extends this methodology with a robustness analysis which can be applied to COSIMA.

It should be noted that the principles of COSIMA are illustrated in the paper along with the treatment of the cases. For a formal treatment, the reader is referred to (Salling et al., 2007).

Several different approaches for assessing the robustness of multi-criteria analysis (MCA) have been proposed. (Jones, 2011) proposes a weight sensitivity algorithm, which produces a set of distinct solutions that characterise the portion of weight space searched. In (Jimenez et al., 2004) a decision support system is presented which operates with imprecise assignments for weights and utilities and defines ranges of the weights where the assessment is robust. (Montibeller and Franco, 2011) use scenarios for assessing the robustness of alternatives under different conditions.

The fact that decisions in the transportation field by nature have a high public awareness, and often involve participation from various stakeholders with different preferences, makes the assessment of robustness in the analysis methodology very relevant. One of the critical points of the socioeconomic appraisal is actually that it often suffers from a failure to represent the stakeholders' dynamic preferences (Musso et al., 2007). However, the presence of these various preferences also creates some challenges for the assessment methodology. It is necessary not only to include the various preferences, but also to arrive at some degree of consensus on which alternative is the most desirable, and when there are very divergent agendas and preferences among the stakeholders, this is not an uncomplicated task. We believe that an appropriately constructed robustness analysis can overcome some of these difficulties by showing how the different preferences influence the total ranking of alternatives. This issue has not been extensively addressed before. But the involvement of several different stakeholders also means that a more practical approach to the assessment is necessary. To include the qualitative or non-monetised impacts in the appraisal, it is necessary for the stakeholders involved to discuss and prioritise the various impacts. For many decision makers and stakeholders, the task of giving weights to various criteria can be very difficult (Alfares and Duffuaa, 2008a). To overcome this, the proposed method uses importance weighting, based on ranking of the criteria using Rank Order Distribution (ROD) weights (Roberts and Goodwin, 2002). This reduces the problem to assigning a rank order value for each criterion; the ROD methodology then converts the ordinal ranking into a numerical weight. It should be noted that there are several different criteria rank-weight methods (see (Belton and Stewart, 2001) and (Alfares and Duffuaa, 2008b)). The COSIMA methodology uses an indicator which expresses the trade-off between the CBA and MCA parts, resulting in a total rate expressing the desirability of each alternative.

However, the proposed procedure for estimating the importance of criteria weights is not limited to the ROD and COSIMA methods described above. It can be applied in general to all the different methodologies of the Multi-Attribute-Utility-Theory family.

To capture the stakeholders' preferences, a decision conference can be held with the purpose of finding a ranking of the desirability of each alternative seen from the point of view of the stakeholders involved (Edwards et al., 2007), (Phillips and Bana e Costa,

C. A., 2005). The decision conference should be tailored to capture all the different perspectives on the specific decision problem and give input to the COSIMA model. One of the advantages with involving the stakeholders (and the decision makers) is that many stakeholders tend to reject decision analysis that appears to provide the "right" answer through a single number coming out of an economic model. They want the flexibility of implementing their own value judgments in the decision (DeCorla-Souza et al., 1997). Furthermore, a decision made during a decision conference has a higher probability of acceptance than a decision made based on a complex decision process (Goodwin and Wright, 2004).

After this introduction, a description of the two case studies follows. Each case study is described in combination with the results from the decision conference held. The design of the two decision conferences differs slightly. While the first case study, Ostlänken, had pre-defined criteria, the second case study, Malmø Light Rail, only had vaguely defined criteria. Because of this, in the second case study, more emphasis was put on elaborating the criteria and achieving a clear understanding of each criterion. See (Jeppesen, 2010) for a more detailed description of the two decision conferences. The results from both the individual and group decisions are discussed and the robustness analysis is described in detail. However, it should be noted that the output from the decision conference is a product which is shaped by the participant's knowledge and values (Edwards et al., 2007).

2 Methods

2.1 Problem definition

This paper considers a deterministic MCDA problem with k alternatives and j decision criteria. The criteria weights reflect the relative importance of each criterion. These weights are normalised by making their sum equal to $1 \sum_{j=1}^{n} w(j) = 1$. Given the specific performance value $V_{MCA}(X_{jk})$of each alternative k (k = 1, 2,..., m) in terms of each criterion j (j = 1, 2,...,n), the overall performance of each alternative k can be calculated as follows:

$$P(A_k) = \sum_{j=1}^{n} w(j) \cdot V_{MCA}(X_{jk}) \qquad (1)$$

It is assumed that input can be obtained from several individuals, where each individual i may rank the criteria in order of importance. Thus, a list of n_i prioritised (ranked) criteria is given by each individual i, who gives each criterion j a rank $r_{i,j}$. Furthermore, it is assumed that a group of individuals can (partly) agree on a list of prioritised criteria. The rank is assumed to be inversely related to weight (rank 1 denotes the highest weight). The problem is to determine to what extent the ranking of criteria influences the overall performance of each alternative. The aim is to determine, on the basis of ROD weights, how robust the alternatives are for changes in the ranking of criteria.

2.2 COSIMA

The purpose with a composite appraisal methodology is to expand the conventional CBA to a more comprehensive type of analysis. The analysis is characterized by including 'missing' criteria which have relevance for the project being examined. These missing criteria usually concern aspects which cannot be evaluated directly in a conventional CBA, but have the potential to improve the final appraisal. In the following, these criteria will be called the MCA part of the COSIMA analysis.

COSIMA consists of an MCA and a CBA part. The outcome of the analysis can be expressed as the total value (TV) based on both parts. This specific model set-up requires that the MCA part is additive to the CBA part. By applying the COSIMA analysis, the project being examined, A_k, is presented in a more satisfactory way by the $TV(A_k)$ than by the net present value of all consequences from the CBA, denoted here as the $CBA(A_k)$. The principles of COSIMA are expressed in (2), where $MCA(A_k)$ represents the MCA part.

$$TV(A_k) = CBA(A_k) + MCA(A_k) \qquad (2)$$

The MCA part is based on the involvement of the stakeholders of the project being examined, in contrast to the CBA part. On the basis of the above, it can be concluded that in a situation with a CBA where $CBA(A_k)$ is not greater than the investment costs C_k ($CBA(A_k)) \leq C_k$), the project will not be economically feasible. However, the investment can still be justified by COSIMA if the total value of the project is greater than the investment costs ($TV(A_k) > C_k$). This can be expressed by $TRR(A_k) > 1$:

$$TRR(A_k) = \frac{TV(A_k)}{C_k} = \frac{1}{C_k} \cdot \left(\sum_{i=1}^{I} V_{CBA}(X_{ik}) + \alpha \cdot \left[\sum_{j=1}^{J} w(j) \cdot V_{MCA}(X_{jk})\right]\right) \qquad (3)$$

$V_{CBA}(X_{ik})$ is the value in monetary terms of the CBA impact *i* for alternative *k* for all I CBA impacts.

$V_{MCA}(X_{jk})$ is a value function for the MCA criterion *j* for alternative *k* for all J MCA criteria.

α is a calibration factor, which expresses the balance between the CBA and MCA part in this specific model set-up.

w(j) is the criterion weight for criterion *j*.

As mentioned above, COSIMA is a model which uses additive value functions for scoring each criterion with respect to each of the alternatives examined. In the work described in this paper, this was done by using pairwise comparisons, but other methods could be used instead. It should be mentioned that COSIMA is a decision support tool for assessing the most desirable alternative among a finite set of alternatives.

2.3 Appraisal framework

The framework for the appraisal methodology proposed is presented in Figure 1.

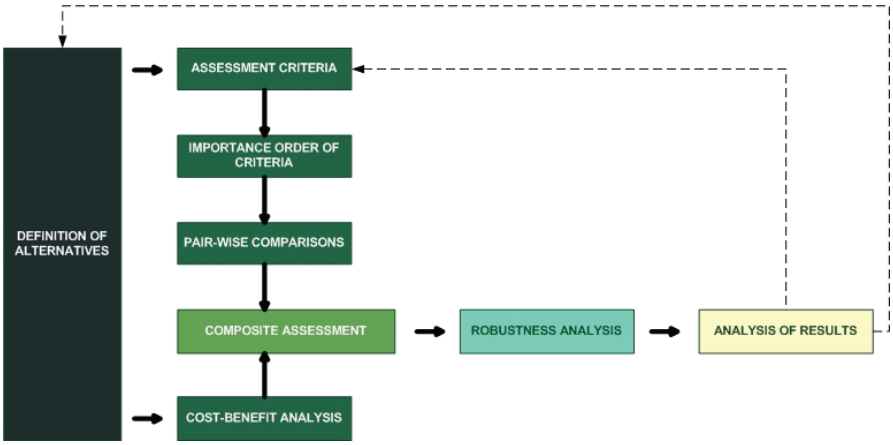


Figure 1 Framework for the appraisal methodology

The basis of the framework is a set of predefined alternatives. All of these alternatives undergo a cost-benefit analysis. Furthermore, a set of assessment criteria are defined. The criteria are ranked in order of importance and each of the criteria is assessed using pair-wise comparison. Using the COSIMA methodology, a composite assessment is made and then the robustness analysis is performed. At the end, the result of the total analysis is presented for the decision makers. If the decision makers feel the need, further iterations of the process can be conducted in which changes can be made to the assessment criteria, the ranking of the criteria, or/and the pair-wise comparisons. If a new alternative is introduced at a later point, the decision makers have to carry out the process from the beginning again, including a cost-benefit analysis for the new alternative.

3 Case studies

To test the proposed methodology, two case studies were examined. The case studies were conducted under the project "An overall evaluation of public transport investments" (Hiselius et al., 2009). In both cases, a decision conference was conducted to analyse the stakeholders' preferences. Both of the case studies were appraisals of public transport systems.

3.1 Case study 1: Ostlänken

A new high-speed railway line is planned in Sweden connecting Stockholm with Göteborg and Malmö. The case study concentrated on the link between Norrköping and Backeby. 4 different corridors were identified and the decision problem was to choose between one of these 4 corridors.

However, each of the four alternatives has both strengths and weaknesses with regard to the multiple objectives which the railway link seeks to meet. A cost-benefit analysis (CBA) was conducted prior to the decision conference. The CBA included the following impacts: investment costs, maintenance costs, operation costs, changes in revenue and travel time. The results of the CBA are shown in Table 1. All of the proposed alternatives have a benefit cost rate (BCR) higher than 1, which means they are all economically feasible alternatives.

Table 1 CBA results for planning corridors for Ostlänken

Corridor	Description	BCR
Red	Follows the existing highway	2
Blue, short tunnel	Follows the existing railway, with a short tunnel going into Norrköping	1.77
Blue, long tunnel	Follows the existing railway, with a long tunnel going into Norrköping	1.54
Green	Goes through untouched natural landscape	1.45

The decision conference was held with delegates from the Swedish railway agency and representatives from the municipality of Norrköping, and it took place in Norrköping. Because of constraints on the time available for the conference, it was decided to use already well documented material as background for the decision conference. The background material used was the Environmental Impact Assessment (EIA), which handles a number of criteria and assesses each alternative's impacts on each criterion. Before the conference, it was decided to work with the criteria found in the EIA and use the assessments made in the report as a starting point for the discussion.

Table 2 Criteria from the EIA report

Criteria	Description
City and scenery impression	Visual intrusion
Cultural environment	Cultural historic development and important cultural environment
Nature	Protection and preservation of important natural and biological diversity
Recreation and outdoor life	Outdoor activities and recreational areas
Health	Barrier effect, noise, vibration, electromagnetic fields and air quality
Natural resources	Raw material, water and climate
Risks and safety	Risk of accidents on and around the railway, transportation of dangerous goods
Disruption (construction period)	Maintaining the other functions of the surroundings during the construction

After a presentation of the CBA and the criteria from the EIA report (Table 2), the delegates were asked to rank the criteria in order of importance. After some discussion they were able to agree on a ranking of the criteria (Table 3).

Table 3 Criteria weights

Criteria	Importance order	ROD weight
Town and landscape scenery	1	0.23
Cultural environment	2	0.20
Nature	3	0.17
Recreation and outdoor life	4	0.14
Health	5	0.11
Natural resources	6	0.08
Risks and safety	7	0.05
Disruption (construction period)	8	0.02

For later sensitivity analysis, each delegate's individual ranking of criteria was also recorded. This step was also taken to ensure fairness, so that each set of weights proposed can be applied later (Labbouz et al., 2008). Then the pair-wise comparisons of the alternatives with respect to each of the criteria were made, and a linear value function was used to transform the comparisons into a performance score on a scale from 0 (worst) to 100 (best). At this stage of the decision conference, CBA and MCA scores had been made available and the COSIMA methodology was then used to combine these two scores into an overall desirability score for each alternative. To combine the two scores, a calibration indicator has to be set (Salling et al., 2007). The delegates of the decision conference agreed to set the indicator to 50%, resulting in an even weight being given to the CBA and MCA scores. The resulting total scores of the alternatives with the calibration indicator set to 50% can be seen in Figure 2.

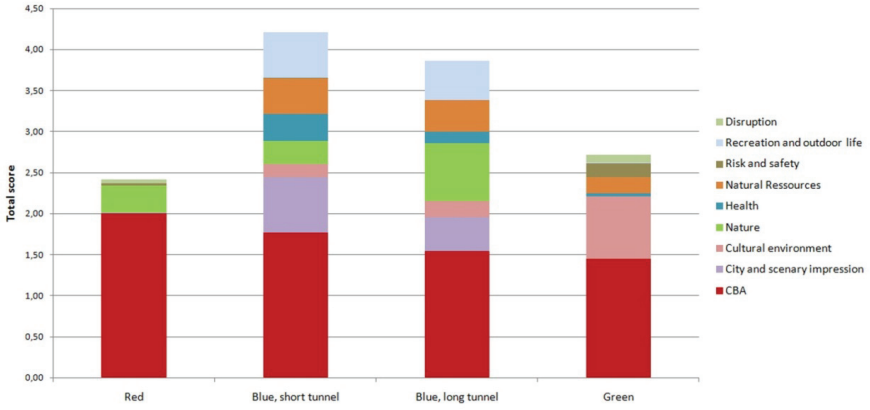


Figure 2 Total scores for the 4 alternatives with a calibration indicator set to 50%

As shown in the figure above, the ranking of the alternatives was: Blue, short (4.21), Blue, long (3.86), Green (2.71) and Red (2.41). However, this ranking of alternatives does not provide any information on how likely it is that a change in the rankings of the alternatives might occur with a change in the criteria weights (Hyde et al., 2005). So an analysis of how robust the decision analysis is with regard to changes in the criteria importance order would be useful. Figure 3 shows the total score for each of the alternatives with the calibration indicator set to 50% for all possible rankings of the criteria and thereby all possible criteria weights. The points marked 'Actual' show the total score when the criteria weights agreed at the decision conference are used. The figure also shows how sensitive the various alternatives are to changes in the criteria weights, with alternative 1 being the most robust alternative and alternative 4 the most sensitive alternative with regard to changes in the criteria weights (the greatest span between min and max).

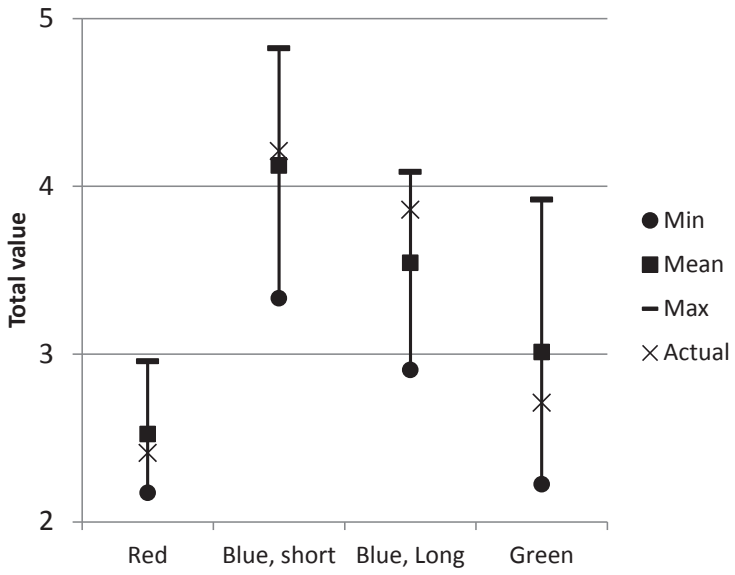


Figure 3 Desirability scores for the alternatives with the calibration indicator set to 50%

The results from the robustness analysis show that alternative 2 (the Blue short tunnel corridor) is most likely to be the most desirable alternative, but alternatives 3 and 4 can also compete when certain criteria weights are chosen. Therefore a further examination of these 3 alternatives is recommended and maybe a more constrained robustness analysis should be conducted with new input from the delegates of the decision conference.

3.2 Case study 2: Malmö light rail

A new public transport system is to be implemented in the town of Malmö, in the southern part of Sweden. The new transport system consists of several light railway lines which are planned to be introduced in stages, and in the first stage, 3 different lines have been selected for further examination. The 3 lines run in different parts of the city and therefore differ in the impacts they have. As a preliminary analysis, the municipality of Malmö had a cost-benefit analysis (CBA) conducted (Hiselius et al., 2009). The CBA was carried out in accordance with the Swedish national standard for transport project appraisal and the impacts included in the analysis were: investment costs, maintenance, operational costs, revenue, travel time, impacts on other modes and land value.

Table 4 Results from the CBA of the 3 light railway lines

Line	BCR
A	2.26
B	3.28
C	1.96

All three alternatives have a BCR over 1 and are therefore economically feasible alternatives. In addition, the municipality also conducted a wider analysis of the 3 lines covering impacts which are not included in the CBA. Not all these impacts could be assessed in a monetary way. This wider analysis of the 3 alternatives was carried out by a working group formed by the municipality and consisted of 10 different stakeholders. However, the municipality had difficulties in making an overall assessment of the 3 lines which took all the impacts into consideration. Consequently, there was a need for a more comprehensive assessment and a decision conference was held to undertake this assessment. All the 10 participants at the decision conference were stakeholders of the project and members of the working group mentioned above.

One of the main focus areas of the decision conference was to clarify the additional criteria that were to be included in the MCA. The working group had produced a long list of criteria which was first narrowed down to 8 criteria in a process in which care was taken to ensure that there was no overlap with the CBA, the criteria were defined as mutually preference-independent, and the group had a common understanding of each criterion. With 10 participants at the decision conference, it took some discussion before a ranking of the criteria in order of importance could be agreed on. The conference decided on the following criteria to be implemented in the analysis (in preference order): capacity, urban renewal, social integration, compact city, the environment, corresponding public transport, impact on car traffic and, finally, impact potential.

As with case study 1, pair-wise comparisons were made for each of the 8 criteria, which resulted in a desirability score for each alternative.

To test the result's robustness for different rankings of the criteria, all possible combinations of how the 8 criteria can be ranked were calculated. The 8 criteria can be ranked in 40,320 different ways and the group consensus on the ranking was only one out of all these possibilities. The simulation was carried out with the same pair-wise comparisons and same CBA information. Furthermore, all combinations were calculated with the calibration indicator in the range from 0% to 80%. A simulation result for alternative B, the alternative with the highest total score, is illustrated in Figure 4.

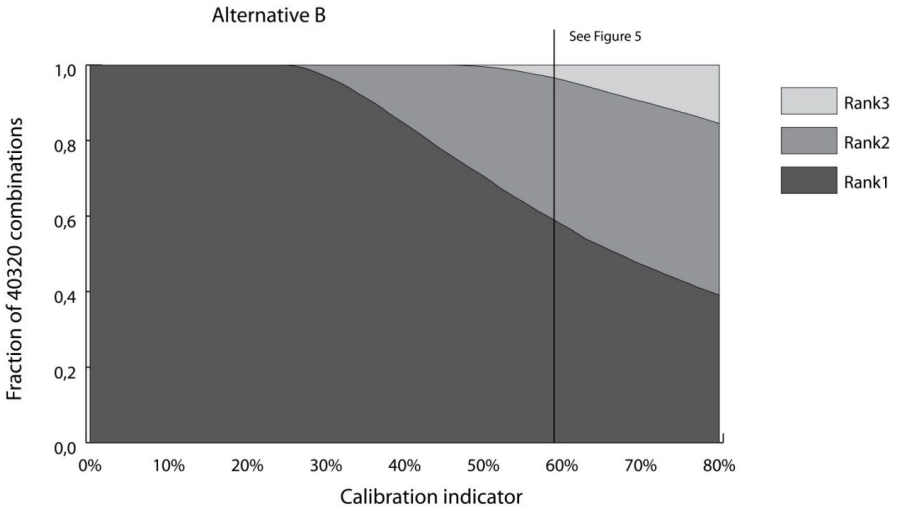


Figure 4 Simulation results for alternative B

The simulation results are illustrated in a way that enables the decision makers to see immediately how robust the total score of the alternative is in regard to changes in the ranking and the weighting of the criteria. For alternative B, it can be seen that if the calibration indicator is in the range of 0% to approximately 25%, it does not matter how the criteria are ranked – alternative B will be ranked as number 1 among the alternatives in all cases. In the case of a calibration indicator set to 60%, as selected by the group, alternative B is ranked as the most attractive alternative in approximately 58% of all possible rankings. The group did choose a ranking which put alternative B as the highest ranked alternative, but other rankings of the criterion would put either alternative A or C as the most desirable alternative. Similar illustrations could also have been made for the other alternatives.

However, the decision conference revealed a tendency in the ranking of the criteria, and this information could be used in the simulation. All the participants' individual ranking of the criteria was recorded and they all ranked the same criterion as the most important, so a simulation with this criterion fixed as ranked number 1 could be made. There was also a tendency with regard to the criterion that was ranked as second most important. These tendencies can be simulated, and the results can either confirm that the decision makers (in this case the participants of the decision conference) have a robust result that indicates one particular alternative as being the most desirable or that the result is sensitive to changes in the ranking of the criteria. Figure 5 illustrates the results from 3 different simulations with the calibration indicator at 60%. The first bar shows the simulation results with all combinations, where in 58% of all the combinations alternative B would be ranked as number 1. The next bar is the result from a simulation with the capacity criterion fixed as ranked number 1. It can be seen that the fraction of all combinations giving alternative B the highest rank is lower than in the simulation with all combinations. The reason for this is that alternative B has the

lowest score among the alternatives for criterion 1 and ranking this criterion as number 1 is then a disadvantage for alternative B. The next bar shows the results for the simulation with capacity and urban renewal criteria fixed as ranked as numbers 1 and 2. This simulation shows a higher fraction of all possible combinations giving alternative B the highest rank (61%). The reason for this can again be found in how alternative B scores on criterion 2. Alternative B has the highest score on the urban renewal criterion, together with alternative A. This gives alternative B an advantage when this criterion is fixed as the second most important criterion.

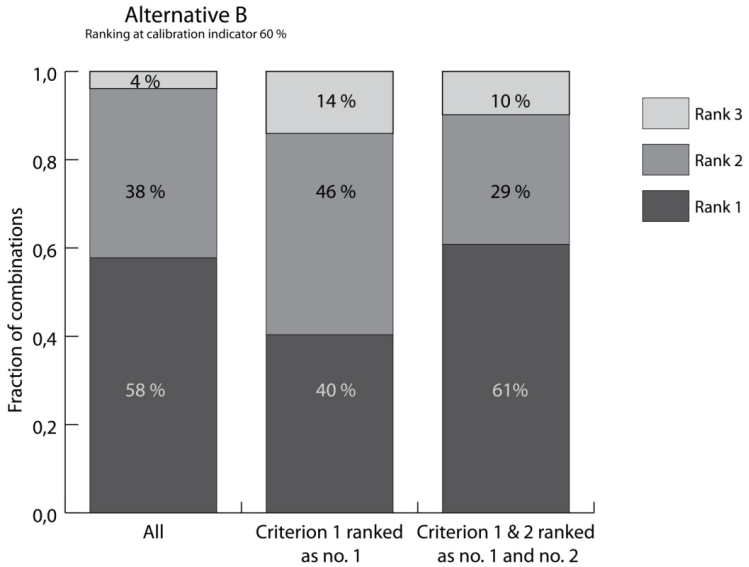


Figure 5 Constrained simulation approach with one or two criterion rankings fixed

This approach provides an opportunity to deal with disagreement among the participants about the ranking of the criteria. If they have difficulties in ranking one or more criteria, a simulation can be made based on keeping all the agreed upon rankings fixed.

4 Discussion

The output of the decision conference is a set of criteria and criteria weights together with assessments of the alternatives with respect to each criterion. The set of criteria and the assessments of alternatives can be seen as the more objective part of the decision problem and defining the decision space; in fact this step can be assessed by experts who have no interest whatsoever in the project being assessed (Hyde et al., 2005). However, the setting of the criteria weights opens up the analysis to a more definite amount of subjectivity (Zietsman et al., 2006) and each setting defines a specific subset of the decision space.

For each alternative examined, an interval score is calculated by changing the calibration indicator used by the COSIMA methodology to combine the CBA and MCA. Knowing how robust the calculated ranking of the alternatives is to changes in the criteria weights is seen as important information for the decision makers. This is discovered by examining all possible rank orders of the criteria, and thereby all possible criteria weights, for a value of the calibration indicator determined by the decision makers. If more than one decision maker takes part in the decision conference, both the calibration indicator for the group of decision makers and for each individual can be examined together. The whole decision space established by the decision conference is then examined, taking into account both the joint and the individual decisions on the calibration indicator.

This paper proposes two variants of a robustness analysis using the rank order of the criteria to estimate the robustness of the decision analysis. In contrast to other techniques for robustness analysis of the weights, the methodology proposed changes all of the weights at once and not just a single weight. The first variant of the robustness analysis methodology is illustrated in case study 1, while case study 2 widens the analysis by taking into account all possible calibration indicators. When the latter robustness analysis is applied, a clear picture of the total decision space is made available for the stakeholders involved as well as the subset the decision conference established. If the group of stakeholders can only agree on one or more criteria weights, a constrained approach is available for examining the robustness of the criteria weights. This approach gives the stakeholders an option of locking some of the criteria weights and only examining alternative weights for the remaining criteria. The outcome of the procedure for examining the importance of criteria weights is a very clear picture of how the project performs as a function of the calibration indicator.

In case study 2, the delegates were asked to rank the criteria in the group before the individual ranking. It can be discussed if this sequence had some impact on the way the different stakeholders ranked the criteria. The reason for this approach was that the main goal of the decision conference was to reach a consensus about the ranking, not to specify the individual preferences. However, the correlation between the group ranking and the individual rankings is so significant that it is clear that the group ranking must have had some impact on the individual rankings. Therefore, it is recommended that the order is changed.

What is a successful degree of consensus among the decision makers is successful? This might be a valid question about this approach. No doubt the answer would be that this is to be decided by the decision makers involved. Complete consensus would clearly make the decision-making more straightforward, but is hardly a goal in itself. In many cases, complete consensus is not achievable, and in these cases the methodology proposed is a valuable tool. This is because the methodology can show whether the failure to achieve consensus has any critical influence on the ranking of alternatives. It should be mentioned that the use of a multi criteria decision model, as proposed in this

paper, would increase the probability of reaching consensus by encouraging communication between the individuals involved (Hall, 2008).

The time available for each of the two decision conferences was only 4 hours. This time constraint had an influence on how much material it was possible to include. As a consequence, both decision conferences were tailored to be able to cover the most critical aspects of the decision problem. The main concern in Case study 1 was the order of importance of the criteria and the pair-wise comparison of the alternatives with respect to the criteria already described. In case study 2, the main concern was to establish well-defined criteria which did not overlap with the CBA and then establish the order of importance of these criteria. In spite of the limited time for conducting these decision conferences, the delegates said that they found the process very useful and that it gave them good input to their further work on the project. As one of the delegates put it: "The time available was enough to conduct a useful analysis and a lot more time would not necessarily have been a lot better".

The proposed method for robustness analysis applied to the COSIMA assessment methodology is based on the fact that the most subjective part of the analysis, and the part that is most up for discussion, is the setting of criteria weights. The robustness analysis examines how robust the outcome of decision analysis is to changes in the criteria weights. Two variants of the method have been illustrated by two case studies: one in which only the criteria weights were changed and another in which the calibration indicator as well as the criteria weights were changed.

The two case studies have shown that value is added to the decision analysis if a robustness analysis is made to examine the consequences of criteria weights different from those decided on at the decision conference. This can be especially useful in public decision-making, because of the large amount of attention there is, e.g. on transport projects. This methodology provides the decision makers with a tool to examine other preferences found, for example, in various interest groups, making it possible to accommodate their positions.

The combination of a technical solution (the COSIMA model) that captures the different perspectives with a social process (the decision conference) that engages the stakeholders led to a successful process in which the stakeholders were confident about the final results.

It can be concluded that the proposed robustness analysis, used in combination with the COSIMA methodology, provides a useful tool for examining how robust the outcome of decision analysis is with regard to changes, allowing the exploration not only of the calibration indicator, but also changes in the order of importance of the criteria and therefore criteria weights.

Acknowledgements

The author wishes to thank the participants in the two decision conferences for making this research possible and the reviewers of this paper. Thanks also to the Swedish Governmental Agency for Innovation Systems for funding this research.

References

Alfares HK and Duffuaa SO (2008a) Determining Aggregate Criteria Weights from Criteria Rankings by a Group of Decision Makers. *International Journal of Information Technology and Decision Making* 7(4): 769.

Alfares HK and Duffuaa SO (2008b) Assigning cardinal weights in multi-criteria decision making based on ordinal ranking. *Journal of Multi-Criteria Decision Analysis* 15(5): 125-133.

Belton V and Stewart JT (2001) *Multiple Criteria Decision Analysis: An Integrated Approach*. London: Kluwer Academic Publishers.

DeCorla-Souza P, Everett J, Gardner B and Culp M (1997) Total cost analysis: An alternative to benefit-cost analysis in evaluating transportation alternatives. *Transportation* 24(2): 107-123.

Edwards W, Miles RF and Von Winterfeldt D (2007) *Advances in Decision Analysis: from Foundations to Applications*. Cambridge: Cambridge University Press.

Goodwin P and Wright G (2004) *Decision Analysis for Management Judgment*. Chichester: John Wiley & Sons.

Hall DJ (2008) Decision Makers and Their need for Support. In: Burstein F and Holsapple CW (eds) *Handbook on Decision Support Systems 1*. Berlin Heidelberg: Springer-Verlag, 83.

Hiselius LW, Barfod MB, Leleur S, Jeppesen SL, Jensen AV and Hjalte K (2009) *Helhetsorienterad Utvärdering Avkollektivtrafikåtgärder*. Lunds Universitet, Tekniska Högskolan i Lund.

Hyde KM, Maier HR and Colby CB (2005) A distance-based uncertainty analysis approach to multi-criteria decision analysis for water resource decision making. *Journal of Environmental Management* 77(4): 278-290.

Jeppesen SL (2010) Use of short decision conferences (DC) in Systemic Intervention. *Systemist* 32(1).

Jimenez A, Rios-Insua S and Mateos A (2004) A decision support system for multiattribute utility evaluation based on imprecise assignments. *Decision Support Systems* 36(1): 65.

Jones D (2011) A practical weight sensitivity algorithm for goal and multiple objective programming. *European Journal of Operational Research* 213(1): 238-245.

Labbouz S, Roy B, Diab Y and Christen M (2008) Implementing a public transport line: multi-criteria decision-making methods that facilitate concertation. *Operational Research* 8(1): 5-31.

Montibeller G and Franco LA (2011) Raising the bar: Strategic multi-criteria decision analysis. *Journal of the Operational Research Society* 62(5): 855-867.

Musso E, Sanguinetti S and Sillig C (2007) Socio-economic impact of transport policies: an institutional approach. In: Haezendonck E (ed) *Transport Project Evaluation - Extending the Social Cost-Benefit Approach*. Cheltenham: Edward Elgar, 95-114.

Phillips, L.D., Bana e Costa, C. A. (2005) Transparent prioritization, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing. Working paper LSEOR 05.75

Roberts R and Goodwin P (2002) Weight approximations in multi-attribute decision models. Journal of Multi Criteria Decision Analysis 11: 291-304.

Salling KB, Leleur S and Jensen AV (2007) Modelling decision support and uncertainty for large transport infrastructure projects: The CLG-DSS model of the Øresund Fixed Link. Decision Support Systems 43(4): 1539-1547.

Thomopoulos N, Grant-Muller S and Tight MR (2009) Incorporating equity considerations in transport infrastructure evaluation: Current practice and a proposed methodology. Evaluation and Program Planning 32(4): 351-359.

Zietsman J, Rilett LR and Kim S (2006) Transportation corridor decision-making with multi-attribute utility theory. International Journal of Management and Decision Making 7(2-3): 254.

8.8 Overview of paper 3

Stability intervals

Title: The Robustness in assessment of strategic transport projects

Authors: Anders Vestergaard Jensen and Inga Ambrasaite

Presented at: The 21nd International Conference on Multiple Criteria Decision Making, Jyvaskyla, Finland, 2011.

9.6.1 Introduction to paper

In this paper the approach for including the CBA information in the MCDA has changed from the previous two papers. The CBA information obtained prior to the conducted decision conference is included directly in the MCDA as a criterion. However, to deal with uncertainty in the construction costs and travel time savings, risk assessment has been carried out by the use of MCS. The output of these risk assessments are certainty values expressing the socio-economic robustness of each alternative. These certainty values along with a probability density function for each alternative form the input for a pair-wise comparison of the alternatives with respect to the criterion socio-economic robustness.

The applied sensitivity analysis reveals the critical criteria for the decision problem and assesses the changes needed for these criteria in order to alter the initial ranking of alternatives.

9.6.2 Main findings

This paper introduces the weight stability intervals (WSIs), elaborated by Mareschal (1988), to be used actively during the DC. In order to inform the DM about the robustness of the decision analysis the WSIs are presented along with the initial assed criteria weight for each criterion that has been identified as being critical, thus changes in the criterion weight can alter the initial ranking of the alternatives. Such a presentation of the sensitivity is presented in Figure 8-5.

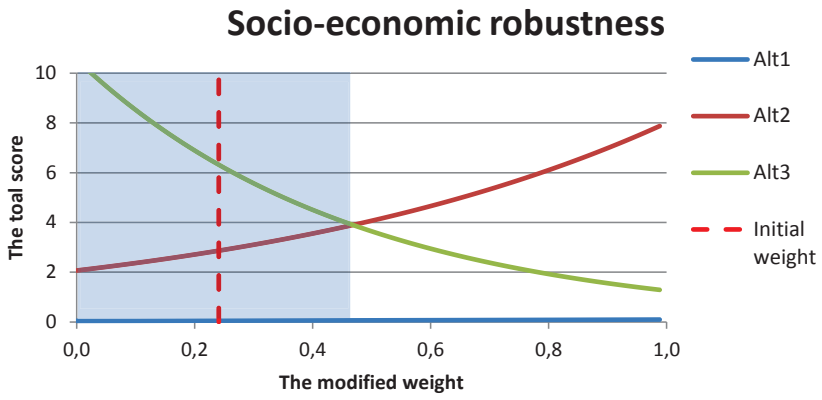


Figure 8-5 Overview of how the changes in the weight for the Socio-economic robustness criteria will affect the total scores of the three alternatives. The blue box indicates the WSI.

The use of WSI can inform the DMs if changes in the ranking of the alternative are likely in case change is made to the CWs. The WSI can even provide information on how much a CW has to change from the initially assessed weight, before a change in the ranking of alternatives occurs. Thereby the DMs have information on potential conflicts among stakeholders as a result of their different preferences. This can lead to further examination of the criteria or to alternation of the alternatives in order to solve disagreement.

This paper contributes to answering research question C as it has been demonstrated how robustness measures can mediate the decision process by providing information about critical criteria and the likelihood of them causing conflicts because of different preferences.

One important issue of the paper is the focus set on the trade-off between very refined MCDA decision models on one side and their function as practical and effective decision support tools on the other side. In order to gain the necessary confidence from the stakeholder one should not use too complex and sophisticated decision models as this could lead to erroneous input from the stakeholders if misinterpreted, resulting in invalid use of the decision support model. The task to resolve is the setting up of a successful operational DSS which employs methods that are theoretically sound and, at the same time, practical, transparent, and effective.

8.9 Paper 3 Stability intervals

The Robustness in Assessment of Strategic Transport Projects

Anders Vestergaard Jensen and Inga Ambrasaite

Abstract

This paper presents a framework for examining how subjectivity, in the form of different stakeholder strategies and stated preferences, can affect the result of an assessment of larger transport projects. In order to assure that these large investments within the transport sector will benefit the most, it is necessary to take the considerable side impacts in the decision making process into account, thus not only the economic effects form the basis of the decision. This complex challenge can be met as concerns assessment methodology with the multi-faceted presented framework, which involves risk analysis on the socio-economic part of the decision problem, Multi-criteria decision analysis (MCDA) to embrace various and often conflicting criteria and sensitivity analysis for taking into account the interests and preferences of different stakeholders. In this paper the criteria weights are examined for estimating the robustness of the different alternatives proposed. The framework is based on participatory MCDA approach by using decision conferences for identifying the preferences of stakeholders. After applying the framework to a case study we conclude that the presented sensitivity analysis can be used for examining the subjective part of the MCDA and its role in the decision support.

Keywords: Multi-criteria decision analysis; Robustness; Participatory process; Criteria weight; Decision conference.

1 Introduction

A large transport project has, in a natural consequence of being large and often influencing a large group of people, a high awareness among the society in general. Thereby, several different stakeholders or stakeholder groups will try to affect the decisions made regarding large transport projects. This together with the fact that larger transport projects also have a wider set of impacts on the society in general results in a very complex decision making problem for the decision makers (DM). When involving several stakeholders in a decision process, it can be difficult to get consensus while often conflicting interests are present (Mendoza and Prabhu, 2005). Each stakeholder will try to promote the solution or alternative that is best suited for their preferences, and it can be difficult to get agreement about a common preferred alternative. One way to overcome some of these complicated issues, is to illustrate the consequences of applying different stakeholder preferences to the decision problem e.g. with different criteria weights and re-calculating the resulting ranking order of the alternatives. This paper deals with this aspect and how it can be used actively in a decision conference (DC).

In the literature there is an agreement that the cost-benefit analysis (CBA) does not provide a sufficient base for decisions about transport projects (Thomopoulos et al., 2009) because of the complexities mentioned above and often conflicting criteria. Furthermore, Beukers et al. (2012) also has examined the CBA to be somehow a black box and not a transparent decision support. The use of multi-criteria decision analysis (MCDA) has been applied for supporting strategic decision making in organisations (Montibeller and Franco, 2011). As Munda (2008) describes it: in assessment of public policies, MCDA is an adequate policy tool, since it allows us to take a wide range of criteria into account (e.g. environmental impacts, distributional equity, ect.), and not simply profit maximization, as a private economic agent would do. Thus the essential meaning of multi-criteria evaluation in a social context is simply tolerance and democracy. Recognizing that formal economic evaluation analyses tend to be inadequate, because objectives for transport projects often are broader than pure economic or market concerns, the MCDA is a valid methodology for assessing these projects (French et al., 2005; Kapros et al., 2005).

Montibeller and Franco (2011) suggest using strategy workshops for getting a better understanding and reach consensus about a decision. Munda (2004) goes even further and states that scientists cannot provide any useful input without interacting with the rest of the society, and vice versa the society cannot perform any sound decision making without interacting with the scientists. The facilitation of such an interaction between scientists and society can be reached by using stakeholders in DC to formulate the decision problem and express preferences. A decision achieved in consensus at a decision conference has a higher possibility for being implemented than e.g. the results from a complex decision analysis that only involves one decision-maker, who later has to justify his decision to an organisation or to the general public. But the final decision is still in the hands of policy-makers and based on many other factors (Zurita, 2006). Banville et al. (1998) presents some findings on how to integrate the concept of stakeholders systematically into an MCDA approach. They define the stakeholders as those who are either both affected by and affecting the problem, those who are acting on behalf of 'clients' and representing them, or those who do not have any direct control over resources viewed as relevant for the problem (e.g. future generations will be effected by today's policies). Musso et al. (2007) also write about the participative MCDA approach and suggest that the elicitation of criteria weights should be made by the stakeholders and reflects their cultural values. This also raises questions about the estimates of values of the variables in the CBA, however, we will in this paper assume these values (or unit prices) to be well established and do not question them further. One of the main advantages of using several stakeholders in an MCDA is that the discussion about the relative importance of the involved criteria is made within the stakeholder group (Macharis, 2007). Thereby clarification about the assessed criteria weights can be made during a DC by the stakeholders who have set the weights and a common agreement among the stakeholders with regard to the relative importance of the criteria is reached.

Several different methods for undertaking sensitivity analysis in the MCDA framework have been proposed. Many have found that to examine the differences in preference weights are of high importance e.g. Salo and Hämäläinen (1992), Triantaphyllou and Sánchez (1997) and Strager and Rosenberger (2006). They acknowledge that variations in the criteria weights (preferences) can have big influence on the results of an MCDA. Mareschal (1988) presents a method for evaluating the consequences of changing the relative importance of the criteria by determining in which intervals the criteria weights can be changed without any altering of the prioritised ranking order of the alternatives. This method reveals how robust the different alternatives are for changes in the criteria weights. Jimémez, et al. (2004) describes how these weight stability intervals (WSI) can be implemented in a decision analysis tool.

Traditionally the CBA has been widely used for assessing transport projects as it is assumed to be an objective decision support tool. However, as stated above the CBA is not sufficient to identify and evaluate all the impacts of a transport project. The MCDA has been suggested to contribute to a more comprehensive assessment. In order to accommodate the critique of the MCDA in regards to its subjectivity (Browne and Ryan, 2011) there is a need to research this subjective part of the decision processes and its role in the decision support. We present a hybrid approach based on CBA, MCDA and WSI of criteria weights for evaluating strategic transport projects. CBA is used for assessing the monetary impacts. MCDA is used to non-marketable and qualitative criteria take into account in the decision making. Sensitivity analysis is used for dealing with the different preferences in the decision process of strategic transport projects. The framework is based on participatory MCDA approach by using decision conferences for identifying the preferences of stakeholders. To illustrate the framework a case study is presented, were the use of WSI are applied in an MCDA using a multiplicative version of Analytic Hierarchy Process (AHP) named REMBRANDT (ratio estimation in magnitudes or deci-bells to rate alternatives which are non-dominated) (Lootsma, 1992). The MCDA of the case study is conducted with input from a DC were the sensitivity analysis is used actively for achieving consensus among the stakeholders. The focus of the paper is on: dealing with subjectivity in assessment of larger transport projects, transparent decision making process and consensus and communicative issues.

This paper is structured in the following way. After this introduction the framework is introduced more specifically. Hereafter, the case study is presented and used to demonstrate the applicability of the sensitivity method in DC. The paper ends with a discussion on how the framework can support a participatory MCDA approach for strategic transport projects as well as the main conclusions and perspectives for the future research.

2 Methodology

The framework of the presented methodology is illustrated in Fig. 1. We assume that in previous analysis a set of alternatives has been generated and all alternatives are valid candidates for fulfilling the purpose of the transport project. The socio-economical

input is a CBA which is conducted and a benefit-cost ratio (BCR) is calculated for each alternative. In order to deal with inaccuracies in the construction costs and travel time savings risk assessment is carried out by use of Monte Carlo Simulation (Salling and Leleur, 2011).The output of this feasibility risk assessment is a certainty value for each alternative, describing the certainty for obtaining a BCR higher than 1 and thereby socio-economic feasible. However, as input from the CBA a BCR can be used instead although one will not get the useful information about how confident the BCR has been calculated. A group of stakeholders are then gathered at a DC to combine technical and social aspects for solving the complex decision problem. The main purpose of the DC is to enable a structured debate between the stakeholders and form a common basis for discussing the decision problem at stake. The DC consists of four steps:

(1) Definition of MCDA criteria: The first task for the stakeholders at the DC is to define a set of criteria for evaluating the different alternatives. In this step it is essential that the facilitator ensures that there will be no overlapping of the criteria, as this can result in impacts being counted for twice in the analysis. This step is a very important step since it can have a big influence on the final outcome of the MCDA, and therefore necessary time should be allocated to this step.
(2) Pair wise comparisons: The score of each alternative with respect to each criterion is determined by pair wise comparisons by using the REMBRANDT methodology. The alternatives are compared two by two under each criterion and assigned with numerical values based on preference.
(3) Relative importance of criteria: The criteria are ranked by Rank Order Distribution (ROD) weights (Roberts and Goodwin, 2002). It is considered to be very difficult to make the stakeholders agree directly upon a weight set of criteria. Therefore all the different weight sets provided by each stakeholder individually, is recorded for later examination.
(4) Sensitivity analysis: The sensitivity analysis is conducted by examining the individual weight set and in that way examining if the difference in preference at the DC, will influence the final ranking of the alternatives. Furthermore for each criterion a weight stability interval is calculated. The results of these sensitivity analysis' are presented for the stakeholders.

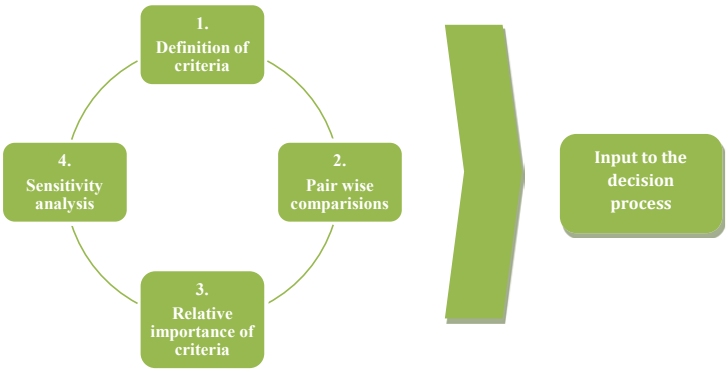


Figure 1 The four steps of the decision conference and the framework of the presented methodology

The DC should be seen as an iterative process where new perceptions about the issues might lead to successive revisions. This iterative and interactive process proceeds until a shared understanding of the issues is obtained. When the stakeholders have reached some sort of agreement, the final ranking of the MCDA can form an input to the decision making process. It is important to notice here that the stakeholders do not have to agree on all of the DC components, but a partial agreement can as well be a valid result from the DC process. Also it should be noted that a DC can be a time consuming process and it is possible, as the case study will illustrate later, that some of the steps of the DC can be conducted in advance e.g. in another DC.

2.1 Robustness measures

The method used in the MCDA is the REMBRANDT method, which is a multiplicative version of the well-known Analytic Hierarchy Process (AHP) developed by Saaty (1977) which attempts to overcome some of the theoretical difficulties associated with the original AHP (Barfod, 2012; Belton and Stewart, 2002). As with other MCDA methods the REMBRANDT method also requires an assessment of the relative importance of the criteria (criteria weights).

The weight of each criterion is a positive number, which represents the relative importance of the criterion for the stakeholder. For assessing the robustness of the MCDA we use the weight stability intervals (WSI) method presented by Mareschal (Mareschal, 1988). The method investigates how the desirability of the alternatives changes, when all the weights of the criteria are kept constant except for one criterion. The weights w_j are normalized to 1:

$$\sum_{j=1}^{k} w_j = 1 \qquad (2.1)$$

The weight w_i may be increased or decreased from its initial value. The bounds within which stability is achieved form the WSI for criterion f_i. We change the relative importance (weight) of criterion f_i (w_i):

$$w'_i = (1 + \beta) \cdot w_i \tag{2.2}$$

All other weights are adjusted so only the importance of f_i relative to the other criteria is modified:

$$w'_j = \alpha \cdot w_j \tag{2.3}$$

And are related and constrained as follows:

$$\alpha = \frac{1-(1+\beta)\cdot w_i}{1-w_i} \tag{2.4}$$

$$-1 \leq \beta \leq \frac{1-w_i}{w_i} \tag{2.5}$$

$$0 \leq \alpha \leq \frac{1}{1-w_i} \tag{2.6}$$

The method ensures that when changing the relative importance of criterion *i*, the relative importance of all the other criteria are unchanged. By only analysing one importance weight at a time, makes the sensitivity analysis more transparent and easier to communicate to the stakeholders at the DC. The method examines to which extend the weight has to be changed from the initially weight, before a shift in the ranking order of the alternatives occur. The initially weight is the weight the stakeholders agreed upon in the DC. A large change in the criterion weight from the initial weight indicates a robust MCDA in regards to the importance of the examined criterion. Some criterion will be fully stable, meaning that by changing the weight for the criterion it is not possible to change the order of the final ranking of the alternatives.

The sensitivity analysis will be conducted, as earlier described, as the last step in the DC along with the presentation of the final ranking of the alternatives.

In the next section, we present a case study which will illustrate the use of the proposed MCDA framework.

3 Results

In the previous section an MCDA framework has been described with a methodology for performing sensitivity analysis with regards to the subjective inputs from various stakeholders. A real world case study involving a new fixed link between Denmark and Sweden has been conducted to test this methodology. The decision problem was part of the EU Interreg project EcoMobility examining e.g. the possibilities for establishment of green logistics and sustainable supply chains in the Øresund region.

The Øresund fixed link between Denmark and Sweden, which opened in 2000, has led to a strong increase in traffic across Øresund. In 2009, an average of 19,500 vehicles and 184 trains crossed the link per day, corresponding to 141% and 125% increase respectively compared to the first full year of operations in 2001 (Øresundsbro Konsortiet, 2010). Moreover, the fixed link across Fehmarn Belt between Denmark and Germany which is expected to open in 2018 will increase these numbers due to more travellers from central Europe through Denmark to Sweden and Norway. Especially, the number of freight trains through Denmark is expected to grow significantly, turning the Øresund fixed link into a bottleneck, because the existing capacity is already close to the limit.

Several different proposals for an additional fixed link between Elsinore and Helsingborg in the northern part of Zealand, Denmark, has been made in the recent years. In the autumn of 2011 three tunnel projects are considered as the main alternatives (Table 1). All are expected to relieve the existing southern fixed link for traffic and thereby free up capacity.

Table 1 Three proposed alternatives for the HH-connection (Larsen and Skougaard, 2010)

HH-connection	Description	Cost (mill. DKK)
Alternative 1	Tunnel for rail (2 tracks), passenger trains only.	9,500
Alternative 2	Tunnel for rail (2 tracks), passenger trains only + tunnel for vehicles (2 × 2 lanes).	24,500
Alternative 3	Tunnel for rail (2 tracks), passenger trains + tunnel for vehicles (2 × 2 lanes) + tunnel for rail (single track), goods trains.	32,500

Socio-economic analysis (CBA) has been undertaken including the following impacts: construction costs, maintenance, generalized user benefits (time and vehicle operating costs), environment and revenue. The CBA reveals that all of the alternatives are socio-economic feasible, all with a BCR over 1. As described earlier to deal with inaccuracies

in the construction costs and travel time savings, risk assessment is carried out by use of Monte Carlo Simulation (Salling and Leleur, 2011). The output of this feasibility risk assessment is a certainty value for each alternative, describing the certainty for obtaining a BCR higher than 1 and thereby socio-economic feasible. This certainty value expresses the socio-economic robustness of each alternative.

A decision conference (DC) was arranged for getting input from a stakeholder group comprised of experts from both Sweden and Denmark. All of the four experts have many years of experience with the transport area and the Øresund region. Prior to the DC, five other criteria, besides the socio-economic robustness criterion, were defined by the research team. This resulted in a criteria list containing the following six evaluation criteria: Socio-economic robustness, Improvement for passenger cars and public transport, Impacts on towns and land-use, Impact on regional economics, Impact on flexibility in logistics and Contribution to EU green corridors. The reason for defining the criteria in advance was merely a question of available time for conducting the DC. The DC was planned to take no more than one day and the research team judged that this time frame did not allow for defining the criteria. However, the criteria list was sent out to all the stakeholders in advance of the DC. At the DC the stakeholders were given the option for making changes in the criteria list either by deleting or adding criteria if they felt the criteria list did not cover the decision problem adequately. However, the stakeholders did not make use of this option.

During the DC the stakeholders performed the pair wise comparisons for assessing the three alternatives with respect to all of the six criteria. Afterwards they ranked the criteria after importance, first, individually, then as a group. The criteria were then assigned with the surrogate ROD weights based on the rankings. In Table 2 the weighted alternative score for each criterion is listed along with the criterion weights agreed upon at the DC. The table also shows the total score of each alternative.

Table 2 Criteria weights and scores for each of the three alternatives together with the total score

Criterion	Weight	The weighted score within criterion		
		Alt1	Alt2	Alt3
C1: Socio-economic robustness	0.24	0.57	1.65	1.06
C2: Improvement for passenger cars and public transport	0.19	0.50	1.30	1.55
C3: Impact on towns and land-use	0.04	0.85	1.11	1.06
C4: Impact on regional economics	0.30	0.44	1.32	1.73
C5: Impact on flexibility in logistics	0.09	0.71	1.04	1.34
C6: Contribution to the EU green corridors	0.14	0.73	0.88	1.57
The total score:		0.06	2.87	6.32

Alternative 2 and 3 clearly surpasses alternative 1. And alternative 3 has the best total score of all three alternatives.

As mentioned above the individual ranking of the criteria was also given by each of the stakeholders (A, B, C and D) and Fig. 2 illustrates the different individual criteria weights together with the common weights.

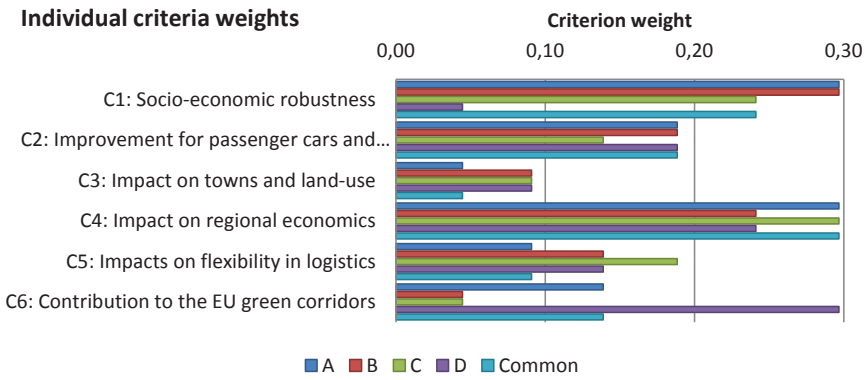


Figure 2 Individual criteria weights

From Fig. 2 it can be seen that the four stakeholders do not have the same preferences with regards to the importance of the different criteria, although there seems to be a common pattern. Most of the stakeholders have criterion C2 as the third most important criteria and criterion C3 as the fifth most important criterion. However, three of the stakeholders have the socio-economic robustness criterion as the most important criterion, but one stakeholder (D) has the criterion ranked as the least important criterion. The argument for the low priority of this criterion is that all of the alternatives are socio-economically feasible, and therefore this criterion is not given a high priority by the stakeholder. This issue was discussed at the DC and after some discussion all of the four stakeholders agreed that the socio-economic robustness criteria should have a relative high importance because of the need to justify the overall decision of implementing one of the alternatives.

After all the total scores had been presented for the stakeholders, the results from the sensitivity analysis were presented for the group. One big advantage of the WSI method is that it is fairly easy and fast to calculate the intervals and thereby the analysis can be presented immediately for the stakeholders.

As it can be seen in Table 3 only changes in the importance of criterion C1 and C3 can alter the initial ranking order of the alternatives. The weight for criterion C3 has to be increased significantly from the initial weight before a change in the ranking order occurs (see also Fig. 4). For C1 the change in the weight is still large, but not as large as for C3 (Fig. 3). From Fig. 3 and Fig. 4 it can also be concluded that the ranking of alternative 1 do not alter when changes are applied of the importance of either socio-economic robustness or impacts on towns and land-use. All of the criteria have a lower

bound equal to zero, so each of the criteria may vanish without altering the initial ranking order. All in all it can be concluded that all of the alternatives are quite robust for changes in the importance order of the criteria.

Table 3 The WSI for each of the 6 criterion in the MCDA

Criterion	Initial Weight	WSI	
		Lower	Upper
C1: Socio-economic robustness	0.24	0.00	0.47
C2: Improvement for passenger cars and public transport	0.19	0.00	1.00
C3: Impact on towns and land-use	0.04	0.00	0.49
C4: Impact on regional economics	0.30	0.00	1.00
C5: Impact on flexibility in logistics	0.09	0.00	1.00
C6: Contribution to the EU green corridors	0.14	0.00	1.00

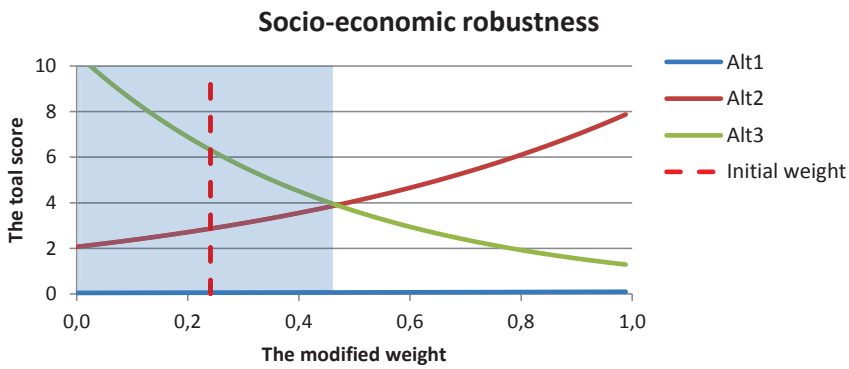


Figure 3 Overview of how the changes in the weight for the Socio-economic robustness criteria will affect the total scores of the three alternatives. The blue box indicates the WSI.

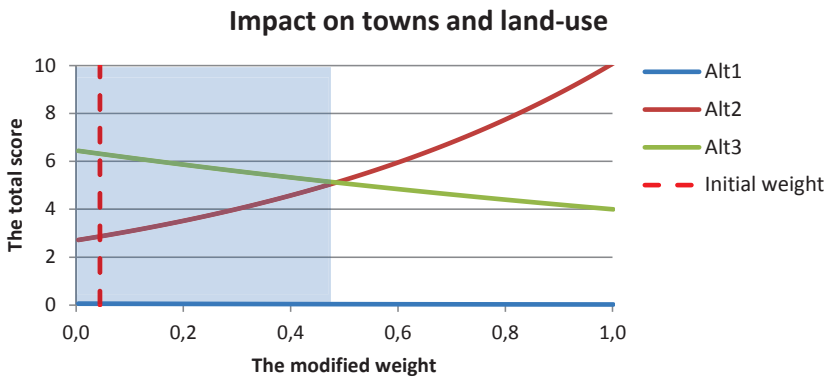


Figure 4 Overview of how the changes in the weight for the Impact on towns and land-use criteria will affect the total scores of the three alternatives. The blue box indicates the WSI.

All of the individual criteria weights are within the stability intervals and thereby it can be concluded that all of the stakeholders in the DC have the same alternative (Alternative 3) as the most preferred alternative. It should be noted that by using the ROD methodology for assessing the relative importance of the criteria it is not possible to give a criterion a weight higher than 0.29. Thereby the stakeholders did not have the possibility to assign a weight that would have resulted in a shift in the most preferred alternative.

4 Discussion

The presented methodology in this paper implements a socio-economic robustness measure in a multi-criteria decision analysis (MCDA), identifying the most desirable alternative and analyse the overall robustness of the assessment by calculating weight stability intervals (WSI) for the criteria weight. Within these intervals the overall ranking of the alternatives with respect to desirability is intact. The methodology has been illustrated by a case study concerning an assessment of a strategic transport project by the use of DC.

From the case study several conclusions can be made. First, the determination of evaluation criteria was in the case done in advance of the DC, however, the stakeholders were given the possibility to change, add or remove the criteria. It could be questioned if the evaluation criteria should come directly from the stakeholders or they should be formulated by the research team. However, it is our belief, as Munda (2004) also found, that the material collected during e.g. interviews and DC could be used as a source of inspiration. But the technical formulation of criteria is a job for the researchers in order to insure the properties such as non-redundancy and legibility. On the other hand the time limitations of a DC promote that the definition of the criteria are done in advance, and this serves as a starting point for the actual definition of the

criteria done by the stakeholders under supervision of the facilitator. Second, as an optimal weighing methodology the Swing weight method should be chosen. As Keeney (2002) notes one of the common mistakes done in making value trade-offs are assessing value trade-offs independent of the range of consequences. But when using DC with various stakeholders the swing method is not a very practical method, and often the stakeholders do not use the necessary time for assessing these weights. The case study also showed this tendency. The stakeholders did not give the weighing much attention and there was a sense of inattention among the stakeholders. It was up to the facilitator to ensure that the stakeholders gave their true preferences in given importance to the different criteria. Using a more complex method e.g. the Swing weight method would probably not solve this issue, actually it could lead to a more imprecise estimation of the stakeholders' preferences. One compromise could be that the stakeholders should give their preferences about the importance of the criteria while being informed about the range of the consequences that their pair wise comparisons gave for each criterion (the interval of the lowest and highest score for the criterion). Third, the calculation of the WSI revealed that changes in importance of only two out of the six criteria could alter the final ranking order of the alternatives. One of these criteria was the socio-economic robustness. Among the stakeholders there were variations of how important they found this particular criterion to be. However, the sensitivity analysis showed that this variation did not result in any shift in the final ranking of the alternatives because all of the stakeholders' assessed weights were within the WSI. Thereby the sensitivity analysis defused a potentially lengthy debate about the importance of this criterion. And the stakeholders could then all be confident with the final ranking of the alternatives, knowing that the impact of the subjective weights had been examined.

As remarked by one of the stakeholders at the DC, if there had been more criteria (than six) it would have been difficult to get an overview of the decision problem. Furthermore, if more than three alternatives should have been assessed the number of pair wise comparisons would increase significantly and thereby make it more time consuming. These observations give reason to remark that assessing a decision problem with several criteria and alternatives the pair wise comparison methodology as presented in this paper is probably not the most appropriate method.

One of the big advantages of the DC is also the process. After the DC the stakeholders stated that they were surprised by the fact that they achieved new insight in a familiar field. Furthermore they were very amazed that in matters where they thought they had strong preferences, in fact their preferences were not that strong. This implies that during the process of the DC the stakeholders reached a better understanding of their own knowledge and preferences. In an MCDA framework the process is of high importance and each step (e.g. a conducted DC) is an input in an iterative decision making process. As Munda (2004) clearly states it "a computation is not a decision". Consequently the output from a DC is not only the ranking order of the alternatives, but also the new insight the stakeholders achieve. And decisions achieved in consensus at a decision conference has a higher possibility for being implemented than e.g. the results

from a complex decision analysis that only involves one decision-maker who later has to justify his decision to an organisation or to the general public. But the final decision is still in the hands of policy-makers and based on many other factors (Zurita, 2006).

Overall the presented methodology proves to be a practical and valuable tool for supporting the decision making by involving stakeholders. The difficult task as Montibeller and Franco expresses it is to "employ methods that are theoretically sound and, at the same time, practical and effective decision support tools" (2011). The trade-off between very complicated decision models on one side and practical and effective decision support tools on the other side has been one of the main concerns for the development of the presented methodology. Using highly complex and sophisticated decision models makes no sense if you cannot get the confidence from the stakeholders using the models. Then there will be a risk of getting an erroneous input from the stakeholders and thereby the analysis will be based on false information. Then you are better off using less complex models, even if more theoretically sound decision models exist. Off course one always has to make sure that the applied models do not misinterpret the stakeholders' preferences.

The focus of this paper has been to formulate a framework which provides a transparent analysis, which can be communicated to DMs. This complex challenge can be met as concerns assessment methodology with the multi-faceted presented framework, which involves risk analysis on the socio-economic part of the decision problem, MCDA to embrace various and often conflicting criteria and sensitivity analysis for taking into account the interests and preferences of different stakeholders. Furthermore the presented sensitivity method can be used for examining the subjective part of the MCDA (here in the form of criteria weights) and its role in the decision support. The decision conference concept is essential for the framework where the engagement of stakeholders and their different preferences provide a common platform for understanding a decision problem and for seeking out the most attractive decision alternative.

Acknowledgments

This study has been conducted as part of the EcoMobility project and is co-funded by the European Union, Interreg IVA.

References

Banville C, Landry M, Martel J and Boulaire C (1998) A stakeholder approach to MCDA. *Systems Research and Behavioral Science* 15(1): 15-32.

Barfod MB (2012) *Optimising Transport Decision Making using Customised Decision Models and Decision Conferences*. PhD thesis, Technical University of Denmark - Department of Transport: Kgs. Lyngby.

Belton V and Stewart JT (2002) *Multiple Criteria Decision Analysis: An Integrated Approach*. London: Kluwer Academic Publishers.

Beukers E, Bertolini L and Te Brömmelstroet M (2012) Why Cost Benefit Analysis is perceived as a problematic tool for assessment of transport plans: A process perspective. *Transportation Research Part A* 46(1): 68-78.

Browne D and Ryan L (2011) Comparative analysis of evaluation techniques for transport policies. *Environmental Impact Assessment Review* 31(3): 226-233.

French S, Bedford T and Atherton E (2005) Supporting ALARP decision making by cost benefit analysis and multiattribute utility theory. *Journal of Risk Research* 8(3): 207-223.

Jiménez A, Rios-Insua S and Mateos A (2004) A decision support system for multiattribute utility evaluation based on imprecise assignments. *Decision Support Systems* 36(1): 65.

Kapros S, Panou K and Tsamboulas DA (2005) Multicriteria approach to the evaluation of intermodal freight villages. *Transportation Research Record*(1906): 56-63.

Keeney RL (2002) Common mistakes in making value trade-offs. *Operations Research* 50(6): 935-945+1077.

Larsen LA and Skougaard BZ (2010) *Appraisal of Alternatives Concerning a Fixed Link between Elsinore and Helsingborg*. M.Sc., Technical University of Denmark: Kgs. Lyngby.

Lootsma FA (1992) *The REMBRANDT System for Multi-Criteria Decision Analysis Via Pair Wise Comparisons Or Direct Rating*. Delft: Faculty of Technical Mathematics and Informatics, Delft University of Technology.

Macharis C (2007) Multi-Criteria Analysis as a Tool to Include Stakeholders in Project Evaluations: the MAMCA method. In: Haezendonck E (ed) *Transport Project Evaluation - Extending the Social Cost-Benefit Approach*. Cheltenham: Edward Elgar, 115-131.

Mareschal B (1988) Weight stability intervals in multicriteria decision aid. *European Journal of Operational Research* 33(1): 54-64.

Mendoza GA and Prabhu R (2005) Combining participatory modeling and multi-criteria analysis for community-based forest management. *Forest Ecology and Management* 207(1-2): 145-156.

Montibeller G and Franco LA (2011) Raising the bar: Strategic multi-criteria decision analysis. *Journal of the Operational Research Society* 62(5): 855-867.

Munda G (2008) *Social Multi-Criteria Evaluation for a Sustainable Economy*. Berlin: Springer.

Munda G (2004) Social multi-criteria evaluation: Methodological foundations and operational consequences. *European Journal of Operational Research* 158(3): 662-677.

Musso E, Sanguinetti S and Sillig C (2007) Socio-economic impact of transport policies: an institutional approach. In: Haezendonck E (ed) *Transport Project Evaluation - Extending the Social Cost-Benefit Approach*. Cheltenham: Edward Elgar, 95-114.

Øresundsbro Konsortiet (2010) 10 Years: *The Øresund Bridge and its Region*. Copenhagen: Øresundsbro Konsortiet.

Roberts R and Goodwin P (2002) Weight approximations in multi-attribute decision models. *Journal of Multi Criteria Decision Analysis* 11: 291-304.

Saaty (1977) Scenarios and priorities in transport planning: application to the Sudan. *Transportation Research* 11(5): 343-350.

Salling KB and Leleur S (2011) Modelling of transport project uncertainties: Feasibility risk assessment and scenario analysis. *European Journal of Transport and Infrastructure Research* 12(1): 21-38.

Salo AA and Hamalainen RP (1992) Preference Assessment by Imprecise Ratio Statements. *Operations Research* 40(6): 1053.

Strager MP and Rosenberger RS (2006) Incorporating stakeholder preferences for land conservation: Weights and measures in spatial MCA. *Ecological Economics* 58(1): 79-92.

Thomopoulos N, Grant-Muller S and Tight MR (2009) Incorporating equity considerations in transport infrastructure evaluation: Current practice and a proposed methodology. *Evaluation and Program Planning* 32(4): 351-359.

Triantaphyllou E and Sánchez A (1997) A sensitivity analysis approach for some deterministic multi-criteria decision-making methods. *Decision Sciences* 28(1): 151-185.

Zurita L (2006) Consensus conference method in environmental issues: relevance and strengths. *Land use Policy* 23(1): 18-25.

8.10 Overview of paper 4

Comparison of robustness measures

Title: Comparisons of Robustness measures- as a communicative means for involvement of Decision-makers

Authors: Anders Vestergaard Jensen

Presented at: 18th International Conference on Urban Transport and the Environment, La Coruña, Spain, 2012.

Published in: Longhurst, J.W.S. and Brebbia, C.A. (eds.) (2012) *Urban Transport XVII – Urban Transport and the Environment in the 21st Century*. Proceedings of Urban Transport and the Environment 2012 Conference, pp. 603-614, WITT Press.

9.7.1 Introduction to paper

This paper makes use of two case studies, the Östlänken and the HH. However, the Östlänken case is treated differently than in paper 2, where the COSIMA appraisal methodology was used. In this paper the CBA forms part of the MCDA as a criterion, which characterises the SIMDEC approach. The paper extends the robustness analysis presented in paper 3 to also including a stochastic approach, which consists of random CWs with or without a preferred ranking of importance of the criteria.

9.7.2 Main findings

The two case studies show that it is possible to produce a transparent and valid analysis of the robustness of an MCDA process which can be used actively by the stakeholders or DMs. By working at different levels of the decision problem and thereby seeking to simplify the SA, it is demonstrated that a clear and informative analysis can be conducted. The different levels examined in the paper are the methodology level and the parameter level.

Three types of simulations are used to assess the robustness of a decision at the parameter level: (1) Random ROD weights without importance order – simulation where criteria are assigned a random ROD weight. Two criteria can have equal weighting. (2) Random order weights – the importance order of the criteria is preserved, but the weights are randomly generated. (3) Weight interval – identifying critical criteria where a change in the weight will result in different rankings of the alternatives.

The result of a simulation using the random order weight is illustrated on Figure 8-6. In this example it is not possible to distinguish if alternative 2 or 3 will be the most preferred alternative.

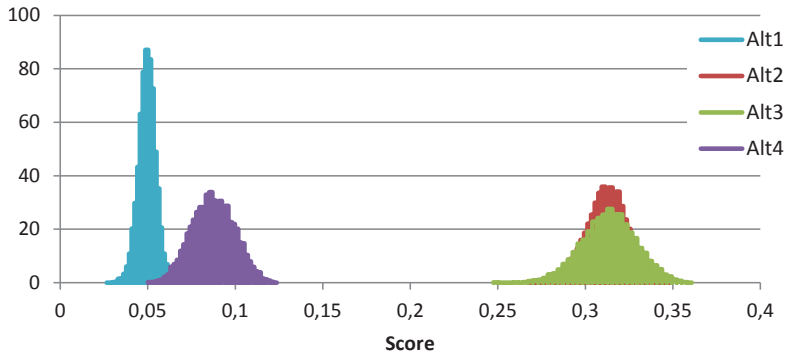


Figure 8-6 Simulation with random generated weights assigned to the criteria preserving the importance order of the criteria

Two different types of aggregation methods have been applied to the two case studies, the AHP and the REMBRANDT method. The results from the two case studies did not lead to a clear conclusion on method bias with regards to influence on results from the criteria weights. However, this finding is positive as different pairwise comparison methods in MCDA should not lead to changes in the results.

8.11 Paper 4 Comparison of robustness measures

Comparisons of Robustness measures as a communicative means for involvement of decision-makers

Anders Vestergaard Jensen

Abstract

Decisions about infrastructure projects or new policies in the transport sector have traditionally been based on cost benefit analysis (CBA). However, society in general is becoming more and more complex and this affects the decision-making process. Decision-makers are confronted with the difficult problem of evaluating potential outcomes and choosing policies to achieve the desired outcomes in the presence of this complexity. This is where multi-criteria decision analysis (MCDA) becomes a useful tool for the decision-makers, because this type of analysis can make an assessment based on a more comprehensive evaluation framework by also taking into account non-quantifiable impacts. This article concerns decision-making in transport projects that involve multiple objectives (MCDA); in particularly it addresses how to measure the robustness of these decisions with regard to views of multiple and diverse stakeholders involved in the MCDA. The paper also discusses the communicative means for involving stakeholders and decision-makers in the decision process. Specifically, it makes a comparison of various measures of decision robustness based on theory and case studies and it also discusses how these measures can be communicated to the decision-makers. Furthermore, it examines how the choice of MCDA methodology can affect the decision robustness. Finally, in summary it gives some recommendations for applying robustness measures.

Keywords: Sensitivity analysis, Multi-criteria decision analysis, Stakeholders, Transport appraisal, Rank Order Distribution, Ordinal Ranking.

1 Introduction

Today's decision-makers (DMs) are expected to be socio-economically responsible when taking decisions – but also socially responsible with regard to sustainability, equity, and so on. In most real-world policy situations, there are many alternatives, many uncertainties, many stakeholders and many consequences of interest (Walker, 2000). This, and the fact that there is usually no single decision-maker, means that getting agreement on a single set of preferences is probably not possible. This is where multi-criteria decision analysis (MCDA) becomes a useful tool for the decision-makers. MCDA should not be seen as a prescriptive answer (often there is no optimal solution), but a transparent and informative decision process which helps to uncover how people's intuitive decision procedures can be informed by a structured rational analytic process (Ananda and Herath, 2009). MCDA help decision-makers choose one course of action from among many complex alternatives, but it is a decision aid and in no way replaces the judgement of the decision-makers.

In public decision-making the inclusion of various stakeholders has become a natural part of the process. The process of public decision-making has to deal with two difficulties: the difficulty of instituting a structured public debate that can serve to legitimize the decisions made and the difficulty of providing appropriate instruments for evaluating investment projects that are transparent for all (Damart and Roy, 2009). This means that scientists cannot provide any useful input without interacting with the rest of society and the rest of the society cannot perform any sound decision making without interacting with the scientists (Munda, 2004). As stated earlier in most real-world policy situations there are many uncertainties and many stakeholders, each with their own preferences. To help deal with these two aspects of decision-making, Funtowicz and Ravetz (1993) has developed an epistemological framework called 'post-normal science". Post-normal science can be described by Figure 1. When both uncertainty and the stakes are small, we are in the realm of normal academic science; here it is safe to rely on standard routines and procedures. When either uncertainty or the stakes are in the medium range, then the application of routine techniques and standardized knowledge is no longer enough. In these cases, skill and judgement are required to adjust the "general knowledge" available to the "special situation" using professional consultancy. Lastly, when the stakes and systems uncertainty are very high, we are in the realm of post-normal science. Petersen et al. (2011) give a description of how the post-normal science paradigm has been put into practice in the Netherlands.

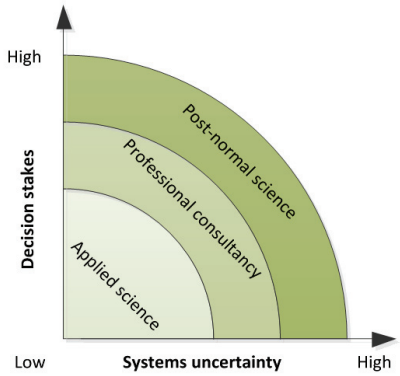


Figure 1 Graphical representation of post-normal science

1.1 Sensitivity analysis

Sensitivity analysis (SA) is aimed at determining how the output of a quantitative analysis depends on the inputs and thus assesses the robustness of the outcome. In MCDA, SA could be a means of explaining to the DM the implications and possible inconsistencies of his judgements. However, in decision problems with high public awareness, as with transport policy, SA could also be a means for taking into account the variety of different preferences.

In MCDA, the weights assigned to the decision criteria represent the importance of the criteria. When criteria cannot be expressed in quantitative terms (such as cost, weight, volume, etc.), it is difficult to accurately represent the importance of these criteria. In a situation like this, the decision-making process can be considerably improved by identifying the critical criteria (those with the potential to alter the ranking of the alternatives) and then re-evaluating more accurately the weights of these criteria. The intuitive belief is that the criterion with the highest weight is the most critical. This may not always be true and, in some particular instances, the criterion with the lowest weight may be the most critical one (Triantaphyllou and Sánchez, 1997).

Several different publications have suggested various methods for assessing the robustness of decisions by applying SA. Many have found that it is of great importance to examine the differences in preference weights (Salo and Hamalainen, 1992; Strager and Rosenberger, 2006; Triantaphyllou and Sánchez, 1997). They all acknowledge that variations in the criteria weights can have big influences on the results of an MCDA. Recently Jessop (2011) has suggested a method for modelling uncertainties using probabilistic weight; Rios Insua and French (1991) proposed an SA method to find competitors for a currently optimal alternative. Butler et al. (1997) and Butler and Olson (1999) proposed a method for doing simulation over the criteria weights while the rank order weights on the measures are maintained, but the weights are otherwise generated at random.

A lot is required of the techniques used to provide decision support in the public domain where there is public high awareness and active involvement of several different stakeholders groups in the decision process. The techniques have to be fast, transparent and clear, and at the same time be valid in their methodology. This particularly applies to decision support in transport planning, where there is an increasing tendency to involve stakeholders in the process of planning and decision support. This paper sets out to describe how to assess the robustness of decision support in a transparent, fast and operational way that can be used directly with the stakeholders during the decision process. So far this has had little attention in the literature, but it is vital if the decision-making is to be able to deal with the complexity of today's society.

The paper is structured as follows: first comes a description of methods for assessing the robustness of decisions supported by MCDA, thereafter follows case studies of specific decision problems supported by the use of MCDA and a measurement of the robustness of the decision aid; finally a conclusion and some perspectives are presented.

2 Assessing the robustness by sensitivity analysis

French, Maule and Papamichail (2009) argue that sensitivity analysis (SA) can support the decision process in eight ways: exploration of consequence models, exploration of the relationship between science and the consequence model, support for the elicitation of judgemental inputs, development of efficient computational algorithms,

design of surveys and experiments to gather relevant data, guidance for inferences, forecasts and decisions, communication and consensus building, and the development of understanding.

All of the above listed reasons for SA are relevant for public decision, but especially communication and consensus building are important. To get both public and political support for decisions, it is essential that the foundation for the decision-making process can be made transparent and easy to communicate. Furthermore, the existence of several different preferences calls for a methodology to examine how decisions are sensitive to differences in preferences.

There are several other reasons for conducting SA especially in an MCDA framework. First is because the MCDA process inherently contains several levels of uncertainty because of qualitative and subjective choices with regard to various parameters. In fact, MCDA has been criticised for not being an exact procedure, so it is very important to test how robust the results are. A second reason is that the SA of MCDA procedure enables the data and the decision-making problem to be explored at greater depths. This provides insight into the nature of the decision problem, unravels its complexities and may even provide recommendations for future analysis. Furthermore, an SA can be carried out to deal with the uncertainty in estimating some of the input figures. The decision-makers may not be able to arrive at a set of criteria weights and may provide a range of weightings or the decision may involve various groups of stakeholders with different preferences about the weights, which can then be analysed. In the same way, the impacts of various options under different criteria may fall within a statistically estimated range that can be incorporated into the analysis (Proctor and Qureshi, 2005).

The need for SA is clear because MCDA includes a comprehensive process involving a rich interaction between human judgement, data analysis and mathematical/computational processes. Errors and unintentional biases can enter at any of these stages, and it is the process as a whole that needs to be robust. Perhaps some of the key points at which such errors and biases may intrude would be the following (Stewart, 2008): external uncertainties, internal uncertainties, choice of model and identification of criteria and alternatives.

Figure 1 illustrates some of these key points at the different levels of uncertainty in an MCDA process. The two top levels of uncertainty are subjected to SA. Here internal uncertainties, choice of model and the identification of criteria and alternatives are examined. Uncertainties about the (external) environment are usually not analysed in SA.

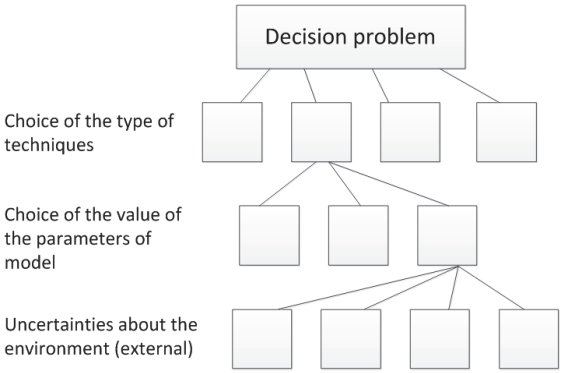


Figure 2 Different levels of uncertainty (adapted from Vincke, 2003)

The methodology proposed in this paper for examining the robustness of MCDA applied to the decision problem operates at different levels of the decision problem. The choice of MCDA model can affect the overall robustness of the analysis. This is with respect to the determination of weight and the method used for assessment of the alternatives' performance with regard to various attributes. This paper restricts itself to examining MCDA methodologies using pair-wise comparisons. This restriction is a result of experience in giving decision support. This paper examines how the choice of assessment methodology can affect the overall outcome of the analysis. This is done by applying the Analytic Hierarchy Process (AHP) and the Ratio Estimation in Magnitudes or deci-Bells to Rate Alternatives which are Non-DominaTed (REMBRANDT) methods (respectively additive and multiplicative methods using pair-wise comparisons) to the same decision problem and examining how robust the results are for changes in the attribute weights. If people are unsure of the weighting methods/principle, a robustness measure to changes in the weights is more relevant with regard to the uncertainty in the weighting. But even if everyone is content about methods, a robustness analysis could be useful to deal with changes in preferences or different ones. Elicitation of weights is usually a time-consuming process and is often controversial, and it is also difficult to arrive at exact weights and to determine consistent intervals between the weights. In such circumstances, ordinal ranking might be a reasonable compromise that makes use of consistent information and often provides output rank order for alternatives similar to the rank order based upon cardinal information (Shepetukha and Olson, 2001). This paper uses the ordinal ranking method Rank Order Distribution (ROD) for assessing criteria weights (Roberts and Goodwin, 2002).

The case studies in this paper are both decision problems in which the various stakeholders had agreed upon the criteria that should be included in the MCDA. In other words, they had all agreed upon the structure of the decision problem. So, the difference between the stakeholders can be described as differences in assessment of the various alternatives with respect to each criterion and differences in weights. The latter is described in the classical work of Edwards and Newman (1982), where they

found resolving conflicts about weights are easier to interpret and resolve than conflicts about the structure of the decision problem, e.g. what criteria to include or exclude.

Three different classes of simulation are used to assess the robustness of a decision:

a. Random ROD weights without importance order – simulation where criteria are assigned a random ROD weight. Two criteria can have equal weighting.
b. Random order weights – the importance order of the criteria is preserved, but the weights are randomly generated.
c. Weight interval – identifying critical criteria where a change in the weight will result in a different ranking of the alternatives.

The three classes can be interpreted as different levels of knowledge. With random weights, there is no knowledge about the relative importance of the weights an extreme case of SA. For random order weights, the rank order of the criteria is maintained, but the weights are otherwise generated at random. Lastly, for weight interval, it is assumed that the weights assessed by the stakeholders or decision-makers are valid, but with some uncertainty and the simulation can identify the critical criteria where a change of weight can result in a different ranking of the alternatives. This information can be useful, especially in the public decision domain where there can be a need to justify the decision for the public. An indication of how critical the criteria are if the weights are changed can defuse a potentially lengthy debate about the setting of the weights (Butler et al., 1997).

To make us of SA actively and directly in the decision process with the stakeholders, the SA must have specific properties. First the SA needs to be a transparent methodology if it is to gain acceptance and the attention of the stakeholders. Second, the SA has to be able to produce measures for robustness fairly fast, with respect to both the set-up and the computation time.

3 Results

Two case studies are used for illustration. Both the case studies concerns large strategic transport projects which are highly complex and also of great interest for the general public. The first case is about the alternatives for a new fixed link between the two cities of Helsingør (Elsinore) in Denmark and Helsingborg in Sweden (referred to as the HH-connection). Here the decision problem was to decide between three alternatives supporting different modes of transport. The second case is an examination of the alignment of a high-speed railway, Östlänken, in Sweden. The decision problem was to choose one of four different alternatives for the alignment on a small segment of the project. Common for both case studies is that a decision conference was convened to bring together some of the stakeholders to obtain input to the decision analysis.

3.1 Case study: HH connection

This case study concerns an additional fixed link between Denmark and Sweden. In the autumn of 2011, three tunnel alternatives were considered as the main alternatives for the HH-connection: Alt1 is a tunnel for rail, passenger trains only. Alt2 is a tunnel for rail, passenger trains only and a tunnel for vehicles. The last alternative, Alt3, is a tunnel for passenger trains, a tunnel for vehicles, and a tunnel for goods trains.

The first SA examined a random importance order for the criteria and the resulting total score for the alternatives. The results revealed that Alt1 would not be able to compete with the other two alternatives in any of the simulations. In Figure 3 the distributions of the simulated scores for Alt2 and Alt3 are shown.

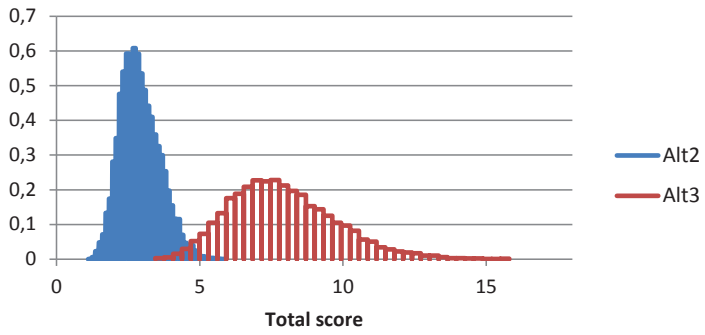


Figure 3 Simulation with ROD weights based on random importance order of the criteria

It can be seen from Figure 3 that Alt3 is most likely to be the preferred alternative and that Alt2 could compete with Alt3 only in a small subset of the simulation results. If the importance order of the criteria assessed by the stakeholders is preserved, the simulation results show the same tendency as with random weights, although the spread of the scores is smaller (Figure 4).

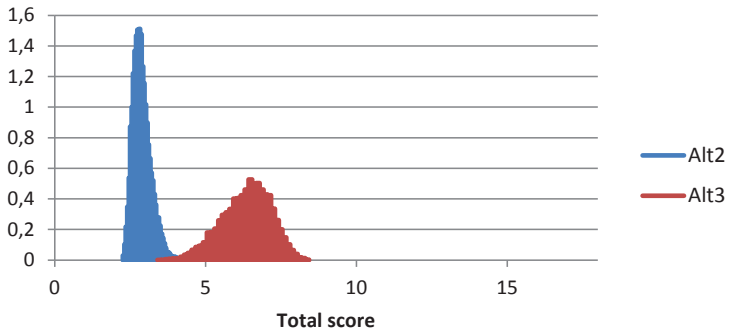


Figure 4 Simulation with random generated weights assigned to the criteria preserving the importance order of the criteria

To pinpoint criteria which are critical for the ranking of the alternatives, SA was performed on the weight stability. The results shown in Table 1 indicate that the 'Socio-economic robustness' and 'Impact on towns and land-use' criteria were critical for the ranking of the alternatives. The column 'Initial' shows the criteria weights used, based on the rank order of the criteria.

Table 1 Weight stability intervals of HH case

		REMBRANDT		AHP	
Criterion	Initial	Lower	Upper	Lower	Upper
Socio-economic robustness	0.24	-	0.47	-	0.48
Improvement for passenger cars and public transport	0.19	-	-	-	-
Impact on towns and land-use	0.04	-	0.49	-	0.15
Impact on regional economics	0.30	-	-	-	-
Impacts on flexibility in logistics	0.09	-	-	-	-
Contribution to the EU green corridors	0.14	-	-	-	-

As can be seen, both the REMBRANDT and AHP MCDA methodologies indicated the same critical criteria. However, AHP is much more sensitive to a change in the criterion 'impact on towns and land-use'.

3.2 Case study: Östlänken

A new high-speed railway line is planned in Sweden connecting Stockholm with Göteborg and Malmö. This case study concentrates on the link between Norrköping and Backeby. Four different corridors have been identified and the decision problem was to choose one of these 4 corridors.

The first SA examined a random importance order for the criteria and the resulting total score for the alternatives. The simulation results showed that the final ranking of

alternatives is sensitive to the criteria weights (Figure 5). Alt1 was the lowest ranked alternative is in most of the simulations and the highest ranked in none of the simulations. For Alt2 and Alt3 the simulation results are more ambiguous. Alt2 tended to be higher ranked than Alt3, but only slightly. Alt4 was ranked as third alternative in most of the simulations, but a more detailed analysis of the results revealed that with some special weight combinations this alternative was the highest ranked. However, this was just in a minor subset of the simulations and only if the REMBRANDT methodology was applied.

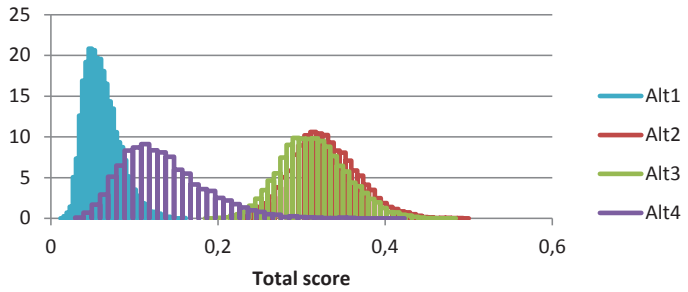


Figure 5 Simulation with ROD weights based on random importance order of the criteria

Figure 6 shows the simulation results when the importance order is preserved. From this figure, some of the same conclusions can be drawn as with Figure 5, but the spread of the scores is smaller for each of the alternatives. The overall conclusion from these first two SAs is that alternatives 2 and 3 are most likely to get the highest scores without being possible to say which of them is the most preferred.

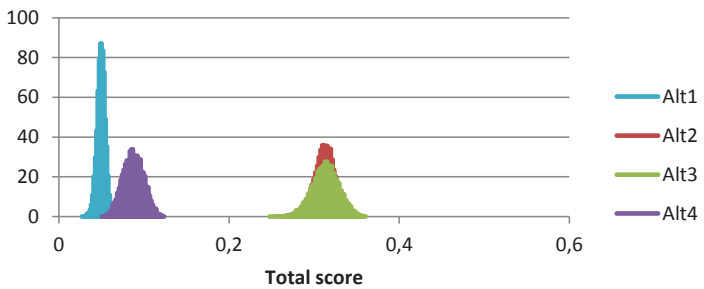


Figure 6 Simulation with random generated weights assigned to the criteria preserving the importance order of the criteria

For the Östlänken case the weight stability intervals for the criteria were also calculated and are shown in Table 2. The column 'Initial' shows the criteria weights used, based on the rank order of the criteria. From the table, it can be seen that a change in several of the criteria weights can affect the ranking of the alternatives. The

critical criteria weights for 'City and scenery impression' and 'Health' are relatively close to the initial weights – especially if the REMBRANDT methodology is applied. A slight lowering of the weight for the criterion 'Natural environment' results in a different ranking. One major reason for the large number of critical criteria in this case is that Alt2 and Alt3 were assessed to be very close to each other. So the critical criteria reflect the fact that changes in these weights result in a shift between which of these two alternatives is most preferred.

Table 2 Weight stability intervals

		REMBRANDT		AHP	
Criterion	Initial	Upper	Lower	Upper	Lower
Socio-economic	0.21	0.37	-	0.51	-
City and scenery impression	0.18	0.27	-	0.49	-
Cultural environment	0.16	0.53	0.06	0.33	-
Natural environment	0.13	-	0.09	-	0.05
Health	0.11	0.16	-	0.22	-
Natural resources	0.09	-	-	0.89	-
Risk and safety	0.06	0.22	-	0.21	-
Recreation and outdoor life	0.04	-	-	-	-
Building time	0.02	0.40	-	0.21	-

With regard to the two MCDA techniques, this case study implies that the REMBRANDT method is more sensitive to changes in the weights than the AHP.

4 Discussion and conclusion

This paper has illustrated an SA technique that combines simulation and rank order to provide the decision-makers with a clear picture of the robustness of the decision at stake. The technique presented is transparent and can be used actively during the decision-making process, as shown in the two case studies.

The two case studies showed that it is possible to produce a transparent and valid analysis of the robustness of an MCDA process to be used actively by the stakeholders or decision-makers. By working at different levels of the decision problem with a view to simplifying the SA, a clear and informative analysis can be conducted.

For the two case studies the random ROD weights approach gives information about how sensitive the decision analysis is for different ranking of the criteria by

importance. This is useful if not all of the stakeholders or decision-makers can agree on one particular ranking order of criteria. Furthermore, the random weight approach, where the ranking order is preserved, showed the effect of using other weights than the ROD weights applied. This can be used to meet criticism of the ROD methodology. Critical criteria can be found by examining the weight stability intervals of the criteria. For the HH case, two critical criteria were found, but a rather large change in the weights from those initially assessed has to be applied to get a different ranking order of the alternatives. For the Östlänken case, several critical criteria were found, but this was a result of the two closely competing alternatives. These results could be used to indicate the criteria which should bed examined further in the decision problem to provide a solid foundation for a decision.

Two different types of aggregation methods were applied to the two case studies: the AHP and the REMBRANDT methods. The results from the two case studies cannot give any clear conclusion about which method is more sensitive than the other with respect to the criteria weights. However, this implies that there could be a difference in the ranking of alternatives resulting from these two methods and researchers need to be aware of this.

It is also important to emphasise possible conflicts about the structure of the decision problem, and not only about the criteria weights. However, the purpose of this paper was to compare and produce some recommendations for SA to be used actively during a decision process in connection with the uncertainties and/or differences in preferences expressed by the criteria weights.

The tools for analysing the robustness of the MCDA must be able to communicate the complexity and associated uncertainties of the decisions and to allow for broad stakeholder participation while integrating all the different aspects of the situation involved. We should employ methods that are theoretically sound and, at the same time, practical and effective decision support tools.

Acknowledgement

The work with the case studies was co-funded by the Interreg IV project EcoMobility.

References

Ananda J and Herath G (2009) A critical review of multi-criteria decision making methods with special reference to forest management and planning. *Ecological Economics* 68(10): 2535-2548.

Butler J and Olson DL (1999) Comparison of centroid and simulation approaches for selection sensitivity analysis. *Journal of Multi-Criteria Decision Analysis* 8(3): 146-161.

Butler J, Jia J and Dyer J (1997) Simulation techniques for the sensitivity analysis of multi-criteria decision models. *European Journal of Operational Research* 103(3): 531-546.

Damart S and Roy B (2009) The uses of cost-benefit analysis in public transportation decision-making in France. *Transport Policy* 16(4): 200-212.

Edwards W and Newman R, J. (1982) *Multiattribute Evaluation*. Newbury Park: Sage Publications.

French S, Maule J and Papamichail N (2009) *Decision Behaviour, Analysis and Support*. Cambridge: Cambridge University Press.

Funtowicz SO and Ravetz JR (1993) Science for the post-normal age. *Futures* 25(7): 739-755.

Insua DR and French S (1991) A framework for sensitivity analysis in discrete multi-objective decision-making. *European Journal of Operational Research* 54(2): 176-190.

Jessop A (2011) Using imprecise estimates for weights*. Journal of the Operational Research Society* 62(6): 1048-1055.

Munda G (2004) Social multi-criteria evaluation: Methodological foundations and operational consequences. *European Journal of Operational Research* 158(3): 662-677.

Petersen AC, Cath A, Hage M, Kunseler E and van der Sluijs (2011) Post-Normal Science in Practice at the Netherlands Environmental Assessment Agency. *Science, Technology, & Human Values* 36(3): 362-388.

Proctor W and Qureshi M (2005) Multi-Criteria Evaluation Revisited. Palmerston North, New Zealand: Proceedings of the Australia New Zealand Society for Ecological Economics Conference. Massey University.

Roberts R and Goodwin P (2002) Weight approximations in multi-attribute decision models. *Journal of Multi Criteria Decision Analysis* 11: 291-304.

Salo AA and Hamalainen RP (1992) Preference Assessment by Imprecise Ratio Statements. *Operations Research* 40(6): 1053.

Shepetukha Y and Olson DL (2001) Comparative analysis of multiattribute techniques based on cardinal and ordinal inputs. *Mathematical and Computer Modelling* 34(1–2): 229-241.

Stewart TJ (2008) Robustness Analysis and MCDA. Porto Salvo: European Working Group "Multiple Criteria Decision Aiding".

Strager MP and Rosenberger RS (2006) Incorporating stakeholder preferences for land conservation: Weights and measures in spatial MCA. *Ecological Economics* 58(1): 79-92.

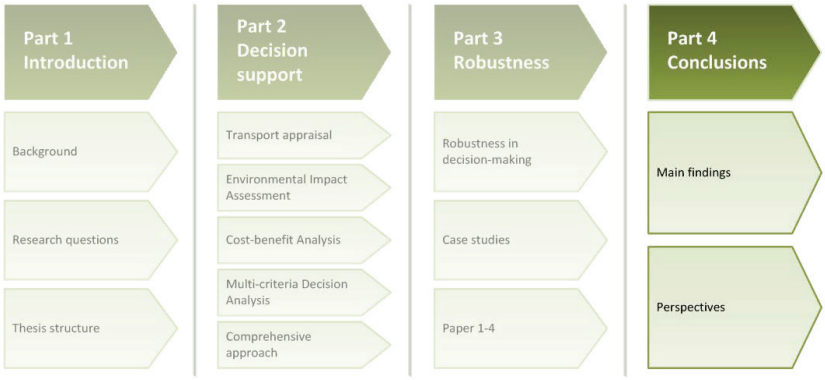
Triantaphyllou E and Sánchez A (1997) A sensitivity analysis approach for some deterministic multi-criteria decision-making methods. *Decision Sciences* 28(1): 151-185.

Vincke P (2003) About Robustness Analysis. Porto Salvo: European Working Group "Multiple Criteria Decision Aiding".

Walker WE (2000) Policy analysis: a systematic approach to supporting policymaking in the public sector. *Journal of Multi-Criteria Decision Analysis* 9(1): 11-27.

Part IV

Conclusions



9 Conclusion and perspectives

This thesis concerns assessing robustness in decision making. The focus is on appraising transport projects which conventionally has been carried out by using cost-benefit analysis. However, some project types such as larger transport infrastructure projects demand that a whole set of consequences are examined which cannot all be handled in a cost-benefit analysis. As stated a comprehensive transport assessment can be made by combining environmental impact assessment (EIA) with cost-benefit analysis (CBA) and multi-criteria decision analysis (MCDA). This PhD study has in this context examined three research questions as detailed below:

A. How to reconcile socio-economic and public acceptance to obtain decisions that are both rational and legitimate?

B. How to involve stakeholders in the appraisal?

C. How to apply robustness as a mean to improve the appraisal?

The examination of these research questions has been two-fold. For each question its relevance has first been argued and afterwards it was treated more specifically with regard to methodology and suggestions. As concerns research questions A and B these have been dealt with by setting focus on working with comprehensive assessment and by embedding this in a decision conference (DC). Compared to the conventional transport appraisal study based solely on cost-benefit analysis this wider type of analysis makes it possible to obtain decisions that can be considered rational, while at the same time being legitimate. This is supported by the various decision conferences conducted during the PhD study and presented in the papers. They all obtained very positive answers in evaluation questionnaires filled in afterwards by the participants. However, not to overrate this feedback it should be interpreted as part of a potential that can be a platform for other upcoming appraisal studies.

Research question C has been studied by setting up and modelling various robustness measures, which are described in detail in the four research papers included in the thesis. The robustness measures are developed to communicate the robustness of the appraisal result so that the decision-makers become aware of the particular robustness – or lack of same – of a decision they take on basis of a conducted comprehensive assessment in combination with a robustness examination. On the basis of the papers dealing with examination of alternatives, it can be concluded that robustness analysis and the robustness measures developed, can assist DMs in their deliberations about which alternative to recommend.

9.1 Main findings

The potential for qualifying transport appraisal as dealt with in this PhD study can be listed in the following summary of the main findings.

1. A methodology for widening the appraisal approach has been presented that makes it possible to take the environmental impacts (EIA) and the non-monetised impacts (MCDA) into account in addition to the conventional CBA approach. One particular aspect of this is that by selecting appropriate criteria it becomes possible to express sustainability in operational terms for actual decision support with MCDA facilitating the participation of stakeholders in the decision process. However, at the same time it is recognised that CBA is necessary for assessing the socio-economic aspects of the decision problem; therefore CBA should serve as an important part of conducting comprehensive appraisal. The first main finding is concerns appraisal that both provide rational and legitimate decision support.

    - To reconcile socio-economic analysis and public acceptance it can be recommended to widen the appraisal methodology.

2. One complication of using comprehensive appraisal consisting of different methodologies, which separately can be found to be complex methods, is that the appraisal process might become very complex and can be characterised as a black box. Therefore it is essential that communication is emphasised and the use of simple and transparent approaches are preferred. In general, this second main finding concerns a balancing of the used modelling approaches.

    - When widening the appraisal methodology, e.g. replacing CBA with EIA+CBA+MCDA+DC, there is a need to balance between the use of sophisticated modelling and emphasising transparency for all.

3. The value of active participation of stakeholders in the decision process has in the thesis been recognised in all the cases applied. The involvement of participants/stakeholders has been facilitated by the use of decision conference; the participatory processes led to generation of requisite decision support models through their concrete interaction in each case. Using a requisite model (seen as created together with the stakeholders) helped to structure the thinking of the individual participants and to generate shared understanding. The third main finding concerns the role of models and group process.

    - The conduct of a DC with the participants representing different stakeholders is seen as a way of combining technical, analytical solutions with social group processes aiming for engaging the different stakeholders with consideration of the variability of preferences.

4. If the DMs can test the outcome of the model in the view of different perspectives by the use of sensitivity analysis and robustness measures. This can both support the group process at a DC and/or examine how the difference

in preferences can influence the decision to be reached. In both cases the DMs can be informed about potential conflicts and take the necessary precautions in order to deal with these conflicts. The fourth and final main finding concerns the relevance of the developed robustness measures.

- The different SA and RA measures developed in the technical part of this PhD study are found to be relevant in supporting the DC group process, as robustness of a recommended solution is of major concern in the final steps of decision making.

9.2 Perspectives

The statements above also express perspectives of the study that can be elaborated further upon. Surely other sensitivity and robustness issues, than those dealt with in this PhD study, may be relevant to explore in the context of comprehensive appraisal. Other issues than the sensitivity with regard to criteria weight, that could be examined are: sensitivity of model parameters and the design of the processes used in the DC. The latter concerns, among others, the issue of how best to facilitate a decision process and how particular procedures could affect the outcome.

One general question could be stated on the basis of the findings of this PhD study: When appraisal becomes more open, participatory, and comprehensive, allowing for full acknowledgement of uncertainty and different preferences, can common recommendations still be formulated and decisions made. To be able to answer such a question, empirical work in appraisal of transport projects involving documentary analysis and interviews with those involved in the process is needed.

The findings of this PhD study need to be refined based on more research and empirical studies, but they are already practicable in transport appraisal and potential to make use of also outside the transport sector.

An overall issue is to further test and refine the robustness measures in upcoming appraisal studies where comprehensive assessment is made use of in combination with the conduct of decision conferences. In this context the findings of this PhD study is seen as a suitable platform for such further research in this field.

LITERATURE

Ackerman F and Heinzerling L (2004) *Priceless / on Knowing the Price of Everything and the Value of Nothing.* New York, NY: The New Press.

André P, Enserink B, Connor D and Croal P (2006) *Public Participation International Best Practice Principles. Special Publication Series no. 4.* Fargo: International Association for Impact Assessment.

Archibugi F (1989) Comprehensive social assessment: An essential instrument for environmental policy-making. In: Archibugi F and Nijkamp P (eds) *Economy and Ecology: Towards Sustainable Development.* Dordrecht: Kluwer Academic Publishers.

Baarsma BE and Lambooy JG (2005) Valuation of externalities through neo-classical methods by including institutional variables. *Transportation Research Part D: Transport and Environment* 10(6): 459-475.

Bana E Costa CA, Stewart TJ and Vansnick J (1997) Multicriteria decision analysis: Some thoughts based on the tutorial and discussion sessions of the ESIGMA meetings. *European Journal of Operational Research* 99(1): 28-37.

Banister D (2008) The sustainable mobility paradigm. *Transport Policy* 15(2): 73-80.

Banister D and Berechman J (2000) *Transport Investment and Economic Development.* London: UCL Press.

Banville C, Landry M, Martel J and Boulaire C (1998) A stakeholder approach to MCDA. *Systems Research and Behavioral Science* 15(1): 15-32.

Barfod MB (2012) Optimising Transport Decision Making using Customised Decision Models and Decision Conferences. PhD thesis, Technical University of Denmark - Department of Transport: Kgs. Lyngby.

Beinat E (2001) Multi-criteria analysis for environmental management. *Journal of Multi-Criteria Decision Analysis* 10(2): 51-52.

Belton V and Stewart JT (2002) *Multiple Criteria Decision Analysis: An Integrated Approach.* London: Kluwer Academic Publishers.

Beukers E, Bertolini L and Te Brömmelstroet M (2012) Why Cost Benefit Analysis is perceived as a problematic tool for assessment of transport plans: A process perspective. *Transportation Research Part A* 46(1): 68-78.

Booth C and Richardson T (2001) Placing the public in integrated transport planning. *Transport Policy* 8(2): 141-149.

Brent RJ (1979) Imputing weights behind past railway closure decisions within a cost-benefit framework. *Applied Economics* 11(2): 157-170.

Bristow AL and Nellthorp J (2000) Transport project appraisal in the European Union. *Transport Policy* 7(1): 51-60.

Brown R (2004) Consideration of the origin of Herbert Simon's theory of "satisficing" (1933-1947). *Management Decision* 42(10): 1240-1256.

Browne D and Ryan L (2011) Comparative analysis of evaluation techniques for transport policies. *Environmental Impact Assessment Review* 31(3): 226-233.

Butler J, Jia J and Dyer J (1997) Simulation techniques for the sensitivity analysis of multi-criteria decision models. *European Journal of Operational Research* 103(3): 531-546.

Colorni A, Laniado E and Muratori S (1999) Decision support systems for environmental impact assessment of transport infrastructures. *Transportation Research Part D: Transport and Environment* 4(1): 1-11.

COWI (2009) *Study Concerning the Report on the Application and Effectiveness of the EIA Directive.* : EU DG ENV.

Damart S and Roy B (2009) The uses of cost-benefit analysis in public transportation decision-making in France. *Transport Policy* 16(4): 200-212.

Danzanvilliers P, Delmarcelle A, Doyle N, van Geldermalsen T, Haaland K, Jäderholm B, et al. (1999) *Economic Evaluation Methods for Road Projects in PIARC Member Countries.* Paris: PIARC-World Road Association.

De Brucker K and Verbeke A (2007) The institutional theory approach to transport policy and evaluation. The collective benefits of a stakeholder's approach: towards an electic multi-criteria analysis. In: Haezendonck E (ed) *Transport Project Evaluation - Extending the Social Cost-Benefit Approach.* Cheltenham: Edward Elgar, 55-94.

De Brucker K, Macharis C and Verbeke A (2011) Multi-criteria analysis in transport project evaluation: an institutional approach. *European Transport* 47: 3-24.

de Montis A, Toro PD, Droste-Franke B, Omann I and Stagle S (2006) Assessing the quality of different MCDA methods. In: Spash CL, Stagle S and Getzner M (eds) *Alternatives for Environmental Valuation.* London: Routledge, 1-133.

Delgado MG and Sendra JB (2004) Sensitivity analysis in multicriteria spatial decision-making: A review. *Human and Ecological Risk Assessment* 10(6): 1173-1187.

Department for Communities and Local Government (2009) *Multi-Criteria Analysis: A Manual.* London: Department for Communities and Local Government.

DfT (2011) *Cost Benefit Analysis.* London: Department for Transport, UK.

Dias LC and Clímaco JN (1999) On computing ELECTRE's credibility indices under partial information. *Journal of Multi-Criteria Decision Analysis* 8(2): 74-92.

Edwards W and Newman R, J. (1982) *Multiattribute Evaluation.* Newbury Park: Sage Publications.

Edwards W and Barron FH (1994) SMARTS and SMARTER: Improved Simple Methods for Multiattribute Utility Measurement. *Organizational Behavior and Human Decision Processes* 60(3): 306-325.

Eliasson J and Lundberg M (2012) Do Cost-Benefit Analyses Influence Transport Investment Decisions? Experiences from the Swedish Transport Investment Plan 2010-21. *Transport Reviews* 32(1): 29-48.

Fischer TB (2003) Strategic environmental assessment in post-modern times. *Environmental Impact Assessment Review* 23(2): 155-170.

Fischhoff B (2010) Judgment and decision making. *WILEY INTERDISCIPLINARY REVIEWS-COGNITIVE SCIENCE* 1(5): 724-735.

Franco LA and Montibeller G (2010) Facilitated modelling in operational research. *European Journal of Operational Research* 205(3): 489-500.

Frank RH (2008) Why Is Cost-Benefit Analysis So Controversial? In: Hausman DM (ed) *The Philosophy of Economics.* Cambridge: Cambridge University Press, 251-269.

French S, Maule J and Papamichail N (2009) *Decision Behaviour, Analysis and Support.* Cambridge: Cambridge University Press.

Frey HC, Patil SR, Patil SR and Frey HC (2002) Identification and review of sensitivity analysis methods. *Risk Analysis* 22(3): 553-578.

Fridstrom L and Elvik R (1997) The barely revealed preference behind road investment priorities. *Public Choice* 92(1-2): 145-168.

Fujiwara D and Campbell R (2011) *Valuation Techniques for Social Cost-Benefit Analysis: Stated Preference, Revealed Preference and Subjective Well-being Approaches.* London: Department for Work and Pensions, HM Treasury.

Funtowicz S and Ravetz J (2008) *Post-Normal Science.* Available at: http://www.eoearth.org/article/Post-Normal_Science.

Funtowicz SO and Ravetz JR (1993) Science for the post-normal age. *Futures* 25(7): 739-755.

Gamper CD, Thoeni M and Weck-Hannemann H (2006) A conceptual approach to the use of Cost Benefit and Multi Criteria Analysis in natural hazard management. *NATURAL HAZARDS AND EARTH SYSTEM SCIENCES* 6(2): 293-302.

Gervásio H and Simões dS (2012) A probabilistic decision-making approach for the sustainable assessment of infrastructures. *Expert Systems with Applications* 39(8): 7121-7131.

Geurs KT, Boon W and Van Wee B (2009) Social Impacts of Transport: Literature Review and the State of the Practice of Transport Appraisal in the Netherlands and the United Kingdom. *Transport Reviews* 29(1): 69-90.

Gissel S (1999) Decision Aid Methods in Rail Infrastructure Planning. Department of Planning, Technical University of Denmark.

Goodwin P and Persson S (2001) *Assessing the Benefits of Transport.* Paris: OECD - ECMT.

Goodwin P and Wright G (2009) *Decision Analysis for Management Judgment.* Chichester: John Wiley & Sons.

Grant-Muller S, MacKie P and Nellthorp J (2001) Economic appraisal of European transport projects: the state-of-the-art revisited. *Transport Reviews* 21(2): 237-261.

Haezendonck E (2007) *Transport Project Evaluation - Extending the Social Cost-Benefit Approach.* Cheltenham: Edward Elgar.

Hage M, Leroy P and Petersen AC (2010) Stakeholder participation in environmental knowledge production. *Futures* 42(3): 254-264.

Healey P (1990) Places, people and policies. *Town and Country Planning* 9(10): 9-10.

Hickman R, Saxena S, Banister D and Ashiru O (2012) Examining transport futures with scenario analysis and MCA. *Transportation Research Part A* 46(3): 560-575.

Holsapple CW (2008) Decisions and Knowledge. In: Burstein F and Holsapple CW (eds) *Handbook on Decision Support Systems 1.* Berlin Heidelberg: Springer-Verlag, 21-52.

Hoppe R (2002) Cultures of Public Policy Problems. *Journal of Comparative Policy Analysis: Research and Practice* 4: 305-326.

Howarth A, Pearce DW, Ozdemiroghu E, Wieringa K, Streeferk CM and de Hollander AEM (2001) *Valuing the Benefits of Environmental Policy: The Netherlands.* Bilthoven: RIVM - National Institute of public health and the environment.

Hyde KM (2006) Uncertainty Analysis Methods for Multi-Criteria Decision Analysis. Doctor of Philosophy, The University of Adelaide: Adelaide.

Janssen R (2001) On the use of multi-criteria analysis in environmental impact assessment in The Netherlands. *Journal of Multi-Criteria Decision Analysis* 10(2): 101-109.

Jay S, Jones C, Slinn P and Wood C (2007) Environmental impact assessment: Retrospect and prospect. *Environmental Impact Assessment Review* 27(4): 287-300.

Jensen AV, Ambrasaite I, Salling KB, Barfod MB and Leleur S (2012) The EcoMobility Modelling Framework for Sustainable Transport Planning. In: Carlsson C, Emtairah T,

Gammelgaard B, Jensen AV and Thidell Å (eds) *Rethinking Transport in the Øresund Region: Policies, Strategies and Behaviours.* Lund: Lund University, 149-164.

Jeppesen SL (2009) Sustainable Transport Planning - A Multi-Methodology Approach to Decision Making. PhD thesis, Technical University of Denmark - Department of Transport: Kgs. Lyngby.

Jessop A (2002) Prioritisation of an IT budget within a local authority. *Journal of the Operational Research Society* 53(1): 36-46.

Joumard R and Gudmundsson H (2010) *Indicators of Environmental Sustainability in Transport.* : INRETS.

Kahneman D and Tversky A (1979) Prospect theory: an analysis of decision under risk. *Econometrica* 47(2): 263-292.

Kao C (2010) Weight determination for consistently ranking alternatives in multiple criteria decision analysis. *Applied Mathematical Modelling* 34(7): 1779-1787.

Keeney RL and Von Winterfeldt D (2007) Practical Value Models. In: Edwards W, Miles RF and Von Winterfeldt D (eds) *Advances in Decision Analysis.* New York: Cambridge University Press, 232-252.

Larsen LA and Skougaard BZ (2010) Appraisal of Alternatives Concerning a Fixed Link between Elsinore and Helsingborg. M.Sc., Technical University of Denmark: Kgs. Lyngby.

Lawrence DP (2007) Impact significance determination - Pushing the boundaries. *Environmental Impact Assessment Review* 27(8): 770-788.

Leleur S (2000) *Road Infrastructure Planning / A Decision-Oriented Approach.* Lyngby: Polyteknisk Forlag.

Leleur S (2012) *Complex Strategic Choices : Applying Systemic Planning for Strategic Decision Making.* London New York: Springer.

Lootsma FA (1999) *Multi-Criteria Decision Analysis Via Ration and Difference Judgement.* Dordrecht: Kluwer.

Lootsma FA (1992) *The REMBRANDT System for Multi-Criteria Decision Analysis Via Pair Wise Comparisons Or Direct Rating.* Delft: Faculty of Technical Mathematics and Informatics, Delft University of Technology.

Lutz E and Munasinghe M (1994) Integration of environmental concerns into economic analyses of projects and policies in an operational context. *Ecological Economics* 10(1): 37-46.

Macharis C (2007) Multi-Criteria Analysis as a Tool to Include Stakeholders in Project Evaluations: the MAMCA method. In: Haezendonck E (ed) *Transport Project Evaluation - Extending the Social Cost-Benefit Approach.* Cheltenham: Edward Elgar, 115-131.

Macharis C, De Witte A and Ampe J (2009) The Multi-Actor, Multi-Criteria Analysis Methodology (MAMCA) for the Evaluation of Transport Projects: Theory and Practice. *Journal of Advanced Transportation* 43(2): 183-202.

Mackie PJ, Jara-Díaz S and Fowkes AS (2001) The value of travel time savings in evaluation. *Transportation Research Part E: Logistics and Transportation Review* 37(2–3): 91-106.

Mackie P (2010) *Cost-Benefit Analysis in Transport: A UK Perspective.* : International Transport Discussion Paper No. 2010-16, OECD.

Mackie P and Preston J (1998) Twenty-one sources of error and bias in transport project appraisal. *Transport Policy* 5(1): 1-7.

Mareschal B (1988) Weight stability intervals in multicriteria decision aid. *European Journal of Operational Research* 33(1): 54-64.

Martinsen JA, Odeck J and Kjerkreit A (2010) *Why Benefit-Cost Analyses Matter Less and how it can be Improved for Decision Making in the Transport Sector – Experiences from the Norwegian National Transport Plan 2010 -2019.* Glasgow: European Transport Conference.

Mateos A, Jiménez A and Ríos-Insua S (2006) Monte Carlo simulation techniques for group decision making with incomplete information. *European Journal of Operational Research* 174(3): 1842-1864.

McFadden D (1976) The revealed preference of a government bureaucracy: empirical evidence. *The Bell Journal of Economics* 7(1): 55-72.

McFadden D (1975) The revealed preference of a government bureaucracy: theory. *The Bell Journal of Economics* 6(2): 401-716.

Morris P and Therivel R (2001) *Methods of Environmental Impact Assessment.* London: Spon press.

Morrissey J, Iyer-Raniga U, McLaughlin P and Mills A (2012) A Strategic Project Appraisal framework for ecologically sustainable urban infrastructure. *Environmental Impact Assessment Review* 33(1): 55-65.

Munda G (2008) *Social Multi-Criteria Evaluation for a Sustainable Economy.* Berlin: Springer.

Musso E, Sanguinetti S and Sillig C (2007) Socio-economic impact of transport policies: an institutional approach. In: Haezendonck E (ed) *Transport Project Evaluation - Extending the Social Cost-Benefit Approach.* Cheltenham: Edward Elgar, 95-114.

Navrud S, Trædal Y, Hunt A, Longo A, Greßmann A, Leon C, et al. (2006) *HEATCO Deliverable 4: Economic Values for Key Impacts Valued in the Stated Preference Surveys.* : European Commission EC-DG TREN.

Nellthorp J and Mackie PJ (2000) The UK Roads Review - a hedonic model of decision making. *Transport Policy* 7(2): 127-138.

Nilsson J (1991) Investment Decisions in a Public Bureaucracy: A Case Study of Swedish Road Planning Practices. *Journal of Transport Economics and Policy* 25(2): 163-175.

Noble BF (2011) Environmental Impact Assessment. *ELS*.

Nyborg K (1998) Some Norwegian politicans' use of cost-benefit analysis. *Public Choice* 95(3-4).

Odeck J (1996) Ranking of regional road investment in Norway. *Transportation* 23(2): 123-140.

Odgaard T, Kelly C and Laird J (2005) *HEATCO Deliverable 1: Current Practice in Project Appraisal in Europe Analysis of Country Reports.* : European Commission EC-DG TREN.

OECD (2011) *Improving the Practice of Transport Project Appraisal.* Paris: OECD Publishing.

OECD (2009) *Regulatory Impact Analysis: A Tool for Policy Coherence.* : OECD publishing.

Olson DL, Fliedner G and Currie K (1995) Comparison of the REMBRANDT system with analytic hierarchy process. *European Journal of Operational Research* 82: 522-539.

Øresundsbro Konsortiet (2010) *10 Years: The Øresund Bridge and its Region.* Copenhagen: Øresundsbro Konsortiet.

Østergaard N and Andersen U (2012) *Millioner Kastet i Dødfødte Bro- Og Vejprojekter.* Available at: http://ing.dk/artikel/125937-millioner-kastet-i-doedfoedte-bro-og-vejprojekter.

Owens S, Rayner T and Bina O (2004) New agendas for appraisal: reflections on theory, practice, and research. *Environment & Planning A* 36(11).

Pearce D, Atkinson G and Mourato S (2006) *Cost-Benefit Analysis and the Environment - Recent Developments.* Paris: OECD Publishing.

Pearce DW and Nash CA (1981) *The Social Appraisal of Projects : A Text in Costbenefit Analysis.* London: Macmillan.

Petersen AC, Cath A, Hage M, Kunseler E and van dS (2011) Post-Normal Science in Practice at the Netherlands Environmental Assessment Agency. *Science, Technology, & Human Values* 36(3): 362-388.

Phillips LD (2007) Decision conferencing. In: Edwards W, Miles RF and von Winterfeldt D (eds) *Advances in Decision Analysis - from Foundations to Applications.* : Cambridge University Press, 375-399.

Phillips LD (1984) A theory of requisite decision models. *Acta Psychologica* 56(1–3): 29-48.

Phillips LD and Bana eC (2007) Transparent prioritisation, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing. *Annals of Operations Research* 154(1): 51-68.

Phillips LD and Phillips MC (1993) Faciliated Work Groups: Theory and Practice. *The Journal of the Operational Research Society* 44(6): 533-549.

Porter T (1995) *Trust in Numbers: The Pursuite of Objectivity in Science and Public Life.* Princeton: Princeton University Press.

Proctor W and Qureshi M (2005) *Multi-Criteria Evaluation Revisited.* Palmerston North, New Zealand: Proceedings of the Australia New Zealand Society for Ecological Economics Conference. Massey University.

Qureshi ME, Harrison SR and Wegener MK (1999) Validation of multicriteria analysis models. *Agricultural Systems* 62(2): 105-116.

Renn O (2006) Participatory processes for designing environmental policies. *Land use Policy* 23(1): 34-43.

Roberts R and Goodwin P (2002) Weight approximations in multi-attribute decision models. *Journal of Multi Criteria Decision Analysis* 11: 291-304.

Saaty (1977) Scenarios and priorities in transport planning: application to the Sudan. *Transportation Research* 11(5): 343-350.

Saaty T (2008) Decision making with the analytic hierarchy process. *International Journal of Services Sciences* 1(1): 83-98.

Sager T and Ravlum I (2005) The Political Relevance of Planners' Analysis: The Case of a Parliamentary Standing Committee. *Planning Theory* 4(1): 33-65.

Saitua R (2007) Some considerations on social cost-benefit analysis as tool for decision-making. In: Haezendonck E (ed) *Transport Project Evaluation - Extending the Social Cost-Benefit Approach.* Cheltenham: Edward Elgar, 23-34.

Salling KB (2012) *UNITE.* Available at: http://www.transport.dtu.dk/subsites/UNITE/english.aspx.

Salling KB (2008) Assessment of Transport Projects - Risk Analysis and Decision Support. PhD thesis, Department of Transport: Kgs. Lyngby.

Salling KB and Leleur S (2011a) Transport appraisal and Monte Carlo simulation by use of the CBA-DK model. *Transport Policy* 18(1): 236-245.

Salling KB and Leleur S (2011b) Modelling of transport project uncertainties: Feasibility risk assessment and scenario analysis. *European Journal of Transport and Infrastructure Research* 12(1): 21-38.

Senécal P, Goldsmith B, Conover S, Sadler B and Brown K (1999) *Principles of Environmental Impact Assessment - Best Practice* Fargo: International association for impact assessment.

Shepetukha Y and Olson DL (2001) Comparative analysis of multiattribute techniques based on cardinal and ordinal inputs. *Mathematical and Computer Modelling* 34(1–2): 229-241.

Sijtsma FJ (2006) Project Evaluation, Sustainability and Accountability - Combining Cost-Benefit Analysis (CBA) and Multi-Criteria Analysis (MCA). PhD thesis, Stichting Ruimtelijke Economie Gronningen: Gronningen.

Spagnoli F (2003) *Homo Democraticus: On the Universal Desirability and the Not so Universal Possibility of Demoncracy and Human Rights.* Buckinghamshire: Cambridge Scholar Press.

Spash CL, Stagle S and Getzner M (2006) Exploring alternatives for environmental valuation. In: Spash CL, Stagle S and Getzner M (eds) *Alternatives for Environmental Valuation.* London: Routledge, 1-20.

Staehr K (2006) *Risk and Uncertainty in Cost Benefit Analysis.* Copenhagen: Environmental Assessment Institute.

Stewart TJ (2008) *Robustness Analysis and MCDA.* Porto Salvo: European Working Group "Multiple Criteria Decision Aiding".

Stewart TJ (1992) A critical survey on the status of multiple criteria decision making theory and practice. *Omega* 20(5-6): 569-586.

Stirling A (2008) "Opening Up" and "Closing Down": Power, Participation, and Pluralism in the Social Appraisal of Technology. *Science, Technology & Human Values* 33(2): 262-294.

Thomopoulos N, Grant-Muller S and Tight MR (2009) Incorporating equity considerations in transport infrastructure evaluation: Current practice and a proposed methodology. *Evaluation and Program Planning* 32(4): 351-359.

Triantaphyllou and Sanchez (1997) A sensitivity analysis approach for some deterministic multi-criteria decision-making methods. *Decision Sciences* 28(1): 151-194.

Tsamboulas D and Mikroudis G (2000) *EFECT - Evaluation Framework of Environmental Impacts and Costs of Transport Initiatives.*

Vaidya OS and Kumar S (2006) Analytic hierarchy process: An overview of applications. *European Journal of Operational Research* 169: 1-29.

van den Belt M (2004) *Mediated Modeling - a System Dynamics Approach to Environmental Consens Building.* Washington: Island Press.

van den Honert R (1998) Stochastic pairwise comparative judgements and direct ratings of alternatives in the REMBRANDT system. *Journal of Multi-Criteria Decision Analysis* 7(2): 87-97.

von Winterfeldt D and Edwards W (1993) *Decision Analysis and Behavioral Research.* Cambridge: Cambridge University Press.

Walker WE (2000) Policy analysis: a systematic approach to supporting policymaking in the public sector. *Journal of Multi-Criteria Decision Analysis* 9(1): 11-27.

Willis KG (2005) Cost-Benefit Analysis. In: Button KJ and Hensher DA (eds) *Handbook of Transport Strategy, Policy and Institutions.* Amsterdam: Elsevier, 491-506.

Woltjer J (2009) Concepts of Participatory Decision-Making in Dutch Infrastructure Planning. In: Coenen, Frans H. J. M. (ed) *Public Participation and Better Environmental Decisions - the Promise and Limits of Participatory Processes for the Quality of Environmentally Related Decision-Making.* : Springer Science + Business Media B.V., 153-163.

Yoon KP and Hwang C (1995) *Multiple Attribute Decision Making - an Introduction.* Thousands Oaks, CA: SAGE.

Zurita L (2006) Consensus conference method in environmental issues: relevance and strengths. *Land use Policy* 23(1): 18-25.

DTU Transport performs research and provides education on traffic and transport planning. It advises the Danish Ministry of Transport on infrastructure, economic appraisals, transport policy and road safety and collects data on the transport habits of the population. DTU Transport collaborates with companies on such topics as logistics, public transport and intelligent transport systems.

DTU Transport
Department of Transport
Technical University of Denmark

Bygningstorvet 116B
Bygning 115
DK-2800 Kgs. Lyngby
Tel. +45 45 25 65 00

www.transport.dtu.dk