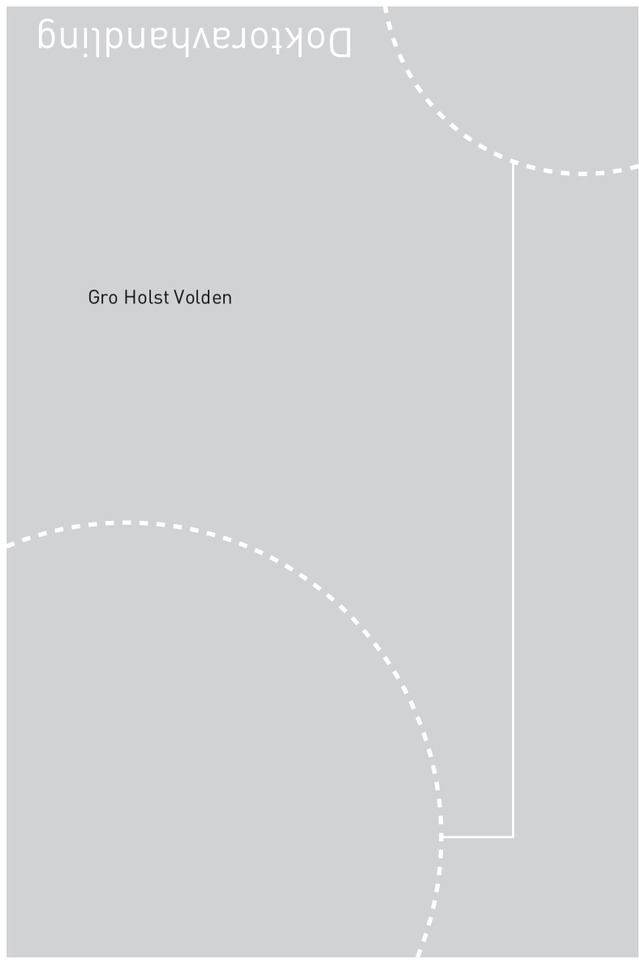
Doktoravhandlinger ved NTNU, 2019:236

Gro Holst Volden

Up-Front Governance of Major Public Investment Projects

NTNU
Kunnskap for en bedre verden



ISBN 978-82-326-4068-3 (trykt utg.)
ISBN 978-82-326-4069-0 (elektr. utg.)
ISSN 1503-8181

NTNU
Norges teknisk-naturvitenskapelige universitet
Avhandling for graden
philosophiae doctor
Fakultet for ingeniørvitenskap
Institutt for bygg- og miljøteknikk

Doktoravhandlinger ved NTNU, 2019:236

NTNU
Kunnskap for en bedre verden

Gro Holst Volden

Up-Front Governance of Major Public Investment Projects

Avhandling for graden philosophiae doctor

Trondheim, November 2019

Norges teknisk-naturvitenskapelige universitet
Fakultet for ingeniørvitenskap
Institutt for bygg- og miljøteknikk

NTNU
Kunnskap for en bedre verden

NTNU
Norges teknisk-naturvitenskapelige universitet

Avhandling for graden philosophiae doctor

Fakultet for ingeniørvitenskap
Institutt for bygg- og miljøteknikk

ISBN 978-82-326-4068-3 (trykt utg.)
ISBN 978-82-326-4069-0 (elektr. utg.)
ISSN 1503-8181

Doktoravhandlinger ved NTNU, 2019:236

Trykket av NTNU Grafisk senter

Up-Front Governance of Major Public Investment Projects

Foreword

This thesis marks the end of a journey that has been challenging and interesting, but also a great learning process. Although I had worked as a researcher for several years, the PhD process has been rewarding in taking my research skills to the next level.

Research conducted by the Concept Research Programme, which is funded by the Norwegian Ministry of Finance, has provided me with a unique opportunity to raise and address the research questions explored in this thesis. Data from the Norwegian quality assurance scheme (QA scheme) have been extensively used. I hope that the present work can be useful in the continued development and improvement of the QA scheme.

The thesis is written as a collection of published articles. The synthesizing part is based on these publications alone and does not introduce new elements. This is in line with established practice at NTNU.

I would like to thank all of my colleagues in the Concept Research Programme. Their direct inputs into my work, encouragement and support in the process, as well as general discussions on project governance, quality assurance and related topics, have been invaluable and a great inspiration on a daily basis. Special thanks are owed to Professor Knut Samset, the founding father of the program, whose thoughts and ideas have inspired me greatly. Professor Samset has co-authored four of the papers and provided valuable inputs into the others. However, the thesis has been an independent work and is my responsibility alone.

Sincere thanks are also owed to my supervisor, Professor Ola Lædre, and co-supervisors, Professor Tore Haavaldsen and Professor James Odeck, for constructive discussions and advice along the way. Tore Haavaldsen talked me into doing a PhD, and today I am glad that he did that.

Finally, I thank my supportive family and friends outside the university, especially my husband Trond, and my sons Ole and Lasse, for bearing with me through busy times.

Trondheim, October 29, 2019

Gro Holst Volden

Contents

List of Tables

List of Figures

Table 1 Papers included in the thesisshows the two book chapters and seven journal papers that have contributed to the conclusion of this thesis. They have all been published following peer review.

Table 1 Papers included in the thesis

No.	Paper Title
1	Samset, K. and Volden, G.H. (2015). Front-end definition of projects: Ten paradoxes and some reflections regarding project management and project governance, *International Journal of Project Management 34*, 297–313
2	Volden, G.H. (2018a). Public funding, perverse incentives, and counterproductive outcomes, *International Journal of Managing Projects in Business 12(2)*, 466–486
3	Samset, K. and Volden, G.H. (2012). The proposal, book chapter in Williams, T. and Samset, K. (eds.), *Project Governance. Getting investments right*, Palgrave Macmillan, Basingstoke, 46-81
4	Volden, G.H. and Samset, K. (2017a). Quality assurance in megaproject management, the Norwegian way, book chapter in Flyvbjerg, B. (ed.) *The Oxford Handbook of Megaproject Management*, Oxford University Press, Oxford, 406-428
5	Volden, G.H. and Samset, K. (2017b). Governance frameworks for major public investment projects: Principles and practices in six countries, *Project Management Journal 48(3)*, 90–108
6	Volden, G.H. and Andersen, B.S. (2018). The hierarchy of public project governance schemes, *International Journal of Managing Projects in Business 11(1)*, 174–197
7	Volden, G.H. (2019). Assessing public projects' value for money: An empirical study of the usefulness of cost–benefit analyses in decision-making, *International Journal of Project Management 37(4)*, 549–564
8	Haavaldsen, T., Lædre, O., Volden, G.H. and Lohne, J. (2014). On the concept of sustainability, *International Journal of Sustainable Engineering 7(1)*, 2–12
9	Volden, G.H. (2018b). Public project success as seen in a broad perspective: Lessons from a meta-evaluation of 20 infrastructure projects in Norway, *Evaluation and Program Planning 69*, 109–117

My contribution to papers for which I was not the sole author:

- Paper 1: Contributed to research design, synthesizing of data and analyses, writing drafts (four of the paradoxes), and editing and finalizing the paper. In total, slightly less than 50% of the work.
- Paper 3: Contributed to design, writing drafts (6 out of 15 subsections), and editing and finalizing the paper. In total, slightly less than 50% of the work.
- Paper 4: Contributed to research design, data collection and analyses, writing drafts, and editing and finalizing the paper. In total, slightly more than 50% of the work.
- Paper 5: Contributed to research design, data collection and analyses, writing drafts, and editing and finalizing the paper. In total, about 50% of the work.
- Paper 6: Contributed to research design, data collection and analyses, writing drafts, and editing and finalizing the paper. In total, more than 50% of the work.
- Paper 8: Contributed to research design, analyses, and editing and finalizing the paper. In total, about 25% of the work.

During the PhD process, I have also contributed to other publications, which are relevant to the thesis to a greater or lesser extent. These are not included in the thesis itself, but have been produced as part of the research process that led to the thesis. Thematically, they overlap with the main papers, but some answer slightly different or narrower research questions. The supporting papers are listed in Table 2 and I refer to them where relevant.

Table 2 Supporting papers

No.	Paper
10	Samset, K. and Volden, G.H. (2013). Major projects up front: Analysis and decision – rationality and chance, paper presented at IRNOP Project Research Conference, Oslo, June 16–18, 2013
11	Samset, K. and Volden, G.H. (2014). Front end governance of major public projects – lessons with a Norwegian Quality Assurance scheme, *International Journal of Architecture, Engineering and Construction, 3(2)*, 1–10
12	Odeck, J., Welde, M. and Volden, G.H. (2015). The impact of external quality assurance of costs estimates on cost overruns: Empirical evidence from the Norwegian road sector, *European Journal of Transport and Infrastructure Research 15(3)*, 286–303
13	Klakegg, O.J. and Volden, G.H. (2017). Governance in public projects: The Norwegian case, book chapter, Müller, R. (Ed.) *Governance and Governmentality for Projects*, Routledge, New York, 129-155

14	Næss, P., Volden, G.H., Odeck, J. and Richardson, T. (2017) *Neglected and Underestimated Impacts of Transport Investments*, Concept Report No. 54, NTNU, Trondheim
15	Welde, M. and Volden, G.H. (2018). Measuring efficiency and effectiveness through ex-post evaluation: Case studies of Norwegian transport projects, *Research in Transportation Business and Management 28*, 33–39

Summary

Traditionally, project success has been interpreted in a narrow, operational sense, concerning the question of *How* to implement the project. More recently, the need to take a wider tactical and strategic view has been increasingly recognized, and includes the questions of *Why* and *What*, both of which concern the choice of concept. A project concept is a construct of thought that is meant to solve a problem or satisfy specific needs, and thus includes the entire business case.

In line with the broader interpretation of project success, the crucial role of the front-end phase of projects has been highlighted in literature. This is the stage from when the idea is conceived until a final implementation decision is made, during which it is still possible to make changes or to terminate the project, at an affordable cost. The key actor in this phase should be the project owner (the financing party), not the project manager. In major state-funded projects the Government is the true project owner on behalf of all taxpayers.

The starting point of this thesis was the need for a better understanding of success in public projects, how to govern them and how to evaluate them, with particular focus on the front-end phase and the choice of conceptual solution. Front-end management and project governance are still under-researched in the literature, and few empirical studies demonstrate experiences relating to governance frameworks.

In the year 2000, the Norwegian Ministry of Finance introduced an overall project governance framework applying to its largest projects, and the scheme was expanded in 2005 to include the choice of concept. The framework is often referred to as the Quality Assurance Scheme (the QA scheme), as it includes two external reviews during the front-end:

1. Quality assurance of the conceptual solution (QA1), performed at the end of the pre-study phase, before a Cabinet decision on whether to start a pre-project.
2. Quality assurance of the cost and management base (QA2), performed at the end of the pre-project phase, before the project is submitted to Parliament for approval and funding.

Together, the two intervention points are intended to cover all the three questions of *Why*, *What* and *How*. The intention is to ensure quality-at-entry before the final decision to fund a project is made.

The QA scheme has provided researchers with an exceptional opportunity to follow and learn from experiences as regards this specific initiative, taken by a government with the aim to improve the success of public projects. Reference to the scheme is therefore used extensively throughout the thesis, and strengths and weaknesses of the scheme are discussed as part of the research.

Three research questions have guided this PhD process. Although the scope of the research is major public projects in Norway and the Norwegian QA scheme, the research questions are

broad and open, as I wanted to develop deeper learning, rather than to test a set of hypotheses:

> RQ1. *What are the main challenges in front-end management of projects?*
>
> RQ2. *How can project success be achieved through project governance frameworks?*
>
> RQ3. *How can project success be achieved through improved appraisal and evaluation methods?*

The thesis consists of nine papers, each of which addresses a different issue and takes a different perspective. Together, they offer an overall picture of the potential for improving project performance through the use of project governance frameworks and better appraisal and evaluation models. Most of the studies reported in the papers were designed as multiple case studies, based on data gathered through document reviews, interviews and observations. Data analysis consisted of rather simple qualitative coding, categorization and summarizing of findings.

The first research question—RQ1 *What are the main challenges in front-end management of projects?*—is addressed in Papers 1–3. These papers are based on literature reviews and a number of Norwegian cross-case analyses of *pre-QA1* projects. Weaknesses and challenges in the front-end of public projects are framed as ten paradoxes in Paper 1. A key message is that strategic project success and front-end management ought to be given more attention. Despite the now well-known recommendations for exploring problems, needs, and alternative solutions (the *Why* and *What* questions), there is still a tendency to jump rapidly to the more narrow and short-term project management issues (the *How* question). Consequently, the financing party ought to make sure that all the three questions are properly addressed. Funding should be considered part of the conceptual solution, and perverse incentives during project initiation should be avoided.

The second research question—RQ2 *How can project success be achieved through project governance frameworks?*—is addressed in Papers 4–6, all of which are based on empirical studies. They discuss preliminary experiences relating to the Norwegian QA scheme compared with similar schemes in other countries. The Norwegian scheme is largely in line with key recommendations from the literature, and results are promising thus far. The use of external QA reviews has been essential. Interestingly, such a simple scheme established at the topmost level has triggered improved practices in ministries and agencies. However, some of the paradoxes in Paper 1 seem more difficult to overcome, and the findings in Papers 4-6 indicate that central government should continue to emphasize the importance of the tactical and strategic aspects. There is an opportunity for countries with similar schemes to learn from each other in order to develop and improve their own schemes. For example, the Norwegian Government may benefit from the experiences of countries that have introduced intervention points at an earlier stage.

A somewhat worrying finding is the large number of projects in which the government did not adhere to QA1 recommendations based on value for money as the overall success criterion. This gave rise to the third research question—RQ3 *How can project success be*

achieved through improved appraisal and evaluation methods?—which is addressed in Papers 7–9. A recurrent theme is that strategic project success is a multifaceted term, and that this ought to be reflected in the appraisal and evaluation models. In particular, "value for money" as measured by the cost-benefit analysis is not accepted by decision-makers as the *only* success criterion. If the cost-benefit analysis is overemphasized by quality assurers, there will be a risk that decision-makers will ignore their advice altogether. Instead, I recommend a framework based on the OECD DAC criteria, and present experiences of its application to 20 completed projects.

The contribution of this thesis is its exploration and synthesis of various challenges and weaknesses in the front-end of projects, in order to present and discuss experiences of the introduction of a fairly recent project governance scheme at Cabinet level, and to demonstrate the need for a holistic evaluation methodology. The findings illustrate the potential to improve project success, not least through external reviews, but I also note the difficulties in overcoming some of the paradoxes. Several papers include practical recommendations for project owners, ministries, agencies, quality assurers, and others involved in the appraisal and management of major public projects.

The findings from this thesis are only a small step, and the knowledge generated may be developed further in various directions. In the years to come, there will be a lot more to learn from the Norwegian QA scheme, through qualitative as well as quantitative studies, when projects subjected also to external quality assurance of the choice of conceptual solution (QA1) are completed and into their operational phase. We certainly need more data on the actual effects of the Norwegian scheme, as well as those of other project governance frameworks. Caution should always be exercised when expanding case study findings to different contexts, and I strongly recommend that more research is done on the implementation of project governance frameworks in other countries and under different circumstances.

1 Introduction

This chapter discusses the background and motivation for the PhD and presents the three research questions that have guided the work.

1.1 Background

Projects have become dominant means of organizing work across sectors and industries. Major projects are increasingly used for delivering public goods and services, such as transport infrastructure, defense acquisitions, public buildings, and major ICT projects. They amount to large investments, and their scale is tending to increase (Flyvbjerg, 2014). This in turn generates a need for interdisciplinary research to ensure competence, successful design, planning, and management of projects and programs.

Project success is a multifaceted term, as noted by Samset (2003). Traditionally, project success has been interpreted in a narrow, operational sense, concerning the question of *How* to implement the project. More recently, the need to take a wider tactical and strategic view has been increasingly recognized, by including the questions of *Why* and *What*, which concern the choice of conceptual solution. These questions need to be addressed *before* proceeding with detailed planning processes, as illustrated in Figure 1. However, tactical and strategic success may be difficult to define and assess, especially in large public projects, since it can only be determined in the long run.

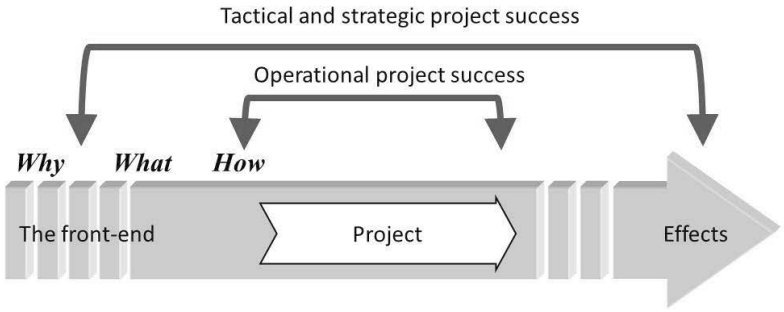


Figure 1 A simplified model of the project in the bigger picture

In the year 2000, the Norwegian Ministry of Finance introduced the Quality Assurance Scheme (the QA scheme) for major public investments, and the scheme was expanded in 2005 to include the choice of concept. The QA scheme is a simple stage-gate process with a top-down review of the quality of project proposals. The Norwegian QA scheme includes two external reviews during the front-end, which are intended to cover all the three questions of *Why*, *What* and *How*. The intention is to ensure quality-at-entry before the final decision to fund a project is made.

Parallel to the QA scheme, the Concept Research Programme was initiated by Professor Knut Samset at the Norwegian University of Science and Technology in 2002. I have been a

member of the research team since 2010 and its research director since 2011. The aim of the Concept Research Programme is to study front-end management and the choice of concept in major public projects, which was (and still is) an under-researched area in the literature. The QA scheme has allowed researchers to follow the largest public projects in Norway over the years, and has been a unique laboratory for research on longitudinal data. This is undoubtedly an interdisciplinary field, and members of the program have conducted separate studies in areas such as public management, project management, portfolio management, economic analysis, planning, decision-making, risk analysis, contract management, theory of incentives, applied logic, and judgmental assessment.

In this thesis I want to gain a better understanding of success in public projects, how to evaluate them and how to govern them, with particular focus on the front-end phase and the choice of concept. Reference to the Norwegian QA scheme is used extensively, and the strengths and weaknesses of the scheme are discussed as part of the research.

A convenient starting point has been to see the front-end phase as consisting of two processes that run in parallel: the analytic process and the decision-making process. The two processes may be interrelated to a greater or lesser extent, as illustrated in Figure 2 (the lines should to be interpreted as timelines, the circles as analyses produced, and the triangles as decisions made). Many challenges affect one or both processes, such as lack of competence among planners, shortage of time and resources for planning, the effect of hidden agendas, and unrealistic estimations of cost and benefits (Samset and Volden, 2013). Only when these challenges are well understood and recognized can project owners decide on the proper governance arrangements, in terms of requirements concerning decisions, analyses and the relationships between them.

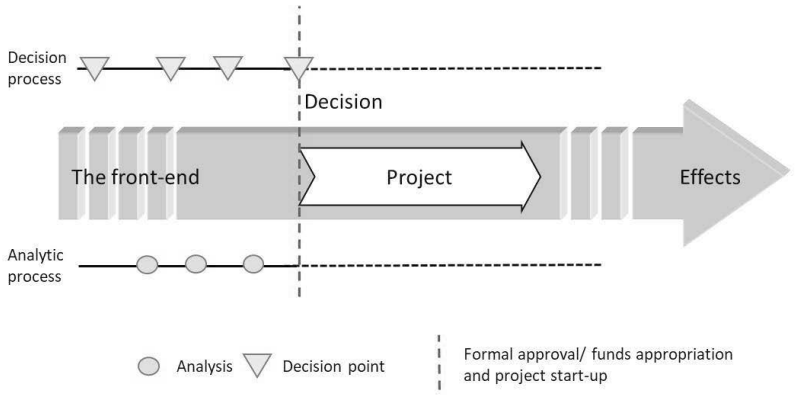


Figure 2 Two processes that run in parallel (based on Samset and Volden, 2013)

1.2 Research questions

The overall motivation for the research for this thesis has been the potential for improving public project success, especially in tactical and strategic terms. Front-end management and project governance are still under-researched in the literature, and few empirical studies have

reported experiences relating to governance frameworks. The Norwegian QA scheme is at the forefront internationally, and experiences based on this scheme ought to be disseminated.

Three research questions have guided this PhD procedure. They are broad and open, as I wanted to develop deeper learning rather than to test a set of hypotheses. The scope of the research is major public projects in Norway and the Norwegian QA scheme, as explained below:

RQ1. *What are the main challenges in front-end management of projects?*

This question is addressed through literature reviews and a number of Norwegian case studies (primarily projects that have *not* been subjected to QA) in which various challenges in the front-end of projects are explored and described.

RQ2. *How can project success be achieved through project governance frameworks?*

This question is addressed through empirical and case-based studies of the implementation of project governance frameworks from the topmost level of government, with the Norwegian QA scheme as the main research object.

RQ3. *How can project success be achieved through improved appraisal and evaluation methods?*

The question is addressed through empirical studies of the use of appraisal and evaluation methods on Norwegian projects under the QA scheme, and supplemented by a conceptual paper. As an economist, I am particularly concerned with the choice of appraisal and evaluation methodology and their use, including the use of cost-benefit analysis, which is an essential element of the current QA scheme. A relevant subquestion is whether there is need for a broader evaluation framework covering multiple aspects of project success.

The research questions are illustrated in relation to the project process in Figure 3. It should be noted that RQ2 and RQ3 are closely linked, as appraisal and evaluation methods should normally be an important part of any project governance framework.

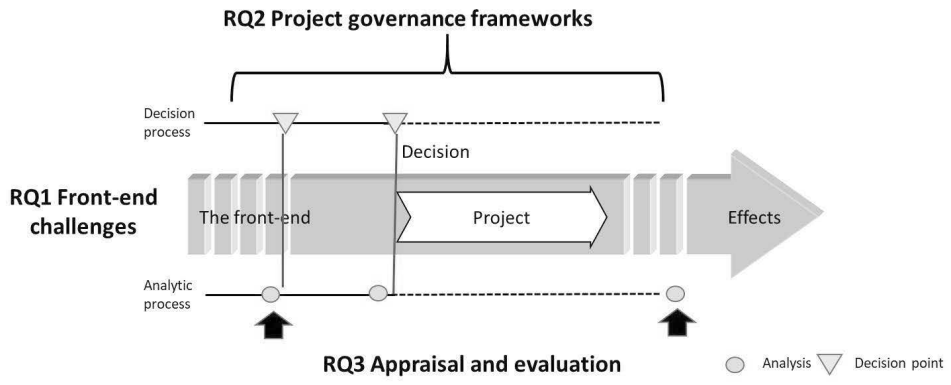


Figure 3 The three research questions

By addressing and discussing the three research questions, new knowledge can be gained about how the success rate of public projects can be improved, particularly in tactical and strategic terms.

1.3 Scope and limitations

It follows from the preceding section that, although the research questions are rather broad, the scope of this PhD thesis is more limited in practice.

The data—geography, time and sectors

The Norwegian QA scheme is the main object of research, which implies that most of the empirical studies are limited by the scope of the scheme. Accordingly,

- I focus on *state-funded* projects rather than private sector projects or projects funded by the municipal level in the public sector.
- I focus on *large* projects rather than smaller ones, as the QA scheme has a threshold level of NOK 750 million.
- With the exception of Paper 5, I study projects in only one country.
- The data are mainly from the period after 2005, when the full QA scheme with two intervention points was introduced. However, the studies also include some reference projects that have not been subjected to external QA, some of which are from the years before 2005.

The above-listed limitations imply that the relevance of the findings cannot be directly generalized to other settings. Nevertheless, they might be of interest beyond the Norwegian context. Many challenges and weaknesses in the front-end of projects are similar across country borders.

In terms of the specific sectors included, data from all types of projects covered by the QA scheme are applied and discussed, i.e. primarily transport infrastructure projects, defense projects, building construction projects, major public ICT projects, and certain major sports events that require financial guarantee from the State. Sectoral differences are discussed when relevant, but the thesis mainly aims to gain generic insights.

Subject matter and topics

The project process, as illustrated in Figures 1–3, is my overall conceptual framework that runs through all of the papers. This framework can be used to explain the thematic scope of the thesis:

- Project success is discussed as a holistic, multifaceted term, which in itself is a response to the narrow perspective often observed in the literature, the media and among practitioners. The main focus is on tactical and strategic success, which is least covered in the existing literature. Thus, projects are seen as measures for delivering specific goals rather than, *inter alia*, as examples of temporary organizations (as there are different ways to see a project, cf. Morris, 2013).

- My main focus is the front-end of projects, whereas the details of project implementation are of less interest.
- Projects are seen from the perspective of society, political decision-makers, user groups, and various other stakeholders, rather than from the perspective of the project team or the implementing agency. This implies that the thesis draws more on economics and the social sciences than engineering and project management.
- For the purpose of some of the papers, the perspective of a "rational" decision-maker (the financing party) is explicitly taken, implying that any adverse incentives at the topmost level are ignored. This simplification is done to make the study manageable. Besides, it can be argued that the complex and often unpredictable political processes are a necessary part of democracy, and difficult to counter by the use of project governance frameworks or improvements in appraisal methodology at the administrative level.
- Measures to improve project governance can be more or less formalized. In this thesis, I focus primarily on the structural aspects of project governance, in terms of formal regulations and requirements, as opposed to the human or attitude-related side of governance.

The terminology in this field may vary and partly be inconsistent in the literature. A list of key terms used in this thesis can be found in Appendix 1, but variants of the interpretations may occur.

The research design is primarily qualitative and case-based. In particular, any assessments of the effects of the QA scheme will be of a qualitative and preliminary nature. The number of projects in the QA scheme is growing, but data will not allow for a direct comparison with projects that were planned and implemented before the QA scheme was introduced or with major public projects under different governance frameworks (e.g., in other countries). Further, it should be noted that these processes take time, in most cases one or more decades. At present, no projects that have undergone the full scheme with two intervention points (starting in 2005) have been completed. In the years to come, more projects will enter their operational phases and new possibilities for conducting quantitative studies of their effects will arise.

1.4 Outline of the thesis

Having defined the research questions that guide this thesis, I proceed in Chapter 2 with a brief review of relevant literature. The chapter also presents the Norwegian quality assurance scheme. It should be noted that a more thorough review of the literature can be found in the papers.

In Chapter 3, I present the research design and methods used in the thesis. Again, a more detailed description of the applied methodology for the separate studies can be found in each of the included papers. Thus, Chapter 3 highlights general issues of methodology that are pertinent to the thesis as a whole.

Chapters 4, 5 and 6 address each of the three research questions through a presentation of the nine papers:

- Chapter 4 addresses the first research question (*What are the main challenges in front-end management of projects?*), with reference to Papers 1, 2 and 3.
- Chapter 5 addresses the second research question (*How can project success be achieved through project governance frameworks?*), with reference to Papers 4, 5 and 6.
- Chapter 6 addresses the third research question (*How can project success be achieved through improved appraisal and evaluation methods?*), with reference to Papers 7, 8 and 9.

It should be noted that the synthesizing part of the thesis is based on the included publications only, and does not introduce new elements. This is in line with established practice at NTNU for doctoral theses written as a compendium of academic papers.

Finally, Chapter 7 summarizes the main findings and discusses the contribution of this thesis to theory and practitioners, and further work.

2 Theoretical background

This chapter presents a review of relevant literature on which this thesis is built. Key topics from the project management literature include how to define successful projects and how to avoid unrealistic or biased estimates of project impact. The importance of project governance arrangements, particularly at the front-end, in securing projects' tactical and strategic success, is increasingly recognized in the project management literature. The thesis builds on this literature, in addition to literature in economics and the social sciences on specific appraisal and evaluation techniques and their use. A brief presentation of the Norwegian quality assurance scheme, which was established to secure quality-at-entry for major public projects, is also provided.

2.1 Different views and perspectives on project success

The term "project success," used as an indicator, is highly complex and aggregated. It may be interpreted differently by various individuals and institutions, in addition to being time-dependent. Three decades ago, Pinto and Slevin (1988), concluded that: "the concept of project success has remained ambiguously defined both in the project management literature and, indeed, often within the psyche of project managers. Until project management can arrive at a generally agreed upon determinant of success, our attempts to accurately monitor and anticipate project outcomes will be severely restricted."

Traditionally, the project management community has focused narrowly on the "iron triangle" of cost, time, and scope (Morris, 2013). In recent years, many authors have argued for the need to take a wider, strategic view on projects. Projects are implemented to deliver benefits and create value for users, parent organizations, and/or society at large (Baccarini, 1999; Hjelmbrekke et al., 2017; Morris, 2013; Shenhar, 2004; Williams & Samset, 2010; Zwikael & Smyrk, 2012). However, the definitions of these overlapping concepts (such as, project benefits, value, and value creation), may also be unclear and multifaceted (Laursen & Svejvig, 2016). In private sector projects, the ultimate goal is to improve the company's profitability, either directly or indirectly through improvements in its competitiveness. In public projects, the commissioner is the government, representing the entire society and its taxpayers. In such cases, the benefits of the project ought to be considered in a user perspective as well as in a broader societal perspective, to ensure that the project provides value for money and contributes to the desired development.

A distinction is sometimes made between projects and programs, where outcomes are achieved only at the program level, as contrasted with outputs at the project level. It has been shown that the risk factors associated with the two levels (and implicitly, the skills required to manage them) may differ (Aritua et al., 2011). However, in line with Williams et al. (2019) this thesis will not emphasize the project/program distinction, but rather argue that the work done here applies to both programs and large, self-standing projects.

Samset (2003) argues that in order to be truly successful, public investment projects should perform well in three perspectives, namely operationally, tactically and strategically, corresponding to the three levels of the project strategy:

1. Operational success is related to the project outputs (i.e., the specific agreed results that should be produced when the project has been implemented).
2. Tactical success is related to the first-order effects, typically as defined by agreed goals for the user effects (corresponding to what Chih and Zwikael (2015) refer to as target benefits).
3. Strategic success is related to the long-term objective, which explains the rationale behind the project.

This categorization has been a starting point for the thesis. A similar categorization is suggested by Baccarini (1999), who defines two levels of project success: project management success, which concerns delivery, and product success, which concerns the outcome. Similarly, Shenhar et al. (2001) offer a chronological sequence of events as a compound definition of project success: (1) meeting time, budget, and other requirements, (2) impact on the customer, (3) benefit to the performing organization, and (4) preparing for the future.

Samset's three-tier definition of project success is simple, but it should be noted that each level of analysis is in itself complex and multidimensional, as illustrated in Figure 4. Operational success, which is assessed in a short-term perspective, means meeting short-term performance targets, such as producing agreed outputs within budget and on time. This is essentially the project management perspective. By contrast, tactical performance is broader and should focus on the extent to which the project has achieved its formal goal, and whether the project is relevant in relation to current needs and priorities, but also whether the project's impact is predominantly positive, taking all positive and negative side effects into account. Clearly, these measures are more ambitious, realization will have to be assessed at a later stage, and there is more uncertainty involved. This is basically the users' perspective. The broadest interpretation of success is associated with the strategic perspective, which could be based on, for example, measures of the long-term economic effect, the extent to which it can be sustained and whether needs are satisfied in the long term. This is normally the investor's perspective, in this case the broader society.

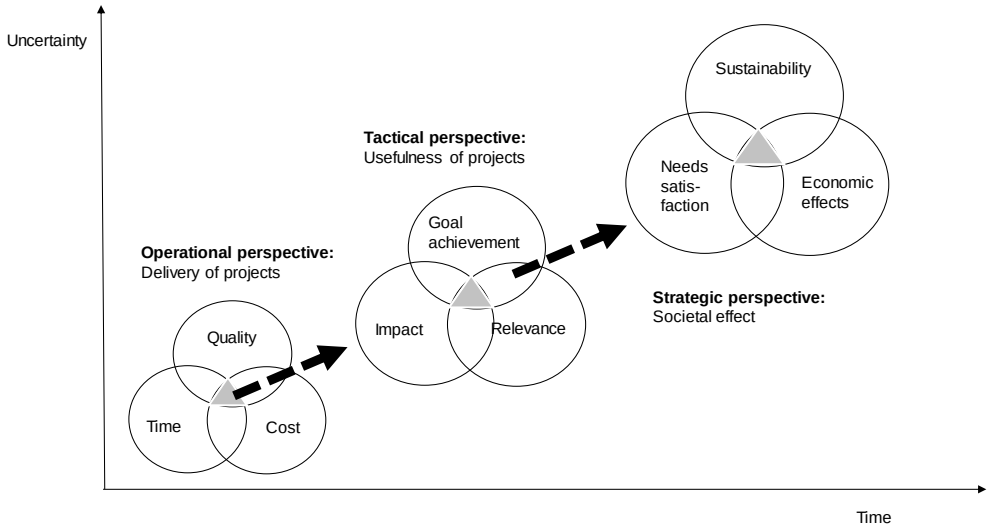


Figure 4 Different measures of success in three perspectives (reproduced from Samset, 2003)

In this thesis I recognize all perspectives, but pay special attention to the broadest perspectives—tactical and strategic project success. The choice of specific evaluation criteria for measuring project success at the highest level is discussed further in Section 2.5.

2.2 The poor track record of public projects

Public investment projects face a number of challenges and have varying reputations. There is broad literature on what Hall (1981) termed "great planning disasters," which are projects with cost overruns, time delays, and either no or very limited benefits, some of which may be so controversial and unfeasible that they end up being closed down or completely redesigned. Most of these studies seem to focus on operational project success. The problem of cost overrun is particularly well documented. Morris and Hough (1987) examined more than 3,500 major public projects across different sectors and countries, and found that cost overruns were typically in the range 40–200% and few projects were completed below budget. A number of recent studies have shown that cost overruns are common in infrastructure projects (Ansar et al., 2016; Berechman & Wu, 2006; Cantarelli et al., 2012; Flyvbjerg et al., 2003; Love et al., 2012; Lundberg et al., 2011; Makovšek et al., 2012; McDonald, 2002; Odeck, 2004).

Even more serious, but less well documented, is the type of problem that occurs when projects do not meet the expectations of users and society because they have failed in tactical and strategic terms. In extreme cases, the investment may be wasted. Flyvbjerg et al. (2003) and van Wee (2007) documented that in numerous transport projects, demand estimates were of equally poor quality as the cost estimates, and benefits were often overestimated. Moreover, environmental and other negative side effects are typically systematically downplayed, whereas regional economic benefits are overestimated (Flyvbjerg et al., 2003). Solberg and Preuss (2007) describe how major sporting events are often justified by

anticipated tourism-related and other wider economic impacts, but fail to realize such effects. Furthermore, some authors note that many ICT projects do not meet the expectations of users and end up being abandoned or reworked (Cicmil & Braddon, 2012; Pinto, 2006).

Project failure is not limited to the public sector, as shown by, for example, Merrow (2011), who documents similar problems in the private sector. However, the public sector has some additional challenges, including the occurrence of multiple objectives, difficulties in measuring success, and having to deal with an array of external stakeholders and the randomness of democratic decision-making processes (Klakegg & Volden, 2017). Public projects are the outcome of a political tug-of-war between stakeholders in society, whose needs and priorities will concur or conflict to varying degrees, and the outcomes of such processes are not always predictable. This is clearly shown in Miller and Lessard's study of 60 international projects (Miller & Lessard, 2000). Also, the public sector has some internal challenges, such as a weakness in establishing strategic visions, lack of professional skills and coordination among levels and actors, as noted by the Organization for Economic Co-operation and Development (OECD) (OECD, 2015). "Mega-projects" have particular characteristics that make them difficult, such as scale, complexity, risk, and the need of social responsibility (Williams et al., 2019).

Cantarelli et al. (2010) offer four explanations for cost overruns (which apply equally to benefit shortfalls), each of which may be relevant to varying degrees in specific projects, but generally they may reinforce each other: technical, economic, psychological, and political.

- *Technical explanations* are related to poor project design, incomplete estimation, and lack of tools, methods, and experience.
- *Economic explanations* concern lack of time and resources invested in the planning phase.
- *Psychological explanations* are based on the concept of planning fallacy and optimism bias.
- *Political explanations* are closely related to stakeholders' incentives to manipulate estimates in order to increase the chance that a specific project will be selected.

In Norway, the latter explanation may be of particular significance, since the local democracy strands strong, while at the same time the local government is financially weak and dependent on the state for funding of local infrastructure. Unfortunately, political determinants are understudied in project research. Miller and Hobbs (2005) discuss how public projects can be affected by deception and irresponsibility if stakeholders pursue hidden agendas rather than strive for openness and social responsibility. Additionally, Bent Flyvbjerg and his colleagues discuss, in several publications, the deliberate underestimation of costs and overestimation of benefits on the part of key stakeholders and project promoters, and have coined the term "strategic misrepresentation" (Flyvbjerg, 2009; Flyvbjerg et al. 2002; 2003; 2009). Other authors discuss similar issues (Locatelli et al., 2017; Mackie & Preston, 1998; Solberg & Preuss, 2007; 2015; Wachs, 1987; 1989). However, as noted by Siemiatycki (2016), these studies have not been very influential in the engineering and project management field, in which cost overruns are still largely explained by "honest

errors." Deliberate manipulation is difficult to prove and especially difficult to distinguish from overoptimism stemming from cognitive biases (Lovallo & Kahneman, 2003; Meyer, 2014).

However, in the development aid sector, adverse and even perverse incentives have long been studied in relation to project selection, which may be explained by "free funding." For a literature review, see Newby (2010) or Wiig and Holm-Hansen (2014). A pivotal study was conducted by Ostrom et al. (2001), which demonstrate serious problems with perverse incentives in Swedish-funded aid projects that resulted in the waste of public funds and adverse side effects such as corruption. An essential message is that it is naïve to believe that a scheme or project meant for the common good will be perceived that way by everyone involved. There will always be some who win and some who lose, and all those involved will adapt in a way that serve their own interests.

2.3 Project governance

Whereas the literature on project management is substantial and dates back to the 1950s, project governance has only recently become an issue of importance in the project management community.

Different interpretations and perspectives

The term governance is derived from the Latin word *gubernare*, meaning "to steer." A key governance issue is that the interests of the implementing agent will not necessarily be aligned with those of the financing party. Principal–agent theory has been useful to understand this constellation (Eisenhardt, 1989). An important part of governance schemes should be to ensure that key decisions are made at the appropriate level. Traditionally, governance has been linked to government and international institutions (public governance). When applied to projects, it should be understood as a system of appropriate checks and balances that enables transparency, accountability, and defined roles in the project, while at the same time supports project managers in delivering their objectives. This corresponds well with what Morris and Geraldi (2011) define as the institutional level of managing projects, which focuses on shaping the context and conditions to support and foster projects.

In the context of a project being undertaken by an organization, a related term is *corporate governance*, which refers to the mechanisms, processes, and relations by which the corporation is controlled and directed. A much cited textbook by Müller (2009) defines project governance as a subset of corporate governance, established to allow projects to achieve organizational objectives and foster implementation that is in the best interests of all stakeholders and the corporation itself. The Project Management Institute (PMI) defines project governance in a similar way, as "an oversight function that is aligned with the organization's governance model and that encompasses the project life cycle [by providing] a comprehensive, consistent method of controlling the project and ensuring its success by defining and documenting and communicating reliable, repeatable project practices" (Project Management Institute, 2017). The Association for Project Management (APM, 2011) has established a set of principles *for the governance of project management*. This, too, should be

seen as an area of corporate governance that concerns projects. However, for the purpose of this thesis, I take the perspective of the broader society and the government, rather than that of a corporation or an organization.

As noted by Ahola et al. (2014), the literature on the topic of project governance is still fragmented. Different conceptual models have been suggested to categorize the various strands of the project governance literature. Müller et al. (2015) distinguish between *project governance* on the one hand and *governance of projects* on the other hand: project governance refers to the governance of a single project, and includes such topics as the project manager's sovereignty and authority, involvement of various stakeholders, and the use of project boards. By contrast, governance of projects refers to governance of groups of projects, and includes matters such as the institutionalization of project management methodologies, reporting systems, project selection techniques, and program and portfolio management. A similar categorization has been made by Too and Weaver (2014)[1] and by Ahola et al. (2014). Williams et al. (2010) distinguish between governance *of* projects, which aims at efficient delivery, and governance *through* projects, which aims at choosing the right concepts and ensuring that effects are realized and are sustainable. This corresponds well with Samset's definition (i.e., the measures necessary to ensure operational success defined by efficiency and cost compliance, and tactical and strategic success in terms of impact on users and society) (Samset, 2003).

Project governance and project ownership are closely related, as the owner should be responsible for introducing a project governance framework. Williams et al. (2019) refer to the owner as the organization or person who ultimately derives the strategic benefits from the project. In major state-funded projects, the Government is the true project owner on behalf of all taxpayers. Christensen (2009) notes that in such projects the government itself should be involved on a strategic level, such as approving very large and critical projects. However, the literature on the project owner's role is limited and partly confusing. Most authors do not seem to focus on the owner as such, but on individuals in their role of "governance agents" as defined by a governance framework. In particular, the project sponsor role has attained focus in the literature (e.g., Crawford et al., 2008; Helm & Remington, 2005; Kloppenborg et al., 2009). Several authors have noted that project ownership (or sponsorship) is multifaceted, and that different perspectives must be balanced, such as the governance of projects and the governance through projects perspectives, the control and support perspectives, and the external and internal perspective on projects (Ahola et al., 2014; Crawford et al., 2008; Olsson & Berg-Johansen, 2016). Crawford et al. (2008) find that the project owner role is played out quite differently in different organizations. Olsson and Berg-Johansen (2016) studied project ownership in seven Norwegian projects and found that the support perspective

[1] Too and Weaver (2014) note that publications that discuss project governance can be classified in two main groups. The first group focuses on governance of single projects, typically involving several actors and stakeholders, when a contract specifies the specific governance arrangements for that project. The second group of publications examines governance models that link different project-related levels (project, program, and portfolio) within an organization, and thus sees project governance as a subset of corporate governance.

was present in all projects, whereas the strategic and external perspective, focusing on the business case and benefits realization, was more or less absent.

Project governance frameworks

During the research for this thesis I focused on governance frameworks, of which the Norwegian QA scheme is one example. As defined by Samset and Volden (2014), governance frameworks for major investment projects comprise the processes and systems that need to be in place on behalf of the financing party to ensure successful investments. This typically includes a regulatory framework to ensure adequate quality-at-entry, compliance with agreed objectives, management and resolution of issues that may arise during the project, and standards for quality review of key governance documents. This thesis primarily addresses the formal roles and regulations, as opposed to the human or mentality side of governance, which Müller et al. (2015) refer to as governmentality[2]. Klakegg and Meistad (2014) refer to the two types of governance as structure-based and relationships-based.

Normally, owners or investors with many projects will establish a common framework to be applied to all their projects. There is no standard format, but project governance frameworks are commonly described in terms of a stage-gate model. Haanæs et al. (2006) suggest that the key elements of a project governance framework should be: clearly defined project phases, clearly defined decision points, a quality assured basis for the decisions, simplicity, and some degree of standardization and common terminology. Similarly, Narayanan and DeFillippi (2012) state that project governance frameworks should incorporate five elements: stage-gate approval processes, formal roles and responsibilities, stakeholder representation, quality assurance, and contracts and sign-offs. Klakegg (2010) argues that the most important governance functions in the front-end of projects are: defining a clear decision-making process, and controlling the quality of documents used as basis for decisions.

Figure 5 is a simplified illustration of a project governance framework that covers the project as a whole, from when the first idea is launched until long-term benefits are realized, and includes decision points and documentation requirements at key gateways. The number and names of the phases, as well as the location of decision points and requirements concerning analyses and their quality, may vary. As discussed above, the front-end phase and the choice of concept need to be highlighted in the framework.

[2] "Governmentality" is a combination of "governance" and "mentality," and addresses such aspects as top management's attitudes and ambitions regarding project work, support and confidence in the project manager, and more generally the cultural values that members of an organization share and respect (Müller et al., 2015).

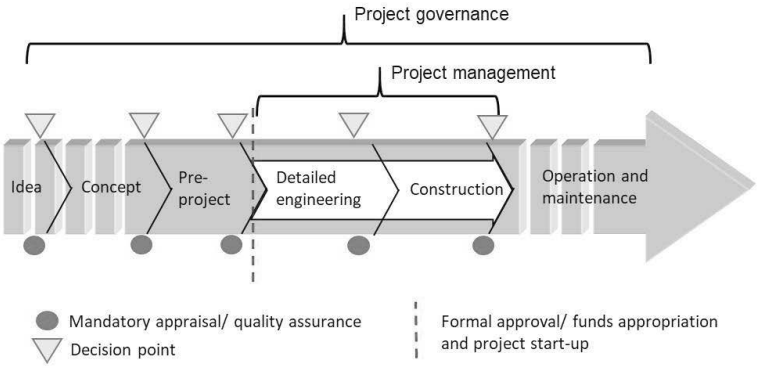


Figure 5 Simplified illustration of a project governance framework

Some authors have emphasized different aspects of project governance frameworks, depending, *inter alia*, on their explanations of project *failure* (see the discussion above about the control perspective versus the support perspective on project governance). However, common to all four explanations identified by Cantarelli et al. (2010) is that they suggest that independent quality assurance of decision documents is helpful. This is essential in ensuring more realistic estimations of cost, risk, and benefits, as well as a transparent planning process.

Project governance models can be more or less detailed. Whereas some require that a number of rules and procedure are followed, others give more autonomy to the project manager (i.e., the behavior-oriented model versus outcome-oriented model). These may also be denoted as bottom-up and top-down (Müller, 2009). The top-down model should be more appropriate in organizations with a high level of trust and a high level of project management skills.

Each organization needs to establish a framework according to its needs, but some standard models exist. For public organizations, a commonly used scheme internationally is PRINCE2® (Projects IN Controlled Environments, see www.axelos.com). The scheme was developed in the UK, originally for ICT projects, but has since been developed into a more general framework, integrating the governance of projects, programs, and portfolios and with an associated certification scheme. Similarly, the Association for Project Management (APM, 2011) has defined four main components of schemes that adhere to their principles for the governance of project management: (1) Portfolio management (ensuring that each project is aligned with key business objectives), (2) Project sponsorship (providing a link between the permanent and the temporary organization, typically by defining a "governance agent" with decision making accountabilities), (3) Project management capability (which is a question of skills, available tools and processes, and a clear mandate (among others), and (4) Disclosure and reporting.

Flexibility is crucial. Although it may be useful to have a common project governance framework for all projects in an organization, Miller and Hobbs (2005) emphasize that the scheme should not be static, as the need for governance may vary across projects and project phases. Müller et al. (2014) identify "organizational enablers" for good governance and governmentality, and their most prevalent finding is the importance of flexibility.

2.4 The front-end and the choice of concept

In line with the broader interpretation of project success and the growing literature on project governance, several authors have highlighted the crucial role of the front-end phase of projects (e.g., Morris, 2013; Shenhar, 2004; Williams and Samset, 2010; Williams et al., 2019). This is the stage from when the idea is conceived, until a final implementation decision is made, and during which it is still possible to make changes or to terminate the project, at an affordable cost. It is often referred to as the vital shaping part of the project. Once the decision to implement the project is made, essential choices will soon become locked, and it may be more difficult and expensive to change the overall design. Governance arrangements should therefore be particularly useful in this phase.

It seems clear that spending resources on early planning (also referred to as front-end loading) typically pays off, and correspondingly, that the roots of problems in later phases are found at the front-end (Williams et al., 2019). A study by the World Bank based on a review of some 1125 projects concluded that 80% of the projects with a satisfactory quality-at-entry"[3] were successful when examined in retrospect, while only 35% of those with an unsatisfactory quality-at-entry achieved success (World Bank, 1996). Similarly, Miller and Lessard (2000) found that the projects that attained the best results had allocated greater portions of their overall costs to their front-end phases. The cost allocations varied from 3% for simple projects to as much as 35% for complex projects. Moreover, it was found that the use of risk analysis was vital in the most successful projects, and that there was a decided advantage in holding an open debate during project planning. A corresponding study of 23 Norwegian projects revealed similar findings (Whist & Christensen, 2011).

Although the significance of the front-end phase is increasingly highlighted, it is still not well understood, as noted by Williams et al. (2019). For example, there is not agreement in the literature as to whether the front-end is a part of the project or a phase that comes ahead of the project. Further, it is not entirely clear which competencies are most important in this phase. Peter Morris (2011:7) writes that "It is evident from an extensive amount of research that management of the front-end definitional stages of projects is of overwhelming importance to their ultimate outcome yet we have little empirical data to suggest how best management competencies here should be improved." However, it is clear that the front-end is the phase in which strategic success or failure is set, and that the key actor in this phase should be the project owner (sometimes referred to as funder or permanent organization), not the project manager (Williams et al., 2019).

Many problems facing public investment projects that have to be overcome in order to achieve project success can be interpreted in terms of deficiencies in either the analytical process or political process or in both. The two processes run in parallel, and may be interrelated to a greater or lesser extent. As shown in Figure 6, there are two extreme situations. In one situation, project planning is predominantly consistent with a technocratic

[3] Quality-at-entry was used as an indicator to characterize the identification, preparation and appraisal process to which the projects had been subjected up-front.

model, in which major decisions follow from sound and comprehensive analyses ("the rational ideal"). In the other situation, the processes may be completely anarchic with unpredictable outcome. Ultimately, the complexity of the decision situation will depend much on whether there is agreement about what one wants to achieve and what constitute the best means to that end (Samset & Christensen, 2017).

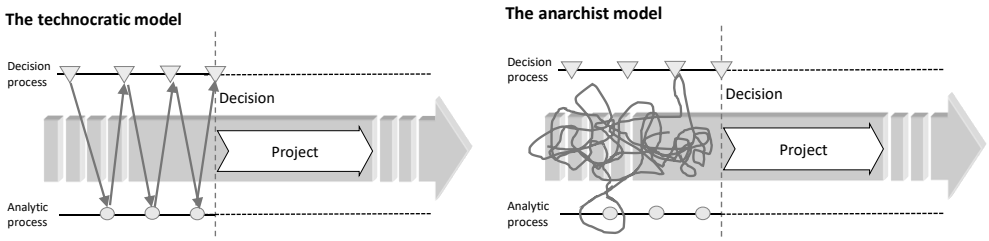


Figure 6 Analytic and decision-making process (based on Samset, 2015)

Williams and Samset (2010) note that the initial choice of concept is of critical importance in projects. The choice of concept includes the entire business case and represents the key decision made during the lifetime of a project, which is likely to have the greatest impact on long-term success or failure. A project concept is a construct of thought that is meant to solve a problem or satisfy specific needs. From the outset, it should be such that several different conceptual solutions might be perceived as possible solutions to the same problem. For example, if the problem is identified as congested roads, three different conceptual solutions could be: (1) a new highway with increased capacity, (2) a public transport program, and (3) congestion pricing.

In practice, it may be virtually impossible to reverse the acceptance of a project concept at an early stage, due to the expectations it has generated. For example, if the citizens of a small town perceive they have been "promised" a new highway, they will hardly accept that the decision is reversed to congestion pricing at a later stage. Cantarelli et al. (2012) use the term "lock-in" to describe the situation in which decision-makers are de facto *committed* long before any formal decision to build is made. For the same reason, van Wee and Rietveld (2013) find it very likely that the extent of cost overruns reported in the international literature is greatly underestimated because most studies compare the final cost with the formal budget. Instead, they should compare the final cost with the estimate at the time of de facto approval. Due to the risk of lock-in, it is crucial that the choice of concept is well-founded and that unviable concepts are sifted out at an early stage.

2.5 Project appraisal and evaluation

It is generally recommended that the choice of concept should be approved on the basis of a business case in which the expected benefits and strategic outcomes are described (Jenner, 2015). The business case provides a rationale for the preferred solution, and is crucial for the management of benefits and cost throughout the project life cycle (Musawir et al., 2017;

Serra and Kunc, 2015). However, the specific analyses to be included in the business case should be open for discussion.

Cost-benefit analysis

Cost-benefit analysis (CBA), which originated from welfare economics, has since the 1960s and 1970s been recognized as the major appraisal technique for public investments and public policy. A number of authors, as well as governments and professional project management bodies, have highlighted the value-for-money perspective and the CBA as a crucial part of the business case for public projects (e.g., APM, 2018; Jenner, 2015; Laursen & Svejvig, 2016; Project Management Institute, 2017; Terlizzi et al., 2017).

The CBA is a tool to determine whether the project's benefits exceed the costs, and can be used to rank projects unambiguously (Boardman et al., 2011). A monetary value is attributed to costs and benefits to the extent theoretically justifiable and expedient. Benefits are interpreted in terms of people's willingness to pay for them, which is grounded in an acceptance of consumer sovereignty, and the costs are defined by the value of the alternative uses of the resources (Boardman et al., 2011).

As a tool, CBA is attractive because it allows systematic assessments and precise recommendations. However, it may also obscure the fact that results depend heavily on uncertain and often subjective assumptions. Mackie and Preston (1998) list 21 sources of error and bias in CBAs of studied transport projects, and conclude that appraisal optimism was the most important. Other studies have found a high level of inaccuracy in key figures such as traffic forecasts (Flyvbjerg et al., 2003; Kelly et al., 2015; Nicolaisen & Driscoll, 2014; van Wee, 2007). Nicolaisen (2012) points to the lack of transparency in CBAs as an important challenge. He argues that more transparent modeling procedures and systematic ex post evaluation would enable better peer review of forecasts, thus reducing the risk of both naïve optimism and strategic behavior.

Some impacts are inherently difficult to attribute monetary values, and this entails a risk that they be ignored or underestimated. An issue that has attracted increasing attention since the start of the new millennium is the wider economic benefits of transport schemes, which are defined as benefits that accrue in the form of productivity gains (Standing Advisory Committee on Trunk Road Assessment (SACTRA), 1999). In several case studies (e.g., Venables, 2007; Vickerman, 2008), researchers have since tried to estimate such impacts and concluded that, in some cases, there is "additionality" to the direct benefits, but that these are context-specific and not relevant in all transport projects. At the other end of the scale, there are also indications that long-term environmental impacts and sustainability issues are being systematically ignored on the cost side (Næss et al., 2017).

Additionally, there is the more fundamental problem of ethical and political aspects being ignored by the CBA. Many authors argue that value for money is too narrow as a measure of projects' tactical and strategic success (Heinzerling & Ackerman, 2002; House, 2000). As an analytical too, it recognizes people's preferences only in their role as consumers, and not, *inter alia*, their political goals and strategies (Mackie et al., 2014). Furthermore, it is an

aggregated success criterion that disregards how costs and benefits are distributed (Nyborg, 2014). Thus, the CBA is of little help in cases in which the public sector has policy objectives that differ from consumers' preferences. A related critique is that the CBA systematically downplays the welfare of future generations, not least due to the use of a discount rate (e.g., Pearce et al., 2006).

The above-discussed weaknesses of the CBA, in particular its normative basis, may explain why its use in the decision-making process has been limited. Several studies have shown that CBA results have little impact on project selection (Annema, 2013; Eliasson et al. 2015; Kelly et al., 2015; Mouter, 2017; Nyborg, 1998). Different measures to mitigate these problems have been suggested, such as transparent and disaggregated analyses, extensive use of sensitivity tests, and separate analyses of distributional effects and other strategic considerations. An outside view is also recommended, such as having an independent third party perform or review the CBA. Additionally, systematic ex post evaluations should be performed to learn about the costs and benefits that can be expected (Flyvbjerg et al., 2003; Mackie & Preston, 1998).

Holistic models

Other appraisal and evaluation models that offer a broader and more holistic approach have been suggested. They may solve the problem of the CBA being a narrow success criterion, but not the problems of inaccurate and biased assessments. In the UK, the Five Case Model is recommended for the assessment of public projects. The model is designed to establish a case for investment by examining the five key areas: strategic case, economic case, commercial case, financial case, and management case (HM Treasury, 2013). At an early stage, the first two cases are highlighted: the strategic case focuses on the case for change and strategic fit, and the economic case focuses on value for money. Any conflicts between the two cases need to be solved and balanced. Later versions of the business case emphasize also the commercial, financial and management cases.

Another model is the goal-oriented evaluation model, which is widely used by development aid authorities and endorsed by the Development Assistance Committee (DAC) of the OECD, which harmonized the terminology among member countries. It is therefore commonly referred to as the OECD DAC criteria (OECD, 2002). Yet another model is the Sustainability Impact Assessment (OECD, 2010), and other frameworks emphasizing sustainability as the key success criterion. However, beyond the agreement that economic, environmental and social issues should be integrated and balanced, sustainability is a highly ambiguous term that is used differently within diverse contexts (Adams, 2006; Hasna, 2010). According to Adams (2006:3), the sustainability concept is "holistic, attractive and elastic but imprecise." A wide range of more specific tools that claim to measure sustainability exist, and typically focus on some elements and not others. Examples are Life Cycle Cost (LCC) analysis, Environmental Impact Assessment (EIA), stakeholder analysis, and scenario tools.

I pay special attention to the OECD DAC criteria and Professor Knut Samset's work on goal-oriented project evaluation (OECD, 2002, Samset, 2003). Clearly, a successful project is one that delivers its outputs and contributes significantly to the fulfillment of agreed objectives.

All projects are explicitly or implicitly based on an assumed set of causal relationships between inputs, project activities, outputs, and outcome, also known as the logic model (Samset, 2003). Drawing the logic model helps clarify the overarching goals, and the capacity and outputs of the project. However, a strong focus on intended effects as defined by the project owner may involve a risk of downplaying any side effects that can be attributed to the project (i.e., impacts beyond the intended effects) (Bakewell & Garbutt, 2005; Gasper, 2000; Sartorius, 1991).

The OECD DAC model stipulates a standardized set of five evaluation criteria, which are associated with the objectives of the logic model. It also attempts to capture any side effects. According to this framework, evaluation highlights (1) the need for the project (*relevance*), (2) whether the uses of resources are efficient (*efficiency*), (3) whether stated goals are achieved (*effectiveness*), (4) what other positive or negative effects may occur as a result of the project (*impacts*), and (5) whether the positive effects persist after the conclusion of the project (*sustainability*). It is primarily used for ex post evaluation. However, as suggested by Samset and Christensen (2017), the early appraisal of an investment case or a project should apply essentially the same evaluation criteria as will be used in ex post evaluation, and thus increase the likelihood of a successful project outcome.

As noted by Picciotto (2013), development projects are not so different from projects in high-income countries. The five criteria reflect lessons of experience, they can be used at project, program and policy level, and they are aligned with the results-oriented stance favored by most countries. A thorough review of the five criteria, which was performed by a group of professional evaluators (Chianza, 2008), concluded that the criteria work well and in particular that they satisfy Michael Scriven's Key Evaluation Checklist for program evaluation (see Scriven, 2015 for the most recent version). Of the five criteria, efficiency is supposed to cover the value-for-money dimension, but as noted by Chianza (2008) there is a tendency to define efficiency in the narrow sense, focusing mainly on cost and time-efficient delivery of the project.

Ex post evaluation—the missing link

Governments expend large amounts of resources on planning and implementing major public projects. Still, systematic ex post evaluations of public investment projects with respect to their outcomes are rarely conducted (Rambøll & Agenda Kaupang, 2016; Samset & Christensen, 2017). Our knowledge of projects' actual value for money, effectiveness, and sustainability may therefore be limited. For example, the International Transport Forum (ITF) refers to the lack of meaningful ex post evaluations of transport investments as the weak link in the assessment of transport infrastructure and policy (International Transport Forum, 2017). This contrasts with the advice of many authors to conduct ex post evaluation to improve appraisal tools and models, and to reduce inaccuracies in estimation of costs and benefits (Mackie & Preston, 1998; Nicolaisen & Driscoll, 2014). Ex post evaluation is also seen as a measure to provide the incentives for cost and benefits management throughout the project's phases (Flyvbjerg et al., 2003). In other words, evaluation can be a means for learning and improvement, as well as for accountability.

There is a considerable body of literature and experience on evaluation in other sectors, from which infrastructure projects could learn. Comprehensive textbooks exist in the field, such as those by Samset (2003) and Rossi et al. (2004). Evaluation is the systematic investigation of the feasibility of projects or other interventions. Evaluation requires certain expertise and application of scientific methods, while at the same time focusing on solving practical problems and being useful to project sponsors, decision-makers and other stakeholders (Rossi et al. 2004). Evaluation became particularly relevant in the USA in the 1960s, when it was associated with the Kennedy and Johnson administrations and the social programs implemented at the time. Its aim was to learn from successes and failures and to improve forward planning. The use of evaluation subsequently spread to other countries and different sectors, particularly to international development aid, where the effectiveness of investments and policy was contested.

There are several explanations for the lack of ex post evaluations of major public investment projects. Political decision-makers have, as per definition, a forward-looking perspective. Users of the infrastructure, who at best only financially contribute a limited share of the investment cost, may see the project as a success once the agreed outputs have been produced, regardless of cost. Long development times mean that there may be a gap between appraisal and evaluation with respect to the standard methodology. Furthermore, personnel, organizations, and systems change over time, and data become more difficult to obtain.

2.6 The Norwegian Ministry of Finance's QA scheme[4]

Against the backdrop of a series of negative experiences relating to major public investment projects in Norway, such as cost overruns, delays and benefits not being realized, the government initiated a study in 1997 to review the systems for planning, implementing and monitoring such projects. The authors (Berg et al., 1999) concluded that the underlying documentation was deficient in a number of projects. For example, defined objectives were vague and either overly ambitious or simply descriptions of tasks. Cost estimates were unrealistic, uncertainties not assessed, and a CBA was either non-existent or had major weaknesses. Further, a challenge in public investment projects was that planning processes were often sectoral and locally based. The front-end phase has typically been a bottom-up process in which ideas are generated locally by those who benefit from the project, and there may be strong incentives to overestimate benefits and underestimate costs. In the studied projects, broader economic analysis was typically conducted at a later stage when the conceptual solution has already been selected (Berg et al., 1999).

Berg et al. (1999) proposed the introduction of an external quality assurance (QA) scheme in the decision phase for the largest public projects. The QA scheme, introduced in the year 2000 and sometimes referred to as the State Project Model, is mandatory for investment projects with an anticipated budget exceeding NOK 750 million (approximately USD 90 million). The Ministry of Finance administrates the scheme, under which qualified external

[4] The description of the scheme is based on Christensen (2009) and Samset and Volden (2014).

consultants are assigned to perform quality assurance of the decision documents before two key gateways (see Figure 7):

1. Quality assurance of the conceptual solution (QA1), at the end of the pre-study phase, prior to a Cabinet decision on whether to start a pre-project.
2. Quality assurance of the cost and management base (QA2), at the end of the pre-project phase, before the project is submitted to Parliament for approval and funding.

These requirements apply to some 20–30 projects per year, mostly in the building, transport, construction, and ICT sectors. Initially, the aim of the requirements was to reduce the problems with cost overruns (i.e., ensure operational project success), now referred to as QA2. In 2005, the framework was expanded to include quality assurance of the choice of conceptual solution, QA1, in order to ensure that the right projects are started and that unviable projects are rejected (i.e., to ensure tactical and strategic project success).

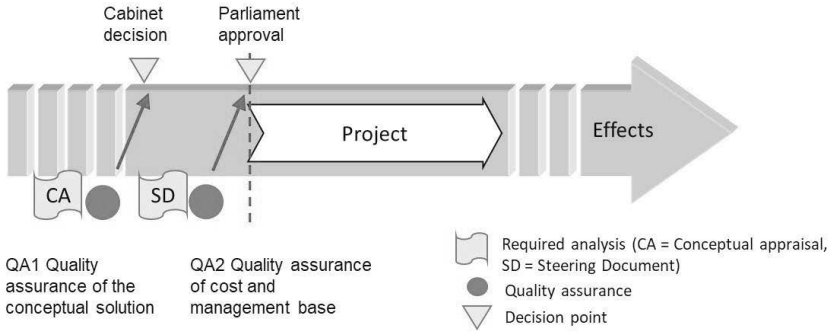


Figure 7 The Norwegian quality-at-entry regime

As basis for QA1, the responsible sectoral ministry or agency ought to prepare a Conceptual Appraisal (CA) report, which is essentially the business case. The CA comprises a needs analysis, project objectives, overall requirements to be fulfilled, an alternatives study in which the "opportunity space" is explored, a CBA of at least two alternatives and the "zero alternative" (no project), and, if relevant, a suggested implementation strategy for the recommended alternative.

The quality assurer examines the documentation and determines whether it provides a sufficient basis for decision-making. If the basis is insufficient, additional information may be requested. The quality assurer then prepares an independent CBA, including a stochastic estimation of the investment cost, and gives recommendations regarding the ranking of alternatives and the decision strategy. The findings and assessments are summarized in a report (the QA1 report), which is presented to both the sector ministry and the Ministry of Finance. Ultimately, the concept is selected through a political process in which external reviewers do not play any role. If an investment alternative is selected, the project proceeds to the pre-project phase. If the zero alternative is selected, the process is terminated.

As basis for QA2, the ministry or agency ought to provide a steering document (SD), including a complete base estimate for costs (and, if relevant, income/revenue), as well as an assessment of at least two alternative contract strategies. QA2 is performed at the end of the pre-project phase, and the aim is to provide the responsible ministry with an independent review of the cost and management base before it is submitted to Parliament for approval and funds appropriation.

In this case too, the quality assurer examines and verifies the documents and requires additional information if necessary. Thereafter, he or she will conduct a separate analysis of the investment cost and give recommendations regarding (1) the proposed budget, including necessary contingency reserves and the target cost for the agency, and (2) the appropriate strategy for managing the project in order to keep within the budget, including the management and authorization of contingency reserves. The proposed cost frame is normally P85, whereas the agency's target cost is lower, normally at the P50 level, in order to avoid incentives to spend contingency reserves. At the portfolio level, it should be expected that the projects hit P50 on average, so that no provisions are spent.

After the QA2 report is delivered, the Cabinet still has two options: either stop the project or allow it to enter the budget process. The final overall project budget is decided by Parliament, and the ministry will then determine the target cost for the executing agency. They are not required to follow the recommendations from QA2.

In effect, the QA scheme involves a fairly simple project governance framework with two decision gates. It aims to ensure that projects are well prepared and at a sufficient level of maturity before being presented to political decision-makers. The scheme only establishes requirements for the type of documentation that must exist before the two decision gates, and requires that the documentation is quality assured. Further, no specific changes to procedures of the various government agencies are required with respect to, for example, project management tools, project organization, and the use of governance agents, thus enabling the agencies to implement their projects as before to a large extent. The QA scheme is a top-down model, as opposed to a bottom-up one (cf. Müller, 2009). However, the anticipation of an external QA is expected to provide the agencies with incentives for high-quality planning processes. Both QA reports are normally made public.

Empirical research on project governance frameworks is still limited and fragmented. In particular, very few studies have been conducted on project governance frameworks at national level (i.e., government level). There are some exceptions, all of which applied data from the Norwegian framework. One is a study by Samset and Volden (2014), who present and discuss early experiences relating to the Norwegian framework, and conclude that the results are promising. Another exception is a study by Williams et al. (2010), who compared the Norwegian governance framework with the British one. They concluded that in all the four case projects examined, the governance framework was useful in its own way, but also that there was some potential for improvement, such as more assessment of the project during the early stages (which has since been introduced in the United Kingdom). Yet another study,

by Klakegg et al. (2015), compared the aforementioned two frameworks with the Dutch framework. The researchers concluded that consistent project governance provides rewards, but also noted that effort must be made to preserve the effect, otherwise it might “wear off.”

3 Methodology

This chapter presents the research design and methods used in the research for the thesis, and highlights general methodological issues that are pertinent to the thesis as a whole. A more detailed description of the applied methodology for the separate studies can be found in each of the included papers.

3.1 Research design

Front-end management of public projects is still a relatively new research area and to some extent unexplored. In the absence of well-established theories and large amounts of quantitative data, which was the case when the research process for this thesis started, an open-ended, *inductive* and *qualitative* approach to data collection and analysis is most appropriate (Samset, 2003; Creswell, 2014). The Norwegian QA scheme has provided researchers with an excellent opportunity to follow and learn from the experiences as regards one specific initiative taken by a government, with the aim to improve the success of public projects.

With this thesis I want to develop deeper learning, rather than to test a set of hypotheses. I want to *explore*, *describe*, and, to the extent possible, *explain* the various lessons from the Norwegian QA scheme. According to Yin (2014), exploring is the least ambitious and explaining is the most ambitious type of study, and case studies are well suited for all three purposes. Some of the included studies are partly also *evaluative*, in which findings are discussed in relation to specific criteria or evaluation models, such as Papers 4, 7 and 9.

Each of the nine papers addresses a different issue and takes a different perspective. Most of the studies on which the papers are founded were designed as multiple case studies, based on data gathered through document reviews, interviews and observations, as explained further in section 3.4. Data analysis has consisted of rather simple qualitative coding, categorization and summarizing of findings, which is often referred to as a general inductive approach (e.g., Thomas, 2006). This is in contrast to more specific methods, such as grounded theory, phenomenology and narrative analysis, which typically require a very specific research philosophy and the use of a technical language.

Although the research was primarily qualitative, the number of projects that undergone external QA is increasing, and for the purpose of two recent papers (Papers 7 and 9) I had access to relatively large numbers of projects. This meant that I could perform categorizations and counting of occurrences. This was primarily done to systematize researchers' assessments, and did not involve statistical analyses. In the future, when the number of completed QA projects is higher and the hypotheses may be clearer, the potential for quantitative analyses should improve.

All the steps in the research process were inevitably shaped by my own assumptions and experiences. An important strategy that was used to avoid subjective bias is *triangulation*. Triangulation involves compensating for the use of simple data collection methods and a simple study design by using several information sources and different methods

simultaneously in order to generate information about the same topics (Samset, 2003). In the included case studies, the use of document reviews, interviews and observation as different methods to illuminate each case, proved useful to reveal any inconsistencies in the data. When performing the interviews, respondents with different perspectives were consulted. Lastly, most of the studies were conducted in collaboration with one or more fellow researchers, with whom I discussed all of the steps in the study, including the research design, checklist for document review, interview guide, interview sampling, interpretation and categorization of findings, making decisions about what was important and less important in the data, and other matters. In most of the studies, we also used as advisers one or more researchers who were external to the study. This collaboration with other experienced researchers was very helpful when learning how to do research, and is assumed to have strengthen the reliability of the studies.

3.2 Research context—the Concept Research Programme

The research for this thesis was conducted under the auspices of the Concept Research Programme. Researchers have followed the Norwegian QA scheme since 2002 and collected project-specific data in an electronic archive called the Trailbase. A number of studies, ranging from single-case studies to multiple-case studies and to some extent quantitative studies have been conducted in a number of areas based on these data.

The Concept Research Programme is concerned with *applied* research and is often referred to as a *trailing research program.* The term trailing research was first introduced by Finne et al. (1995) as a model for program evaluation. The idea was to have a team of scientists trailing a program in real time and providing feedback to the clients who owned the program. However, the Concept Research Programme is more independent of its clients and users than suggested by Finne et al.'s definition (Finne et al., 1995). In this case, the financing party is the Ministry of Finance, and the users are the ministries and agencies affected by the QA scheme, as well as the consultants conducting the QAs. None of the parties can instruct researchers in their work. Furthermore, the user group is very wide in terms of the sectors and organizations involved.

In this thesis, I have made use of information from the Trailbase, drawn on previous studies, and benefitted from cooperation within Concept Research Programme's research team.

3.3 The research process

A research process is commonly said to comprise the following stages (Dudovskiy, 2018):

1. Selection of the research area
2. Formulation of research aim, objectives and RQs or developing hypotheses
3. Performance of a literature review
4. Selection of methods for data collection
5. Collection of primary data
6. Data analysis

7. Drawing conclusions
8. Completion of the research.

These steps fit well with the process that has been followed in the research for this thesis, but the process has been somewhat more iterative than sequential in practice.

My professional and personal interests are relating to the front-end phase and what governments can do to ensure that project concepts that will benefit society the most are selected. With this in mind, the process started with an open approach, and the research questions were defined somewhat later. I also started the literature review at an early stage, to establish the context of the research and to develop more specific research questions, but the review was expanded continuously and new searches were conducted for each of the nine papers.

Each of the nine papers included in this thesis represents a separate work that included a set of stages similar to those mentioned above. Seven papers report empirical studies, all of which included a different set of cases, adapted to the more specific research questions defined for that particular study. The definition and scope of the nine papers follow from my interests and the availability of empirical information. A wide range of other topics could have been addressed within the scope of the Norwegian QA scheme, but I had to prioritize due to time constraints.

Research questions 1–3 are repeated as follows:

> *RQ1 What are the main challenges in front-end management of projects?*
> *RQ2 How can project success be achieved through project governance frameworks?*
> *RQ3 How can project success be achieved through improved appraisal and evaluation methods?*

The scope of this thesis is major public projects in Norway and I draw extensively on experiences relating to the Norwegian QA scheme. Please refer to the discussion of scope and limitations in Sections 1.2–1.3.

The first research question (*What are the main challenges in front-end management of projects*) is rather broad, and it had to be addressed in an exploratory and descriptive nature. My focus in this phase was on gaining an overview of various challenges in the front-end management of major projects. Three papers were published in this process. Paper 1 presents a literature review and synthesis of a number of Norwegian cross-case analyses. Paper 2 goes deeper into one of the challenges, namely the perverse incentives resulting from concentrated benefits and no liabilities for the privileged groups. The paper is based on a multiple-case study. In contrast to this real-world situation, Paper 3, which is a conceptual paper, presents some overall principles concerning "how it should be," based on various recommendations from the literature.

RQ2 concentrates on one type of solution to some of the problems by asking *How can project success be achieved through project governance frameworks?* The Norwegian quality assurance framework was the natural starting point for the investigation. The question is

addressed in three papers (Papers 4, 5 and 6), all of which report empirical, case-based studies, with a focus that is more descriptive and explanatory than purely exploratory. Paper 4 presents and discusses experiences relating to the Norwegian QA scheme, Paper 5 compares the scheme to similar schemes in other countries, and Paper 6 looks at ministries' and agencies' governance schemes at the subordinate level.

The third and last research question was formulated to explore the term "strategic project success": *How can project success be achieved through improved appraisal and evaluation methods?* The choice of appraisal and evaluation methodology should normally be an important element of a project governance scheme, and therefore this research question is closely related to the former. Papers 7 and 9 report empirical and case-based studies, in both cases with a relatively large number of projects. Paper 7 addresses the use of value-for-money assessments and Paper 9 proposes and tests a broader framework based on a goal-oriented evaluation model. Paper 8 is a conceptual paper that elaborates on one of the criteria in the broader framework, namely sustainability, which is often difficult to discuss across professional and academic disciplines.

In the last part of the research process the perspective broadened again, when summarizing and discussing the findings from all the papers and relating them to the existing knowledge base. This thesis marks the final step, in which the findings from all nine papers are seen together. The research process in this thesis progressed as shown in Figure 8, which is freely adapted from Blumberg et al. (2011) and Swales (1990).

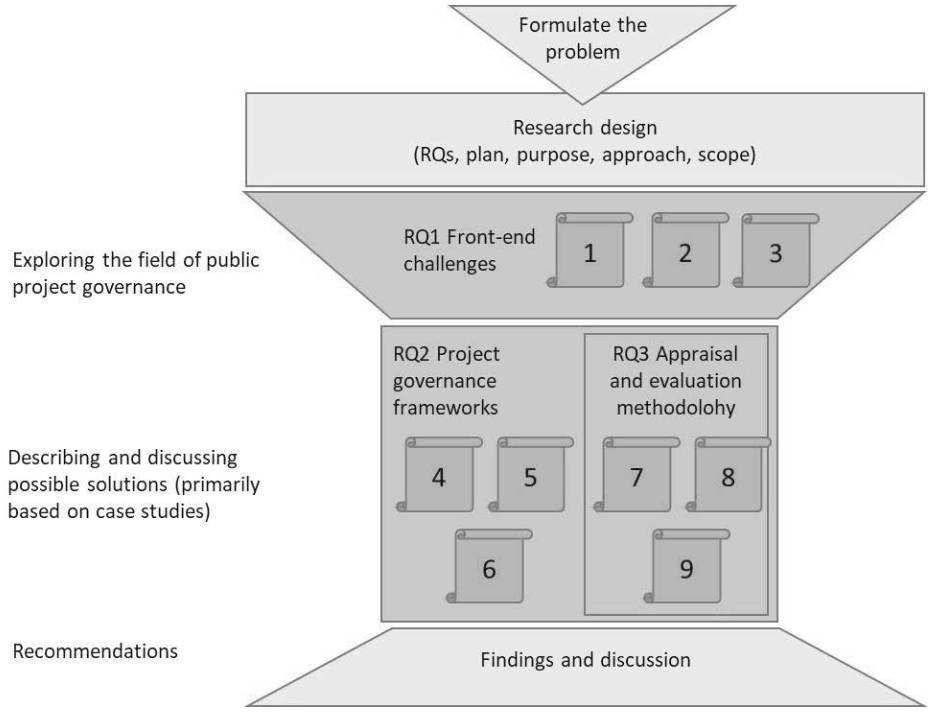


Figure 8 Research process and publications

It should be noted that although the papers are placed under one research question in Figure 8, some of them have results that are intended to answer more than just one research question.

3.4 Methods used

As indicated in the preceding section, most of the nine papers apply an inductive and qualitative research design, based on case studies. I and my co-authors have had access to considerable amounts of secondary data on a number of projects, and had excellent opportunities to collect our own primary data on these projects, primarily through interviews.

A description of each method of data collection and analysis in the research process is presented in the following subsections. In addition, Table 3 provides an overview of the methods used in each paper.

Literature review

As a researcher, one always needs at an early stage in the research process to review the literature on one's topic in order to understand the problem, to know what other researchers have done and how they define key terms. I have not published a separate literature review paper as part of this PhD, but have continuously maintained and expanded a pool of references over time and cited them in all of the papers.

The literature search and review process ended up being somewhat more iterative than first planned, but I used the five steps specified by Blumberg et al. (2011) as a guide:

1. Build an information pool
2. Apply filter to reduce pool size
3. Rough assessment of sources to further reduce pool size
4. Analysis of the literature in the pool
5. Refine filters or stop search.

I started out with a broad orientation and searched through Internet-based search engines (in particular, Google Scholar and the Norwegian libraries search engine, ORIA) with rather generic search terms, such as project success, project value, project governance, governance framework, project owner, public project, and project's front-end. This provided an introduction to the field, and led to the identification of particularly relevant authors, journals and articles. Thereafter, it was relatively easy to follow references in current documents back to the origins of the line of research. A separate literature review was also conducted for the purpose of each paper, following a similar procedure.

Clearly, there is no guarantee that the kind of narrative literature review is exhaustive and unbiased. However, the field of project governance and front-end management of major projects is still relatively new and the amount of literature is limited. Hence, I considered that gaining an overview of the material was manageable.

It should also be noted that just before completing this thesis, a systematic literature review on the front-end of projects, funded by the Project Management Institute, was published as a

journal paper (Williams et al., 2019), and was very useful as a final check on whether I had missed any important references.

Case studies

Case studies were used extensively, as this form of research fitted well with my research questions and data. In most cases I did not have clear hypotheses in advance, and needed to rely on multiple sources of evidence to understand the research units and their context. As noted by Yin (2014), case study research should be the preferred method when (1) the main research questions are "how" and "why" questions, (2) the researcher has little or no control over behavioral events, and (3) the focus of study is a contemporary (not historical) phenomenon in its real-world context. Flyvbjerg (2006) highlights the role of cases in knowledge accumulation, and suggests that if one has a strategic choice of cases (e.g., extreme cases, maximum variation cases, critical cases, and paradigmatic cases), it is possible to generalize from very few cases or even from a single case.

For most of the included papers in this thesis, multiple case designs have been applied, which are generally stronger than single-case designs. A different set of cases were selected for each paper, depending on the research topics addressed in that paper. The cases were at different levels: country, sector and project:

Paper 4 focuses solely on one *country*, Norway, which was selected because it is a critical case (assumed best case). Paper 5 includes Norway and five other countries, all of which were selected because they too are at the forefront in developing project governance systems. Paper 6 uses the *sectors* as cases. In the paper, project governance in the relevant ministries and government agencies in each sector was studied. *All* of the main sectors affected by the QA scheme were included, as well as two other sectors for comparison.

The remaining empirical papers use individual *projects* as cases. I and my co-authors wanted to cover typical QA1 projects and to include projects from all sectors. However, since the QA1 scheme was introduced several years *after* QA2, no QA1 project has yet been completed. Therefore, all of the QA1 case projects were still in their front-end phases when we studied them. However, completed (QA2) projects were useful for the study of what happens in the longer term, in a situation *without* QA1. For example, in Papers 1 and 2, in which we aim to describe challenges in front-end governance of projects, we use several sets of projects, most of which were not subjected to QA1, and some were strategically selected to illustrate particular problems (i.e., extreme cases).

As noted by Samset (2003), a case study is not really a data collection technique like the others discussed in this section, but usually involves a combination of other methods. In a case study, a researcher will have to rely on multiple sources of evidence, and the data will need to converge in a triangulating fashion. The most important sources are discussed below.

Document reviews

A document review involves collecting and analyzing existing data about the cases.

In terms of documents, I have been fortunate to have access to Trailbase. This archive includes data on more than 200 Norwegian major public investment projects and is still growing at the time of writing. Trailbase includes a wide range of data that are structured and made available for researchers. They include the QA reports (QA1, QA2, or both), appraisal and planning documents from the ministries and agencies, termination reports including the final cost (completed projects), and ex post evaluations (for a sample of projects in their operational phases).

A general risk was that secondary data might not fit the purpose of my specific study, and therefore the decision as to whether to include a specific document had to be based on a critical assessment of the relevance and quality of information. However, the QA reports are unique since they include the same information and were produced at the same time in all projects' life cycles, by external experts commissioned by the Ministry of Finance. The documents produced by ministries and agencies are also largely of good quality, but some reservations should be made. Based on these documents, a wide range of issues can be addressed and trends in the data material over time can be identified.

The limitations in the data material stem from the fact that the number of QA1 projects is still limited and none of them have yet been finalized (as already noted). The documents from Trailbase were therefore supplemented with other available data, not least government-issued documents, which were openly available on the government's website, but also newspaper articles, reports and statistics where relevant. In general, there is a considerable amount of available data on the large public projects, which often attract much attention from the media and the public.

However, for Papers 4–6, in which the studied cases were countries and sectors, the availability of written documentation was somewhat more varied, which was a finding in itself. In some cases, we had to rely more on spoken information.

Interviews

All of the case studies were based on interviews, in most cases held with experts from ministries, agencies and quality assurers to discuss the relevant topics relating to the Norwegian QA projects. Interviews were useful to obtain deeper understanding, and since the interviews were conducted after the document reviews, we were able to present some key findings and ask the interviewees to comment on them.

The ministries and agencies with projects under the Cabinet's QA scheme, as well as the quality assurers, have appointed a contact person, who should be available for questions about the QA projects and processes. I had access to this list of contact persons, which was a useful starting point for identifying interviewees (for snowball sampling, see e.g., Dudovskiy, 2018). This was cost-effective but also useful when recruiting experts in the organizations that otherwise would have been difficult to identify. However, using key people involved in the QA scheme as a point of access to the organization may in principle have introduced a risk of bias. Informants from other organizations were therefore also applied, as needed. In

most cases, it was essential to cover different (possibly conflicting) perspectives, such as ministry versus agency, ministry/agency versus quality assurer, and different roles in projects.

All interviews were semi-structured, implying that they started with a list of questions (sent out by email in advance), but also allowed for conversation and follow-up questions. With a few exceptions, the interviews were conducted in person, normally at the interviewee's workplace. In several cases, the interviews were conducted jointly by me and a fellow researcher. For Paper 9, a focus group interview was conducted with a group of experienced evaluators. This is an efficient method for generating information, where participants provide mutual inspiration and comments, so that more can be achieved in the course of a few hours than by conducting far more lengthy individual interviews (Samset, 2003). The process was well-planned and led in order to reduce the risk of negative group effects, although the avoidance of such effects can never be guaranteed.

Anonymity of all study participants was ensured. All interviews were transcribed, and the interviewees were given the opportunity to see and comment on the transcriptions.

My own observations

Observation, which involves collecting data through observing, was useful as a supplementary source of data. I did not apply the method systematically with checklists and observation record forms, but made general observations as part of other data collection methods, and received additional information on, for example, the surroundings and people's reactions to them.

For Paper 9, a meta-evaluation of 20 project evaluations was conducted. Each evaluation was based on a number of sources including site visits and the evaluators' own observations. I was personally involved as an evaluator in two of the project evaluations, and could therefore draw on my observations from these processes when conducting the meta-evaluation.

Furthermore, as a trailing researcher in the Concept Research Programme, I regularly participate in different arenas and meet representatives of the same organizations to discuss, formally and informally, similar issues to those covered in the thesis. One such important arena is the yearly QA Forum organized by the Ministry of Finance. Observations from these fora were useful for several of the included papers. In my experience, there is generally not much conflict between what people say in interviews and in more informal discussions, but these informal conversations add to the data and make them richer. However, as noted by Samset (2003), observation is a demanding method, and the quality of the results will depend on the observer's experience, skills and objectivity. In order to reduce possible sources of error and bias, I discussed the outcome of the observations with fellow researchers and in several cases also with the participants during interviews.

3.5 Analysis of the data

The collected various data relating to the cases (such as interview transcripts, notes and documents) were subsequently analyzed. When analyzing qualitative data, the intention is to reduce the data, display it and draw conclusions (Thomas, 2006).

Creswell (2014) lists six steps in data analysis and interpretation:

1. Organize and prepare the data for analysis (i.e., transcribe the interviews, type up notes and observations, and combine them with documents).
2. Read through all the data and reflect on its overall meaning.
3. Start the coding process. Make a list of topics as one goes through the material, and collate them into broader categories.
4. Generate detailed descriptions for the case studies. Build additional layers of analysis when possible. Reduce the number of categories.
5. Advance how the description and themes will be represented in the qualitative narrative.
6. Interpret the data and compare them with literature or extant theories. New questions can be suggested.

My analyses corresponded closely with these steps, and can alternatively be thought of as a simple coding scheme following the basic principles outlined by Weber (1990). The data analysis in each study was guided by a set of broad themes or research questions, which also typically formed the structure of the interview guides. The analyses were carried out by readings and interpretations of the data, first for a few "pilot cases" and thereafter for all cases, to allow for the cross-case analysis. This was done manually, either in Microsoft Word or in Excel. Throughout the process, categories were documented by writing down my comments and reflections (memos).

In some cases, I had expectations about the findings or thoughts about causal links or other relationships between categories, but not to the extent that specific hypotheses had been formulated to be tested in advance. The themes and categories that I ended up with deviated to a greater or lesser extent from those that I started out with.

For some of the papers (most notably Papers 7 and 9), I relied heavily on the study of documentation (CA/QA1 reports and ex post evaluations respectively in the two aforementioned papers). This included some simple categorizations and counting of occurrences. However, these were basically qualitative analyses, partly based on my own judgments of the documents, and therefore there was no need for the analyses to be quantified for statistical purposes.

3.6 Summary of methods used

The methods used for data collection and analysis in all nine papers are briefly summarized in Table 3. Further explanations can be found in each paper.

Table 3 An overview of methods used in each paper

No	Paper topic	Type of paper	Lit. review	Case studies	Document review	Interviews	Observations
1	Ten paradoxes	Literature reviews + summary of own empirical work	x	Several cross-case studies are presented, each with a different set of projects	x	x	x
2	Public funding, perverse incentives	Empirical	x	Nine projects (extreme cases)	x	x	
3	The proposal	Conceptual	x	n/a			
4	Quality assurance in Norway	Empirical	x	One country (critical case)	x	x	x
5	Governance frameworks in six countries	Empirical	x	Six countries (critical cases)	x	(x)	
6	The hierarchy of governance schemes	Empirical	x	Six sectors (total population)	x	x	(x)
7	Projects' value for money	Empirical	x	58 projects in their front-end phase (total population)	x	x	(x)
8	The concept of sustainability	Conceptual	x	n/a			
9	Meta-evaluation of 20 projects	Empirical	x	20 projects in their operational phases (typical cases)	x	x	x

3.7 Quality of the research

Although the research for this thesis was not intended to verify causal relations, it was nevertheless essential to persuade the readership that the findings are worthy of attention. Lincoln and Guba (1985) use the term *trustworthiness* when they discuss the quality of qualitative research. However, I apply the conventional terms reliability and validity.

In qualitative research, which gives considerable room for judgment, a major concern is subjectivity and potential bias, which may affect both reliability and validity negatively. More generally, qualitative studies are often considered "weak evidence" and biased towards verification. However, it should be noted that the question of bias applies to all methods,

including, for example, the choices of categories and variables in a quantitative study and the structure of a questionnaire. Flyvbjerg (2006) notes that researchers who conduct in-depth case studies more often report that their preconceived views were wrong, than do researchers who conduct quantitative studies. He argues that falsification, more than verification, characterizes the case study.

Reliability

Reliability refers to the quality of measurement, and the test is the extent to which the same answers will be obtained by a second researcher, using the same research methods under similar conditions (Dudovskiy, 2018). This refers to the quality of the data, but also to the coding and categorization process to obtain findings and draw conclusions. As noted by Samset (2003), reliability is closely related to the terms independence and impartiality, meaning, for example, the extent to which facts and judgments are kept separate, and the extent to which the research allows different stakeholders to be equally heard.

An important mechanism used to secure reliable data during the research for this thesis was making sure that reliable sources of information were consulted; this applied both to the document study and to the sampling of interviewees. Normally, interviewees who had first-hand experience of the phenomenon in question were chosen. Efforts were also made to avoid uncertainties and misunderstandings on the part of study participants. The interview guides were discussed and tested in advance to make sure questions were clear, so that interviewees would interpret them in the same way. As an interviewer, I tried to be as neutral as possible; for example, I did not overreact to an interviewee's responses. In many interviews, two researchers were present and/or made audio recordings and went through the material together afterwards. Detailed transcriptions were made of each interview to ensure that the informant's voice was still visible, and not only my interpretations of it.

To avoid subjective bias in data collection and analysis, triangulation was used extensively used (as discussed in Section 3.1).

Validity

Validity concerns whether we measure what we intend to measure and refers to the closeness or fit between an intellectual construct and the things we measure empirically (Samset, 2003). A distinction is often made between internal and external validity. While the former seeks to establish a causal relationship within the data, the latter defines the domain to which the findings can be generalized (Yin, 2014).

Although not intended to test a set of hypotheses about cause and effect, qualitative studies also concern patterns in the data and relationships between variables. Such relationships may have alternative explanations and it is therefore important to avoid "jumping to conclusions." Instead of seeking to "control for other factors, we should find other ways to consider, and preferably rule out, other theories. Triangulation was crucial in this respect. By combining several data sources and data collection methods, involving several analysts in the process, and using multiple cases, I tried to overcome the bias that comes from single-methods, single-

observer, and single-theory studies. In this way, I tried to validate information obtained from one source by checking other sources.

It should also be noted that some of the studies included a relatively large number of cases. Although validity in a qualitative study can never be guaranteed by increasing the number of data points, it is reassuring when the findings from similar projects are consistent.

Furthermore, the findings were discussed thoroughly with fellow researchers and external peers, and in many cases also with stakeholder groups (typically by going back to the interviewees and inviting them to read and comment on parts of the report. In most of the studies, I collaborated with fellow researchers in a multidisciplinary team, in which each researcher had a different professional background, experience, and perspective. All of the papers based on the studies have been published following a peer-review process and most of them have been discussed in conferences with international researchers.

Nevertheless, the findings can and should not be used alone to make predictions about a population outside the study, such as projects in other sectors, other countries or other time periods. As noted by Yin (2014), case studies can be used to expand and generalize theories (i.e., analytic generalizations), but never to extrapolate probabilities (i.e., statistical generalizations). The best I can hope for is that some lessons have been identified from the cases that may be of generic nature. This must be discussed in relation to previous studies. Valid results can only be established through a series of replications and validations, and this thesis is therefore only a small step. Further research is needed to clarify to what extent the lessons from this study are of a general nature or, for example, to what extent they are country-specific.

4 Challenges in projects' front-end phase (RQ1)

This chapter addresses the first research question—*What are the main challenges in front-end management of* projects—with reference to Papers 1, 2 and 3. Paper 1 presents the main challenges and weaknesses in the front-end of public projects, framed as ten paradoxes. Paper 2 elaborates on one of the paradoxes, namely what are termed perverse incentives, which may result when benefits are provided without corresponding liabilities for the beneficiaries. Finally, Paper 3 discusses a set of principles regarding how project proposals should be developed.

4.1 The front-end definition of projects—ten paradoxes (Paper 1)

Paper 1 is co-authored with Professor Knut Samset and based on more than ten years of research in the Concept Research Programme. It summarizes and discusses significant challenges and weaknesses in the front-end of public projects, based on literature and findings from a number of case studies of Norwegian projects, including an in-depth study of 23 major public projects (for details, see Whist & Christensen, 2011). The focus is on the effect of the analytic and decision-making processes during the front-end phase. Most of the project cases had not been subjected to external quality assurance under the Norwegian QA scheme, but were implemented before the scheme was introduced. It follows that the findings probably do not reflect the present-day situation with improved QA procedures.

The findings are framed as a set of ten paradoxes. The term "paradox" is used here to describe situations with a counterintuitive result, at least as seen from an overall societal perspective. It should be noted that some paradoxes are multifaceted, some are partly overlapping, and that a list of paradoxes can never be complete.

The paradoxes are:

1. **The success paradox: Success is measured in operational terms only, rather than the wider, strategic perspective.** Projects that are completed with considerable cost overrun and behind schedule generate negative media attention and even public inquiries, irrespective of their relevance and good value for money. By contrast, projects may avoid negative attention if completed on budget, regardless of their strategic success.
2. **The paradox of the significance of front-end management: Less resources are used up-front to identify the best conceptual solution (project governance), than to improve performance during implementation (project management).** The choice of conceptual solution often originates in the mind of an individual, based on intuition and experience, rather than systematic analysis of, for example, problems, needs, and requirements. By contrast, comprehensive planning and analysis is associated with the project once the choice of concept has been made.
3. **The paradox of early information overflow: Decisions are based on masses of detailed information rather than carefully selected facts and judgmental**

information relevant to highlight the essential issues. Spending resources in the front-end pays off. However, the crucial issue is not the volume or precision, but the type of information needed. The priority should be to establish an overall perspective based on a targeted search for information. Experience shows that large amounts of detailed data at an early stage may result in what is referred to as "analysis paralysis." Instead of opening up the opportunity space, it may lock decisions into an initially preferred concept.

4. **The paradox of the unexplored opportunity space: The choice of conceptual solution is made without systematically scrutinizing the opportunity space up-front**. There is much evidence to suggest that in many cases the chosen concept is not necessarily the most effective solution to the initiating problem. In many cases, the process started out with a predetermined solution, without exploring other options. This is referred to as path-dependency and illustrated in Figure 9.

5. **The paradox of strategic alignment: Strategy and alignment of objectives are highlighted as essential, but in many cases the internal logic of causality and probability of realization are erroneous**. Alignment of objectives is an exercise to define the causal link between the project outputs and the outcome and long-term benefits of the project. Unfortunately, such alignment is not always done. Objectives are missing or unclear, and there may be design faults at different levels, such as too many, overly ambitious and even conflicting goals.

6. **The cost estimation paradox: The focus is on the final cost estimate (the budget), while early cost estimates are overlooked**. Studies of cost estimation have found that initial estimates were substantially underestimated, whereas the final budget was more realistic. There is much to suggest that this may have resulted in the approval of projects that otherwise should have been rejected at an early stage. Nevertheless, project planners, decision-makers and researchers tend to focus primarily on the accuracy of the final budget, after the conceptual solution is decided.

7. **The paradox of disregarded analyses of costs and benefits: Detailed estimation of cost and benefits is commonly done up-front, but disregarded by decision-makers**. Substantial amounts of resources are devoted to cost-benefit analyses, especially for transport projects. However, the estimated value for money had no significant impact on the selection of projects in Norway. On the contrary, many unprofitable projects were realized. Obviously, decision-makers emphasize other aspects, but these are not included in the analyses.

8. **The paradox of** "**predict and provide": The tendency is to choose a** "**predict-and-provide" strategy rather than explore alternative solutions**. A variant of Paradox 4 is that in the case of congestion problems, need is often defined narrowly as a need to increase capacity. However, excess demand for public services and infrastructure should be expected when offered free-of-charge to citizens. In some cases, there may be goals for a different development. Project owners need to clarify the needs and goals that should apply to the project.

9. **The paradox of perverse incentives: Availability of public funding with no financial obligations for the beneficiaries may cause perverse incentives and result in counterproductive projects**. Different actors may have vested interests in certain projects being chosen, with no incentive to opt for the most socially beneficial or cost-effective alternative. This may result, *inter alia*, in supersized projects, positively biased business cases, and the selection of projects that turn out to be complete failures.
10. **The paradox of myopic decisions: Long-term viability is the intention but the planning horizon is too short, resulting in suboptimal choices**. The study of project appraisals shows that that needs and benefits are often assessed in a short-sighted and static perspective, trends are extrapolated without discussing alternative scenarios, and significant risk factors, such as political risk, are not identified and discussed. Such practices may lead to decisions that society will regret in the future.

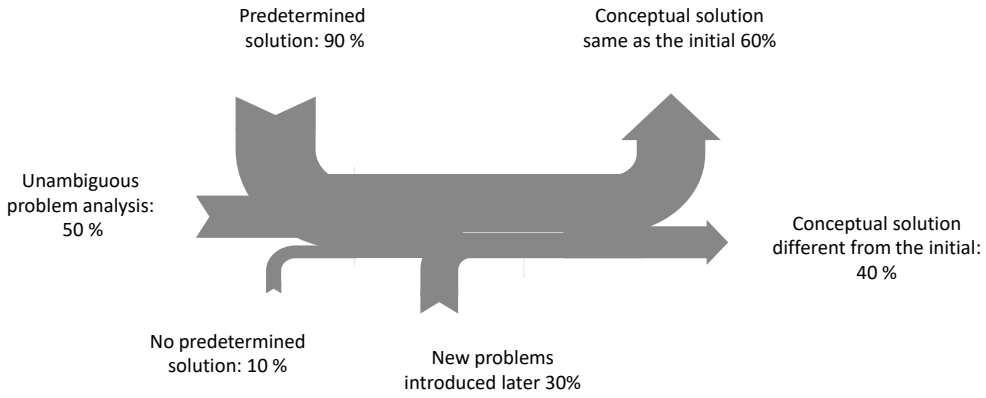


Figure 9 Path dependency up-front (based on Whist & Christensen, 2011).

The overall picture is that there are certain recurring deficiencies in both analytic and decision-making processes, and that the potential for improvement is huge. Moreover, it seems that flaws in both processes may be correlated, and further, that projects with many such deficiencies in the front-end phase tend to end up being less relevant to society.

4.2 Perverse incentives in the front-end of state-funded projects (Paper 2)

The second paper focuses more closely on Paradox 9, namely the adverse incentives that may occur on the part of the groups of beneficiaries when benefits are targeted and costs are born by the broader society. Since project approval and funding is discretionary, the alternative is for the intended beneficiary group to receive nothing. Hence, it will opt for the project as long as the gross benefits are non-negative, regardless of the cost. The problem is well-known from development aid, where projects' strategic success is often poor, and in some cases, it is not even clear whether they fulfill a genuine need.

The paper uses a simple conceptual framework based on principal–agent theory to explore the problem. In its simplest form, the model includes two parties: the national government (principal) and the group or community that receives the major portion of the benefits (agent). The problem is generated by the combination of

(1) misaligned interests, caused by the imbalanced distribution of costs and benefits
(2) asymmetric information: the beneficiaries know best their needs, as well as contextual factors affecting cost and risk.

A case-based research design with purposive sampling was chosen, with eight Norwegian projects from different sectors, and one development aid project that served as a reference case. All of the case projects were characterized by misaligned interests between central and local level, often with the municipality in a key role, and the information about local conditions was clearly asymmetric.

There were clear signs of agency problems in the cases. Early cost estimates were typically unrealistic, and the benefits to users and society were overestimated. Further, project initiatives that might have been sensible in the first place, grew out of proportion and became overdimensioned as a result of requirements introduced by the groups of beneficiaries after de facto approval of the project. Substantial resources were spent on lobbying. The government had surprisingly high confidence in the information provided by the initiating party. Only in one project did the government, admittedly very late in the process, demand an external review of the business case. In that particular project, the independent analyses came to an entirely different conclusion from the one commissioned by the group of beneficiaries.

Four of the projects were considered outright failures in strategic terms and should never have been approved, and three of them would probably not have been approved had it not been for the perverse incentives. The development aid project was an extreme case, wherein the gross benefits for the target group were negative. In the Norwegian projects, user values were generally positive, but often low. In one case, the group of beneficiaries accidentally ended up with the bill and was forced into bankruptcy.

The paper concludes with a discussion of relevant measures to prevent the emergence of perverse incentives. According to the principal–agent model, the solution is two-fold: (1) the beneficiary group's interests should be brought into alignment with the government's interests as much as possible, and/or (2) the information flow from the beneficiary group to the government should be improved in order to reduce the asymmetry. Conceivable measures of the former type would be to require co-financing or other incentives for cost control and accountability. The most high-powered incentive scheme would be to allocate a lump sum that the beneficiary group could allocate freely. Flyvbjerg et al. (2003) argue that the state should grant a general allocation (to the local administration or a state agency) and require that project selection meets certain objective criteria, such as value for money. Solutions to reduce asymmetry in information would be third-party reviews of project proposals, open hearings, and systematic ex post evaluations to learn about the level of impacts that can be expected.

4.3 The project proposal (Paper 3)

The third paper provides recommendations on how to develop a project proposal, in order to avoid the challenges discussed in the preceding section.

Projects ought to be defined and assessed in terms of their success in multiple dimensions. The use of a standardized set of success criteria might be expedient, such as the five OECD DAC criteria: *efficiency*, *effectiveness*, *impact*, *relevance*, and *sustainability* (OECD, 2002). It is argued in the paper that in the earliest phase, strategic performance should be the main issue, and particular attention should be paid to the choice of concept.

Ideally, a project proposal should include the following steps and elements, as illustrated in a simplified way in Figure 10:

- A needs analysis, mapping all stakeholders and affected parties, and assessing the project's relevance in relation to needs and priorities in society. Needs ought to be expressed in sufficiently general terms to give latitude for alternative solutions.
- A specification of overall requirements that need to be fulfilled when the project is implemented (e.g., functional, operational and economic).
- An overall strategy defining the project's goal and purpose (first-order and long-term effects), with emphasis on consistency, realism and verifiability.
- Specification of alternative solutions (concepts) to realize the strategy. The opportunity space (or outcome space in Figure 10) is usually larger than envisioned.
- An alternatives analysis involving at least two alternative concepts and the zero option (doing nothing). The analysis needs to evaluate the alternatives with emphasis on relevance, sustainability and value for money.
- Ranking the proposed alternatives and providing recommendations regarding the decision strategy and implementation strategy for the investment case.

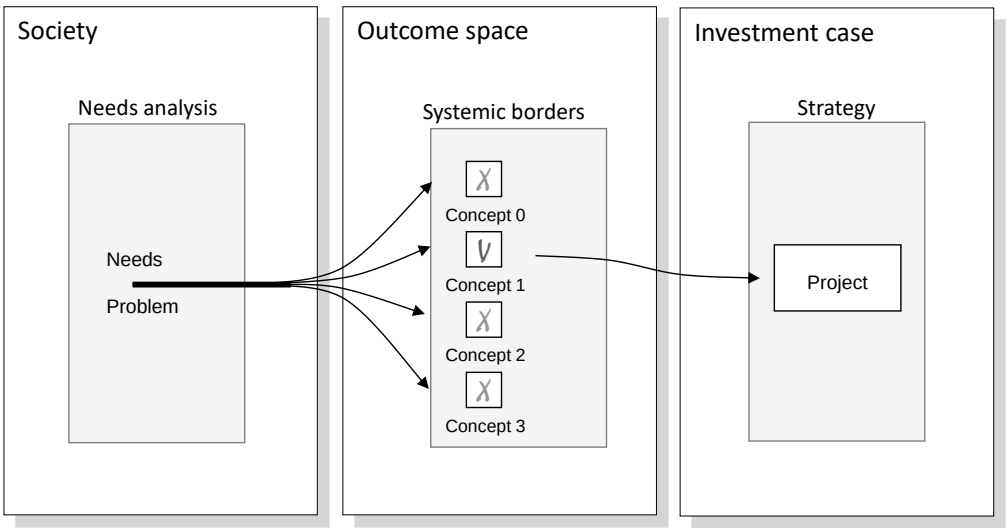


Figure 10 Assessment of alternative concepts (from paper no. 3)

The above-listed initial analyses, which are broad and often qualitative, have been proved crucial. Premature lock-in to a preconceived less appropriate conceptual solution can pose a major risk to project success. Therefore, a comparatively larger share of resources needs to be expended to develop the conceptual solution up-front, before detailed plans and forecasts are prepared for the chosen concept. A challenge may be that availability of information is limited at this stage. However, the early assessments need not require large amounts of detailed and accurate data, but rather they capture, messages of, for example, common sense, knowledge of user needs, markets, and laws and regulations permit, and a rough but realistic cost estimate (typically based on experience from similar completed projects). More detailed considerations relating to project design are better left to project management, who will be allowed a certain amount of flexibility within the strategic framework outlined, to operationalize what has been approved.

A necessary part of every project proposal is a plan for how to meet the project's financial obligations during its life cycle. An important message from Paper 3 is that decisions relating to financing may have an impact on the project's strategic performance, through their effect on relevance, sustainability (i.e., financial) and value for money.

First, as discussed in Paper 2, cost sharing between government and beneficiaries may be a solution to problems of local agents' perverse incentives, the argument being that agents will not promote poor and unviable projects if they have to bear the costs. In order to have the desired effect, such a co-funding requirement ought to be announced in advance and be credible.

Second, funding itself may involve either a cost or efficiency loss, which may be considerable and should be included in an assessment of the investment's value for money. Whether funded locally or centrally, the ultimate choice is between general tax collection and user fees. The collection of taxes creates an efficiency loss in the economy, which in Norway

is estimated as approximately 20%. User fees will normally reduce demand to a suboptimal level and thereby reduce project benefits ex post. In addition, there are administrative costs of a separate collection system. However, in the case of negative external effects of consumption, user fees may enhance efficiency. The question is therefore which option is least inefficient in the specific context. In Norway, this means that if the costs of user fees could be kept below 20% of the necessary funding, user fees could be a feasible solution.

Third, the project's relevance and sustainability depends heavily on it being *accepted* throughout society. This includes the sharing of the financial burden. Therefore, also the distributional impacts of different financing mechanisms need to be considered as part of the alternatives analysis.

Fourth, an option for public projects is to bring in *private capital* of some form. This is often expected to improve project flexibility and thereby improve the project's cost and time efficiency. An experience from such contracts is that they work best when the entire project life cycle is taken into account in the contract ("life cycle models") and when the transfer of risk to the private party is real and credible. However, ultimately, a public project must be financed either by taxes or by user fees, as discussed above.

4.4 Discussion

The combined findings from Papers 1–3 can be applied to answer RQ1: *What are the main challenges in front-end management of projects?* Whereas Paper 1 takes a broad view on challenges and weaknesses in the front-end, Paper 2 explores one of the problems in more depth, and Paper 3 suggests a number of measures necessary to master the process of front-end management, while at the same time they may be used to demonstrate the gap between theory and practice.

Several key themes run throughout, which overlap to varying extent. Most importantly, there is a tendency to ignore the crucial assessment of problems, needs, opportunity space, and the choice of concept, and instead jump to the more detailed, and often quantitative and data-intensive analyses of one specific conceptual solution.

There are numerous explanations for such weaknesses, and different explanations may overlap. Planners, who are often engineers and economists, may be reluctant to question fundamental issues that can be seen as what they consider part of the political sphere. I have heard planners say "We are producing analyses, not guesswork," implying that they are more comfortable working with tangible measures and clearly defined tools and methods, than with multidimensional and qualitative assessments of success criteria that may be unclear and not even agreed. There may also be cognitive shortcoming to planners' ability to apply creative thinking and long-term perspectives, as well as their understanding of fundamental uncertainties. Another quite likely explanation is that project initiators (who often commission the analyses) see it as in their interest to explore one specific conceptual solution, and restrict the terms of reference accordingly, or, even worse, they do not demand an early project appraisal at all.

Perverse incentives on the part of project initiators have been identified as one of the challenges, with potentially adverse consequences. The absence of liabilities, such as co-funding, is a simple and obvious risk factor of which project funders should be aware. This is not to say that the paradox of perverse incentives occurs more frequently than the other nine paradoxes. I have chosen to pay special attention to this paradox because it is largely unexplored in the literature, and I thought it would be interesting to test a simple principal–agent model on the Norwegian data. By contrast, many of the other paradoxes have been thoroughly explored in the literature, and in separate studies under the auspices of the Concept Research Programme.

Principal–agent theory has definitely inspired the project management and project governance literature, but has primarily been used in discussions of how to motivate project managers and contractors, and less in relation to the front-end phase. Paper 2 demonstrates that the model is well-suited for assessing the risk of perverse incentives at the front-end, as measured by the combination of clearly misaligned interests and asymmetric information. In line with, for example, Eisenhardt (1989), I do not claim that the principal–agent theory is sufficient to understand front-end challenges in public projects, nor that it should be used alone. However, it provides a useful new perspective on project initiation that ought not to be ignored.

Still, perverse incentives may be difficult to prove, and their effects, such as a biased business case, could also be the result of overoptimism stemming from cognitive weaknesses. However, countermeasures to deal with perverse incentives could be useful in the case of cognitive bias. In particular, reviews by independent experts would be helpful in both cases.

It should be noted that most of the projects studied in Papers 1 and 2 were selected because they were assumed to be extreme cases, and not necessarily representative of current projects. Nevertheless, I think there is good reason to believe that the findings are relevant to the larger group of public projects in Norway before the QA scheme was introduced. Many of the same challenges and weaknesses in front-end management have been confirmed in other Norwegian studies, most notably the government-initiated study that preceded the establishment of the QA scheme (Berg et al., 1999; Whist & Christensen, 2011). Furthermore, my own observations and informal discussions with a number of project experts indicated that the cases were not necessarily extreme, and many of the challenges in the cases were observed also in other public projects.

Whether the findings are relevant to other countries is less clear. In Norway, the local democracy stands strong, while at the same time the local government is financially weak and dependent on the state to finance local infrastructure. This may have given rise to serious problems with adverse incentives on the part of local initiators. However, Norway is not unique in this respect. It is not uncommon to see examples of, for example, roads, hospitals, universities, and sporting events, that are largely “local public goods,” that are being funded at the national level. Generally, the findings presented here support many of those found in the international literature, as presented in Chapter 2, such those by as Flyvbjerg et al. (2003).

From a comparison of theory presented in Paper 3 with practice presented in Papers 1 and 2, it seems that some key points for improvement are as follows:

- The business case should be presented to decision-makers early enough to prevent premature lock-in to an unjustified concept.
- Incentives for project initiators ought to be brought in line with society's interests as much as possible. There is a need to deal with adverse incentives relating to discretionary assessment and approval processes.
- Analyses should be transparent and overseen by independent experts.

All three questions of *Why*, *What* and *How*, should be properly addresses. This implies that the underlying problem or need (*Why*) should be explicitly analyzed and justified, and not simply taken for granted. Further, the opportunity space should be fully explored, and alternatives assessed with respect to society's strategic success criteria (*What*). Funding should be considered part of the conceptual solution, to the extent that it may affect strategic project success. Only after these issues have been dealt with, should the *How* question be addressed.

The introduction of a technocratic project governance framework is rarely desirable, but at least we need measures to avoid an anarchist type of approach (see Figure 6 in Chapter 2). An apparently reasonable compromise in front-end assessment and assurance of major projects would be for the first step to identify and eliminate the worst alternatives (the *What nots*). These are "low hanging fruits" and proper action can give a high reward with little effort. The next step should be to seek for feasible alternative concepts, but within reasonable limits, and not necessarily crave for the best, since the case will nevertheless be handed over to decision-makers in the political domain to conclude.

5 Project governance frameworks (RQ2)

This chapter presents Papers 4–6, which address the second research question: *How can project success be achieved through project governance frameworks?* The reference here is the Norwegian QA scheme. Paper 4 presents lessons learned from the scheme 15 years after QA2 was introduced, and 10 years after QA1. Paper 5 compares the Norwegian framework with frameworks applied in five other countries in recent years. Paper 6 discusses the relationship between a project governance framework introduced by the Cabinet, and frameworks at lower levels in the public administration. All three papers are co-authored with fellow researchers.

5.1 Quality assurance in megaproject management—the Norwegian way (Paper 4)

Paper 4 presents the Norwegian governance framework for public megaprojects and lessons learned as of 2015. The framework involves external quality assurance of key decision documents at two gateways: (1) before the Cabinet's choice of conceptual solution (QA1) and (2) before Parliamentary approval and appropriation of funds (QA2).

QA2 was introduced in the year 2000 and involves quality assurance of the cost estimate and management base. The aim was to ensure operational project success. After 15 years, nearly 200 projects had been subjected to QA2 reviews, almost 70 of them had been completed and in their operational phase. The data show that the majority (79%) were completed within or below the cost frame. With cost frames largely in line with what was recommended by the quality assurers (i.e., the P85 estimates), the result is quite close to what should be expected statistically.

Similarly, 48% of the projects were completed within or below the target cost (commonly set at P50). These results suggest that at the portfolio level, the Norwegian government is now effectively controlling costs in major investment projects. Moreover, there is no indication that contingency reserves are spent unnecessarily. Considering these results, the QA2 exercise seems to have helped to improve cost management and ensure cost control in major public projects in Norway. However, it should be noted that the first set of completed projects may be somewhat biased in the sense that they were implemented in a time-efficient manner, possibly because they were rather uncomplicated. Further, caution should be exercised when comparing the results with the situation in the 1990s and in other countries, where cost estimates may be set at an earlier stage in the project life cycle. One of the features of QA2 is that projects are now more mature when they are approved by Parliament, and that cost estimates are more realistic.

The framework was expanded in 2005 to include quality assurance of the choice of conceptual solution (QA1). The paper also discusses contemporary experiences based on about 70 projects that have so far been subject to a Conceptual Appraisal (CA), followed by an external QA1 review. None of these projects have been finalized to date. One noticeable feature of the CA/QA1 scheme is that it provides a more systematic approach to the early

assessment of project ideas than before. Rather than opting directly for the technical solution, planners now have to apply a broader perspective and explore the societal aspects of the proposed project. Thus, the opportunity space has been broadened during the appraisal process. Interesting debates have been triggered regarding, *inter alia*, what defines a conceptual solution. Should different technical solutions to the same problem be accepted in the analysis or should there be different conceptual solutions in the broadest sense? Whatever the answer, since the QA1 scheme has introduced this issue on the public agenda, it has had a noticeable effect on analysts, politicians and the public. There is little doubt that the quality of the CA reports has improved steadily over time, and that there is convergence towards a common sound practice. The same trend can be observed in the QA1 reports: quality assurers have gained years of experience and shown a positive learning curve.

Overall, the experiences are positive and the outcome thus far is promising. The governance framework is institutionalized in central government, and ministries and agencies responsible for megaprojects have gained considerable experience from the QA processes. The QA1 scheme was controversial in the beginning, and some feared that political decisions could be dominated by the involvement of private consultants. However, today, most ministries and agencies recognize the benefits of the scheme. The choice of conceptual solution is a political decision to be made by the Cabinet, whereas the consultants' role is restricted to asserting the quality of the documents' supporting decisions. Furthermore, ministries, as well as the Cabinet, are now encouraged to have a more direct influence in the early stages of the process.

A review of QA1 reports that has been completed to date (about 70), demonstrates that quality assurers and the sectoral ministry agree on the ranking of concepts in only one-third of the cases. Not surprisingly, in these cases, the Cabinet commonly endorses the recommended conceptual solution. In the remaining two-thirds of the cases, the quality assurers often recommend the zero alternative or a more economically feasible concept. In such cases, the outcome is less predictable. The Cabinet more often follows the recommendation made by the sectoral ministry than the QA1 recommendation. In other cases, the project proposal is either sent back for a revised CA appraisal or is withdrawn; alternatively, the Cabinet chooses the zero option or a completely different conceptual solution.

The results indicate that the Cabinet is now better informed and advised by the external quality assurers, but also that the final choice of project concept largely remains a political decision. However, over time, it may become more difficult for the government to choose conceptual solutions that are clearly ineffective or poor value for money. Also, there is reason to believe (although difficult to prove) that some of the most poorly conceived investment proposals are now screened out before they even reach the CA/QA1 stage.

Indirectly, one can also infer some of the spinoff effects in government, industry and academia following the introduction of the scheme. We find a clear trend of improved project practices in central government since year 2000, as discussed further in Paper 6. Generally, there is growing awareness in government regarding the need to improve the quality of

decision documents in general, and broaden the scope of analyses to include alternative concepts. Sectors not subjected to the QA scheme, notably health authorities, electric utilities and the municipal sector, have voluntarily introduced modifications of the Ministry of Finance's scheme. We can also see that front-end management has become an issue within the community of professionals in project management, and training courses are now being offered by a number of institutions, agencies and consultants.

The paper also discusses some unresolved challenges. First, there is still room for improvement regarding how the opportunity space is defined and explored in the CA reports. A recurrent problem is that the conceptual solution is de facto selected before the CA process, either because of path dependency in the agencies or due to political constraints and limitations. This may indicate that CA/QA1 sometimes comes too late, or that the front-end phase is subjected to a series of restrictions.

Second, it has been claimed by ministries and agencies that quality assurers give disproportionate attention to economic considerations and that economic impacts ought to be balanced against the achievement of various political goals.

Third, there is a concern that increases in the project content, scope and cost take place during the process that follows QA1. In the current version of the scheme, there is no obligation for ministries and agencies to consider the cost estimate at QA1 as binding for the subsequent pre-project phase. If the cost estimate increases considerably between QA1 and QA2, the assessment of the project's tactical and strategic success from the QA1 stage may no longer be valid. To date, only eleven projects have undergone both QA1 and QA2, and researchers should focus on increases in the cost estimate between the two control points.

5.2 Governance frameworks for major public projects in six countries (Paper 5)

Paper 5 compares the Norwegian QA scheme with similar schemes introduced after the turn of the new millennium in five other OECD countries: the UK, Denmark, the Netherlands, Canada (Quebec Province), and Sweden. The backdrop in all cases was negative experiences from past projects, especially with regard to cost overruns and delays.

Common to all schemes is that they are intended to improve project governance by central government and are applied to projects that involve particularly high cost, risk, complexity, or highly innovative solutions. They are applicable to some or all sectors where the state is responsible for infrastructure projects. Three countries (including Norway) have introduced a general threshold value defining which projects should be involved, whereas the other three decide whether to include projects on a case-by-case basis.

The paper focuses merely on the structural elements of project governance. The reference point has been a scheme adhering to the recommendations from literature, including both the overall principles of good governance and the more specific recommendations concerning project governance schemes.

The six schemes have many common characteristics. They all apply a stage-gate model at project level, including independent quality assurance reviews of project documentation at specified decision points, and placing key decisions, as well as the responsibility for managing the scheme, at a high level in the system (see Figure 11). There are also a number of differences, for example, with regard to the parties involved and their roles, comprehensiveness, flexibility, and organization.

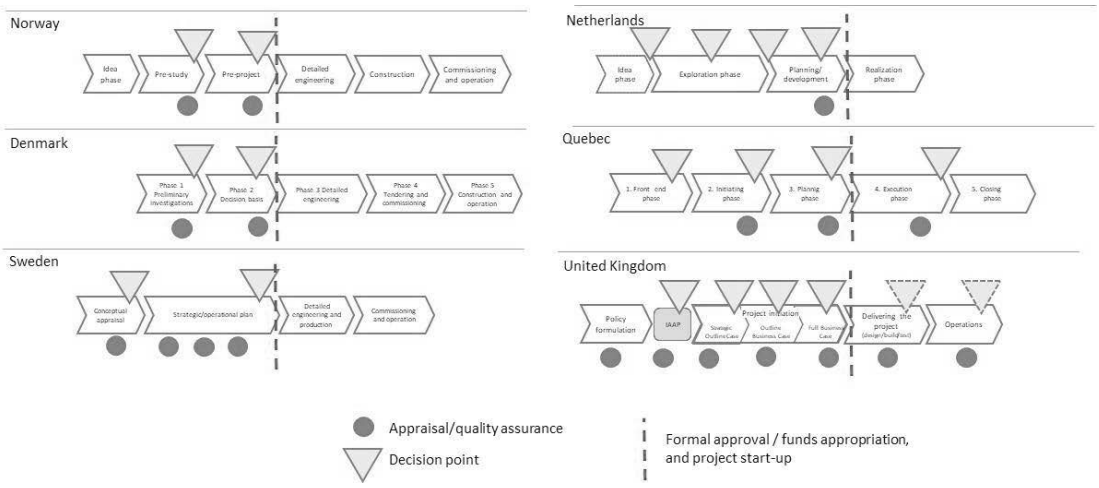


Figure 11 Stage-gate models in six countries (variant of Figure 6 in Paper 5)

As far as the number of decision points is concerned, Denmark and Norway have the simplest schemes, with only two decision points. However, the QA reviews in Norway are rather time-consuming, inasmuch as the quality assurer is required to perform his or her own independent analyses, and not only scrutinize the work that has been done. The UK has the largest number of decision gates, which are not limited to the front-end, but include follow-up points during project implementation and closure. However, the reviews are simpler, and the model is also flexible with respect to the number of intervention points and their scopes.

The government plays a key role as decision-maker in all six countries, primarily with regard to the choice of project alternative. In the Scandinavian countries, the final approval is elevated all the way to the parliamentary level. Presumably, this has to do with these countries having relatively small populations, but also because they normally have minority governments and thus need support at the parliamentary level. The quality assurance function is performed by parties independent of those who conduct the appraisals. In Norway and Denmark, external experts from the private sector are employed, whereas in the other countries designated public bodies are used, or QA is performed internally within government agencies.

It is interesting to observe that several countries have expanded and strengthened their schemes over time. Generally, the purpose of most of the schemes initially related to the

operational aspects of project success. Later, a somewhat broader perspective, including the choice of conceptual solution, was adopted. In recent years, some of the other five countries have introduced a formal decision gate at an even earlier stage than QA1 in Norway, which is interesting, given the consideration that QA1 comes too late in some cases.

As far as cost control is concerned, a key element of the Norwegian governance scheme has been the introduction of a cost frame and a specified, lower target cost for the agency. The other five countries do not seem to apply contingency reserves at project level, but they may add a national supplement. Further, Norway and Sweden are alone in using probability-based estimation. In the UK, an uncertainty level is chosen for each case (e.g., P50 if central government is willing to assume a high risk of cost overruns or if the project forms part of a large portfolio, otherwise higher), and optimism bias correction factors are used, based on rules of thumb tailored to the chosen uncertainty level.

Overall, the schemes seem to be fairly consistent with the recommendations from the literature. One exception is that, in general, the schemes do not require co-funding from those who will benefit from the projects. This is done only in the Netherlands. Further, only Norway highlights transparency at the project level; all QA reports, as well as final costs and other project results, are published online. In the other countries, project results may be published at the group level. Furthermore, there is potential for improvement in several countries with regard to integrating the portfolio perspective. The Norwegian governance scheme focuses on requirements applicable to individual projects only. In the UK, a central government unit is responsible for compiling data on all infrastructure projects in the portfolio, thus making it possible to analyze and manage them collectively.

All schemes are relatively recent, and therefore it is too early to determine with any certainty their impact and degree of success. The ultimate question is whether some schemes are more effective than others in improving project delivery, as well as outcome, and to what extent an effective scheme in one country can be adopted by other countries. This will be a topic for future research.

5.3 The hierarchy of public project governance frameworks (Paper 6)

The Norwegian QA scheme constitutes the “top layer” of a hierarchy of governance measures, and applies only to the country’s largest projects. Clearly, the success of public projects depends profoundly on what happens at the lower levels, where projects are planned and implemented. The aim in Paper 6 is to comprehend public project governance as a hierarchical system, as well as the relationships between project owners at three levels, namely cabinet, ministry and agency, as illustrated in Figure 12.

It is implicitly assumed that the QA scheme is supplemented by more specific governance arrangements in ministries and agencies. By demanding high quality from the top level, a trickle-down effect could be expected, also in the form of higher standards at lower levels.

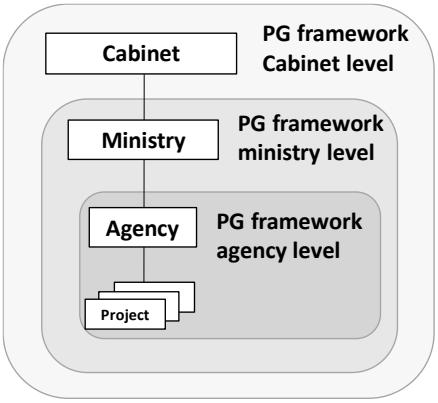


Figure 12 The hierarchy of public project governance frameworks

In the empirical study reported in Paper 6, I and my co-author investigated project governance arrangements at ministry and agency levels in seven government agencies and seven ministries in Norway, covering the sectors with most of the projects under the QA scheme. As in Papers 4 and 5, the focus was primarily on the structural aspects of project governance.

The findings show that the studied sectors had invested heavily in building project competence and capacity since the turn of the millennium. All agencies and one ministry had introduced stage-gate models with defined phases, decision points, roles and responsibilities, and independent (though rarely external) quality assurance. Further, the portfolio perspective was prominent in all agencies. The models seem to have become increasingly comprehensive over time, with associated guidelines, templates, and procedures.

The governance frameworks mainly adhere to the recommendations from project governance literature. However, most of them do not cover the earliest phase, where the initial idea occurs and is developed into a conceptual solution. Some of the interviewees referred to this phase as the political or strategic phase, which is not considered part of the project. Only the Ministry of Defence has come up with clear instruction as to how project ideas should be handled and endorsed before transferring them to the conceptual phase. According to our interviewees, the project ideas occur in various ways. In some cases, they are identified at the political level, or follow from a new policy or reform. However, equally often, an idea "occurs" at agency level, sometimes in close dialogue with internal or external user groups, and based on more or less systematic needs assessments.

The project sponsor role is commonly used, but is often delegated to an individual at a fairly low level in the organization. Another observation is that the project sponsor is often appointed rather late in the project process, while not having a role in the front-end phase, and sometimes not in the operational and maintenance phase. Project boards are widely used, but more often than not they act not as boards, but rather advisory groups, and they too are established late in the project process, after the decisive choice of concept has been made.

A striking finding is that that with one exception (the Ministry of Defence), the agency level takes the most active role as project owner and initiator of governance arrangements. Today, all the agencies are project-based organizations. By contrast, the ministries do not seem to relate to their subsidiary agencies in terms of their projects, programs and portfolios. The exception is the largest projects that undergo the Cabinet's QA scheme, for which the ministries see themselves as responsible for ensuring that the requirements set by the Cabinet are met. Some ministries also routinely consider whether the Cabinet's QA scheme should be applied to projects just below the QA threshold. In addition, ministries govern projects in informal ways to the extent that they find it necessary, typically depending on the project's scope, complexity and political risk.

We have not found any readily apparent inconsistencies between governance arrangements at the three levels of the hierarchy. Particularly, the schemes at the lower levels are well adapted to the Cabinet's scheme and ensure that the largest projects are well prepared for the two control points, QA1 and QA2. However, with the Ministry of Defence as the exception, the ministries' formal involvement is limited in most cases. Governance tasks thus seem to be extensively delegated to the subordinate agencies. Furthermore, governance models at agency level have a somewhat narrow and operational focus. Generally, these findings indicate that there is a risk that governance tasks will be handled from an internal, technical or administrative perspective, rather than representing user groups and the broader society.

5.4 Discussion

Taken together, Papers 4–6 elucidate RQ2: *How can project success be achieved through project governance frameworks?*

The three papers presented in this chapter are empirical papers that discuss the features of recent governance frameworks introduced to improve public investment projects. They discuss the extent to which the frameworks are on in accordance with recommendations from literature, and one paper (Paper 5) explores similarities and differences between countries in terms of their characteristics. Paper 6 introduces the term "project governance hierarchy," arguing that public project governance frameworks may involve several administrative layers that function jointly to improve the probability for successful projects. Both Papers 4 and 6 discuss experiences and preliminary effects relating to the Norwegian QA scheme. Some remaining problems and weaknesses are also highlighted, which indicate that the paradoxes discussed in Chapter 4 may not be easy to overcome.

It should be noted that the scope of the research is the Norwegian QA scheme (with the exception of Paper 5). Thus, caution should be exercised when transferring the findings to other countries and contexts.

The governance frameworks are well adapted to recommendations from literature, and frameworks in different countries have much in common

The Norwegian scheme is a top-down regulatory regime that requires two external quality assurance interventions before the project can be approved. It aims to ensure that the need for the project is genuine, the opportunity space is explored, the choice between alternative project concepts is subject to political assertion and decision, and that the project is mature and well-defined when presented to Parliament. Openness and transparency is an essential part of the scheme. With a few exceptions, most notably the lack of a portfolio perspective and the lack of requirement for risk capital from those who initiate and benefit from the project, the scheme is in line with literature recommendations (e.g., Müller, 2009; Narayanan & DeFillippi, 2012). It is essentially a simple framework, based on an approach of limited interference in existing practices and procedures in ministries and agencies.

The comparative study of project governance frameworks in six OECD countries presented in Paper 5 revealed many similarities between the frameworks in the studied countries. Thus, there seems to be a common conception internationally about what constitutes a suitable project governance framework. To a large extent, the studied schemes require the same types of assessments at each intervention point, they require independent quality assurance, and that decisions are elevated to the highest level in the political system. All schemes were originally established to ensure cost control (the *How* question), but have since expanded their perspective over time to focus increasingly on the front-end and the choice of concept (*Why* and *What*). This is in line with the development in literature and the perspective of the project management community as discussed by, for example, Williams et al. (2019). Some of the countries have introduced intervention points at an even earlier stage than QA1 in Norway.

A notable difference between the schemes is that the schemes in the Scandinavian countries are relatively simple in terms of the number of intervention points, and that they do not intervene significantly in existing processes and practices. The schemes in the other three countries are more ambitious, extensive and behavior-oriented, with more follow-up points, also during the implementation phase. It is not yet possible to determine whether one approach is more effective than the others. The fact that there are several different governance schemes in operation is positive, and they might inspire alternative ways of organizing and implementing such schemes in the future.

The introduction of a simple framework at the top level may generate advances at lower levels of the project governance hierarchy

The Norwegian QA scheme constitutes the "top layer" of a hierarchy of governance arrangements, and applies only to the very largest projects. Paper 6 explores public project governance as a hierarchical system, and the relationship between project owners at three levels: cabinet, ministry and agency. We demonstrate that the simple model at the top level has *been* matched by more detailed requirements, guidelines, and training at the lower levels, and has facilitated the appearance of self-regulatory schemes. In light of the general "projectification" of society, it is difficult to determine whether the improvements at lower

levels are caused by the QA scheme itself, but at least it seems to be an essential triggering factor. Especially, the use of external quality assurance and the principle of transparency seem to have had a disciplining effect on the subordinate levels.

As noted by Klakegg and Volden (2017), Norway is characterized by a consistent democratic tradition, an egalitarian society, and a high level of education, which makes a strong foundation for delegating authority downwards in the hierarchy. Despite being rather simple, the QA scheme was initially somewhat controversial, as it entailed the involvement of external parties in matters handled by experts in the ministries and agencies. A more behavior-oriented scheme that also interfered in the more detailed project procedures and practices in the agencies would probably have been a lot more controversial.

The Norwegian scheme seems to promote operational project success. It is too early to determine its impact on tactical and strategic success.

The empirical evidence shows that projects under the QA2 scheme have been quite successful in operational terms. In particular, the majority of projects have kept within their cost frames as defined on the basis of QA2. This is confirmed by Odeck et al. (2015), who studied cost overruns in road projects before and after the introduction of QA2. It may not be possible to prove the relative importance of the various elements of the QA scheme in explaining these results. Most likely, it is a combination of the requirement to use stochastic cost-estimation methodology, the presence of external QA, and the use of a lower target cost for the agency.

The favorable results concerning cost compliance should not be confused with evidence that the selected projects represent tactically and strategically successful projects. This is addressed in QA1, the effects of which cannot yet be proved. We do know that planners are now obliged to take a broader perspective on needs, opportunities, and social costs and benefits in projects' front-end, and that the quality of the CA reports (the business case) has improved over time. The disciplining effect of external reviews, leading to more realistic estimates of both costs and benefits is discussed also in Paper 7. However, we also observe that the recommendations made by ministries and agencies (CA reports) and those made by the quality assurers (QA1 reports) often diverge, and that decision-makers tend to follow the advice from the ministries and agencies more often than they follow the advice of external consultants. A follow-up question, which is also discussed in Paper 7, is whether the QA1 reports take a different perspective (from the CAs) which is considered less relevant to decision-makers.

We certainly need more data on the effects of the Norwegian scheme, as well as those of other project governance frameworks, on project success at tactical and strategic levels. This ought to be an important topic for further research.

There are still signs of some of the paradoxes

Papers 4 and 6 point to some weaknesses of the Norwegian scheme that ought to be followed up. One weakness is the risk that the opportunity space might not be effectively defined and explored, and especially that low-cost solutions are often not included among the alternatives

assessed. A possible explanation may be that CA/QA1 comes too late and the project has developed too far when it reaches the Cabinet. In studying the lower levels of project governance hierarchies, we find that the agencies' own frameworks often take a narrow, operational perspective and that they ignore the crucial front-end phase. Even if a government agency should have a procedure for how to assess and prioritize early project ideas, it will in effect take a more limited perspective than will the parent ministry and the Cabinet. Paper 6 concludes that there is a potential for improvement concerning owner involvement, particularly in the earliest stages. In light of this, it is interesting that some other countries have introduced an intervention point at the idea stage, as noted in Paper 5. A study of lessons from such "very early" intervention points would be very desirable.

Another weakness noted by some informants is that QA1 focuses narrowly on the cost-benefit analysis, implying a risk that the recommendations are not considered relevant and are thus ignored by decision-makers. As shown in Figure 4 in Chapter 2, strategic project success has several aspects to it, such as sustainability, needs satisfaction and economic viability. This topic is discussed further in the next chapter.

A third concern is the indication that projects enter the pre-project phase without a clear commitment to the original conceptual solution. QA2 focuses on the cost estimate and the project management base, not on the balancing of benefits and costs, thus implicitly assuming that this is taken care of at the previous stage. However, in reality, many decisions are made throughout the pre-project phase that affect both sides of the equation. The tendency is for costs to increase between QA1 and QA2, but we know little about the development in benefits. Regardless, projects are normally approved by Parliament after QA2, possibly due to what Cantarelli et al. (2012) refer to as lock-in. This brings us into the more general issue that concerns a need for benefits management (or value management) throughout the whole project life cycle. Ensuring strategic project success requires more than a sound choice of conceptual solution made at some point in time. Perhaps the Norwegian scheme, which focuses merely on the front-end phase, is too simple after all? This is too early to ascertain, but these issues ought to be followed up as more projects go through both control points and are completed.

In sum, we realize that some of the paradoxes listed in Section 4.1 may still be present after the introduction of the Norwegian QA scheme. Much has been done to highlight the front-end and the choice of project concept, and to reduce the risk of perverse incentives. However, the research undertaken for this thesis suggests that several of the paradoxes may be difficult to overcome and that would require more effort.

6. Appraisal and evaluation methods (RQ3)

The third and last research question was formulated to explore the term "project success" with particular focus on the tactical and strategic levels: *How can project success be achieved through improved appraisal and evaluation methods?* As noted in Paper 4, it has been claimed that the QA1 scheme gives disproportionate attention to economic considerations, and that a broader set of success criteria might be more relevant to decision-makers. The choice of appraisal and evaluation methodology applied in the governance scheme is therefore essential; thus, RQ3 is closely related to RQ2.

Paper 7 examines how cost-benefit analysis is applied in 58 case projects, and maintains that a broader set of criteria is needed. Paper 8 explores the term "sustainability," and offers a conceptual framework for understanding it across disciplines. Finally, Paper 9 recommends and tests a compound evaluation framework of six criteria.

6.1 Assessing public projects' value for money: the usefulness of cost-benefit analyses (Paper 7)

The cost-benefit analysis (CBA) concerns the relationship between resources invested and benefits achieved (i.e., the project's value for money). The project management literature and professional project management bodies consistently highlight the use of CBA as basis for project selection. However, less attention has been paid to the quality and utility of CBAs as seen from the decision-makers' perspective.

Based on the wide, but fragmented literature on weaknesses and challenges in CBAs, Paper 7 offers three broad explanations for why CBAs may not be considered useful by decision-makers (see Figure 13):

1. *The normative underpinning* is considered too narrow. The CBA recognizes people's preferences only in their role as consumers, and is a framework for measuring efficiency, not equity or goal achievement.
2. *Uncertainties and measurement problems* are common, such as omitted impacts, forecasting errors and valuation errors, which causes inaccurate and unreliable results.
3. *Appraisal optimism exists*, in the sense that not only are CBAs inaccurate, but also they may be biased on the optimistic side. The complexity and often limited transparency of the methodology may reinforce problems of low trust and confidence.

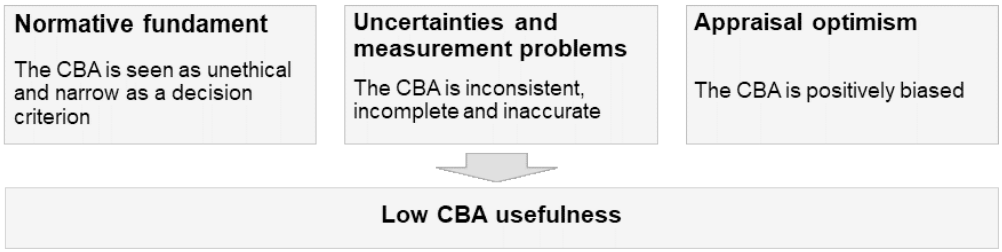


Figure 13 Three types of weaknesses that lead to restricted usefulness of CBA

Furthermore, Paper 7 presents an empirical study, which demonstrates the weaknesses described above, and is based on CA and QA1 reports from 58 Norwegian projects, and the impact of CBAs' usefulness in decision-making.

The findings are quite promising with regard to the monetized benefits and costs, and indicate that the second and third types of weaknesses have largely been avoided. The studied CBAs are relatively comprehensive and consistent, uncertainties are well handled, especially regarding the capital cost, and the level of transparency is acceptable. The estimated net present value (NPV) of the recommended alternative is not systematically more positive in the CAs than in the QA1 reports, which indicates that the NPV is not manipulated by project initiators to make projects appear more profitable. The fact that an external review will be performed seems to have had a disciplining effect on ministries and agencies.

However, the part in the CBAs concerning non-monetized impacts demonstrates serious weaknesses. Non-monetized benefits are often essential for the final verdict of projects' value for money. Nevertheless, I find that such benefits are commonly inconsistently interpreted, assessed, or communicated. Quality assurers often accuse the CAs of exaggerating non-monetized benefits or including elements that are not true economic benefits. This in turn is related to the first type of weakness (i.e., value for money being too narrow as decision criterion), which has not been resolved in a satisfactory manner in the studied CBAs.

Decision-makers consider CBA a vital part of the business case for major public projects. This was found through direct measurement (interviews) as well as indirectly (revealed adherence to recommendations). They value the quality and credibility of CBAs, and find the QA1 intervention useful in determining the quality of the CAs. However, they also prefer a decision base that includes more than the assessment of value for money. Unfortunately, as noted above, these other considerations are often incorrectly referred to as non-monetized impacts, rendering it unclear what has been measured (i.e., whether value for money or some other confounded criterion).

The paper offers a set of practical recommendations to increase CBA usefulness:

1. Relevant perspectives beyond value for money (e.g., equity, sustainability and goal achievement) should be presented as part of a comprehensive business case, but

clearly separated from the CBA. Any conflicts between the perspectives should be identified, but the final balancing between them ought to be left to decision-makers.

2. Alternative solutions to the problem at hand should be assessed; in other words, the opportunity space ought to be explored. Project promoters may not have the right incentives to include simple and low-cost solutions among the alternatives.
3. Completeness and consistency are important quality requirements in CBA, which comprise, for example, impact categories included, the extent to which impacts are monetized, and the choice of parameters. Although all projects are unique, the findings from the research for Paper 7 indicate that there is room for more standardization.
4. Uncertainties need to be identified and presented as part of the CBA, to the extent that they are likely to affect the ranking and recommendations.
5. Non-monetized impacts may be as relevant as monetized impacts. They should not be ignored, overvalued or mixed with other perspectives than the value for money perspective. Methodological improvements, as well as guidelines for assessing non-monetized impacts, are required.
6. Measures to prevent optimism bias on the part of project promoters are recommended. Transparency and external quality assurance worked well in the studied projects.
7. Analysts' professional competence and qualifications in CBA methodology are key.
8. Understandability and communication are important aspects of transparency in reports, and relevant to decision-makers who are not CBA experts.

6.2 On the concept of sustainability (Paper 8)

Paper 8, which is co-authored with fellow researchers at NTNU (including two of my supervisors), is a conceptual paper that explores the concept of *sustainability* as success criterion in public investment projects. Sustainability is a broad and ambiguous term, and the corresponding term "sustainable development," is even more confounding. Its many elements can be categorized in different ways, but the literature seems to agree that three broad dimensions or pillars need to be considered: the environmental, the economic and the social. Sustainable development can be seen as a holistic success criterion that comprises almost all other success criteria, or as a somewhat more restrained criterion that takes a strategic perspective and may itself be part of a holistic model.

The paper argues that a feasible starting point is the definition provided by the OECD, which has introduced the concept of sustainability into the domain of project management. The definition of a sustainable project is as follows (originally in the context of development projects, but here presented in more general terms): *The continuation of benefits from [an intervention] after [the project] has been completed. The probability of continued long-term benefits. The resilience to risk of the net benefits flows over time* (OECD, 2002: 36).

The definition includes some explicit characteristics (e.g., long-term, intended benefits, balancing of costs and benefits, and resilience to risk). It does not indicate any particular analytic tool that can be used to determine whether a project is sustainable.

From the literature, we find that sustainability in projects is interpreted and used differently by different parties, such as engineers, architects, economists, politicians, users, and environmentalists. For example, within the realm of environmental sustainability, some use the term in order to describe environmentally friendly asphalt qualities. Others insist that road construction for personal vehicle use is not sustainable.

The paper proposes a framework to understand and discuss the term across professional and academic disciplines. The idea is that three different analytic levels exist for assessing sustainability—operational, tactical and strategic—in the same way as projects have goals an all three levels. Thus, at the strategic level, we may ask whether sustainability is best assured by improving the road infrastructure or by some other investment in the proposed area. Correspondingly, at the tactical level, we conduct sustainability assessments of different alternative pathways. At the operational level, sustainability may refer to the conditions for road workers or the choice of asphalt quality. We need to acknowledge the difference between doing the projects more sustainably and choosing more sustainable projects. In addition, it is essential to balance the environmental, economic and social impacts of an investment project in order to ensure that it is sustainable. Figure 14 shows how the three types of impacts can be assessed on each of the three analytic levels.

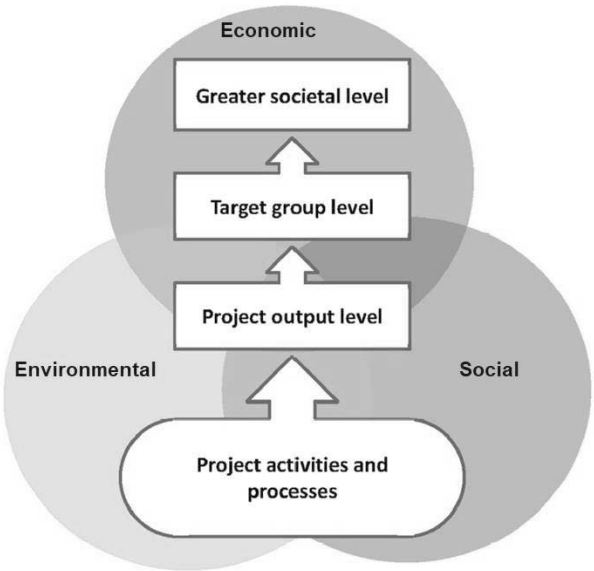


Figure 14 Sustainability assessment at project level (from Paper 8)

Paper 8 goes on to examine a number of analytic tools and analyze how they can serve to assess the sustainability of large public investment projects. Some cover only one pillar and

one level of analysis, whereas others attempt to cover several. No single analytic tool seems to exist that manages to grasp all aspects of the sustainability of an investment, and balance the different perspectives. Thus, for example, the CBA is not sufficient as a measure of sustainability, as discussed also in Paper 7.

Generally, social sustainability is considered the most challenging of the three pillars, being difficult to quantify, since it concerns the well-being of citizens, as well as society as a whole.

The main message of the paper is simply that analysts should always refer to the level of analysis that they apply. To quote from the paper: "Apples are apples, and economic impacts on the strategic level are economic impacts on the strategic level." Although sustainability assessment is complex and multifaceted, a clarification of what we are talking about ought to contribute to better dialogue and a more well-founded understanding.

6.3 Public project success as seen in a broad perspective: lessons from a meta-evaluation of 20 infrastructure projects in Norway (Paper 9)

Paper 9, the last paper, aims to demonstrate the importance of and need for a comprehensive evaluation approach to measuring the success of large infrastructure projects. A generic six-criteria evaluation framework is proposed, based on the OECD DAC framework frequently used to evaluate development assistance projects. The paper focuses on ex post evaluation, with the starting point that infrastructure projects in developed countries are rarely evaluated ex post, and that our knowledge about their various impacts is surprisingly limited. Moreover, as suggested by Samset and Christensen (2017), the early appraisal of an investment should ideally apply the same criteria as will be used in ex post evaluation, and thus increase the likelihood of a successful project outcome.

The six criteria are shown in Figure 15, in relation to the project process (i.e., the assumed set of causal relationships between inputs, project activities, outputs, and outcome). One criterion (efficiency) measures operational project success, another (effectiveness) measures tactical success, and the remaining four (other impacts, relevance, sustainability, and benefit-cost efficiency) measure strategic success. Together, the six criteria cover intended and unintended effects alike, as well as goal-oriented and value for money perspectives, and explicitly asks about sustainability.

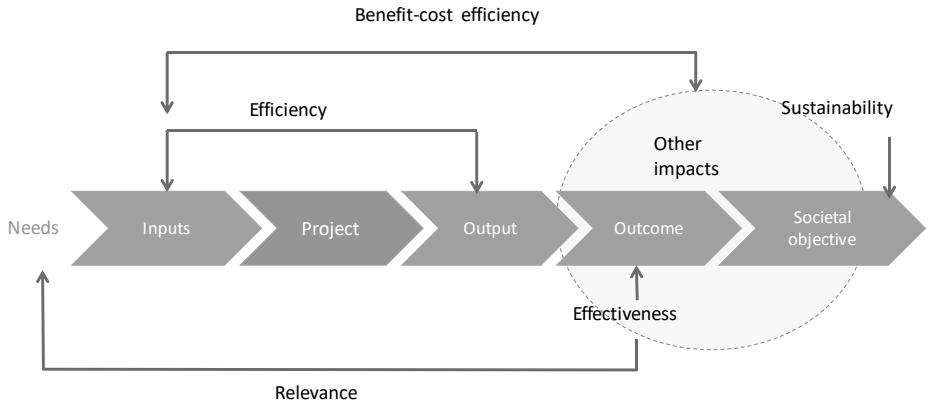


Figure 15 The six evaluation criteria (Figure 2 in Paper 9)

Clearly, the specific indicators used to assess each criterion will depend on the project. Thus the model is flexible, but also sufficiently standardized to allow for comparisons between projects. The purpose is to give an *overall* assessment of public projects' success, as well as a reasonably comprehensive assessment of each criterion. The assessment for each criterion is summarized by setting a score between 1 and 6.

The framework was applied to 20 completed projects during the period 2012–2017: 8 road projects, 5 buildings, 3 railway projects, 2 ICT projects, and 2 defense projects. As of 2017, they constituted about half of the QA2 projects that had been completed and in operation for at least five years (which was assumed to be long enough for the effects to materialize). The Concept Research Programme was responsible for the evaluations, but several evaluators from research institutions and consultancies in Norway and Sweden were involved.

The first part of the paper presents an aggregation and synthesis of evaluation results. It demonstrates that the majority of the 20 projects were successful, as considered by the evaluators, especially in operational terms, with an average score of 4.7 on their "efficiency." Their performance was somewhat more varied in tactical and strategic terms. In some cases, there were conflicts between the criteria, such as when projects had a high degree effectiveness, but also negative side effects, or when some of the projects scored high on relevance and sustainability, but lower on benefit-cost efficiency. In some such cases, it is probable that the project could have been redesigned or scaled up or down in order to achieve improved results overall. However, it should be noted that none of the projects had been through a comprehensive assessment and quality assurance of the choice of concept (QA1).

The second part of the paper discusses lessons from the evaluations, and assesses the feasibility of the evaluation framework. The overall conclusion is that the framework worked well, and the six criteria cover the main aspects of public project success. Some of the evaluators who were used to a narrower project management perspective or to cost-benefit analysis alone noted that the set of four strategic criteria were "eyeopeners."

The meta-evaluation based on the 20 individual exercises identified some improvement points, and revealed a need for clear requirements and guidelines regarding how the evaluation teams should be put together, how the criteria should be understood, and common principles for score-setting (which was largely based on the teams' subjective assessments). Thus, the 20 evaluations provided useful experiences and a basis for consolidating the model and practice.

6.4 Discussion

Papers 7–9 contribute to answering RQ3—*How can project success be achieved through improved appraisal and evaluation methods*— from various perspectives.

The business case, or project appraisal, is commonly seen as important in ensuring strategic project success. It should be presented to decision-makers at an early stage, and present and explore different solutions to the problem at hand. Further, it should focus on strategic issues and be reliable and relevant. However, some findings from previous papers indicate that the term "strategic issues" may be understood differently, and it is not clear how to ensure that the project appraisal establishes a credible and useful basis for decision. The last set of three papers studies the *content* of project appraisal and evaluation in more depth.

The three papers differ in several ways. Paper 7 concerns experiences relating to the narrower CBA, whereas the other two discuss broader frameworks. Furthermore, whereas Paper 8 is conceptual, Papers 7 and 9 are empirical and present quite a lot of data from ex ante and ex post evaluation respectively.

Again, it should be noted that the scope of the empirical papers is state-funded projects in Norway under the QA scheme, and that generalization of the findings should be made with caution.

The CBA alone may not hold as basis for decision

The main focus of quality assurers as part of QA1 has been to ensure projects' value for money, based on CBA. With the introduction of QA1, and efforts made to promote the quality and accountability in CBAs, it seems that in Norway decision-makers largely trust the CBAs conducted as part of the scheme. In particular, I do not find that perverse incentives on the part of project initiators have materialized in the estimation of costs and benefits. These findings are in line with those presented in Papers 2 and 4 about how external reviews may have a disciplining effect on project initiators.

However, notwithstanding value for money being a useful and relevant decision criterion, Paper 7 also finds that the CBA is not accepted as the *only* decision criterion. Ideally, the conceptual alternatives included in the analysis should all be in line with goals and requirements to the same extent, and/or all of these goals should be fully captured by the value-for-money measure. However, in reality, the alternatives may achieve agreed goals to different extents, and those goals may not necessarily support economic efficiency. For example, the CBA does not distinguish between intentional and unintentional impacts, and largely ignores moral and social issues, as well as impacts for future generations. Therefore,

regardless of how excellent and unbiased the CBA is, its usefulness in decision-making is limited by the ethical and normative fundament on which it is built.

Paper 7 demonstrated how ministries and agencies try to solve these problems by adding qualitative elements to the NPV. However, in doing so, they tend to mix "true" economic elements with those that do not really belong in a CBA, with the result that it is unclear what has been measured. The conclusion presented in the paper is that the CBA ought to be supplemented by, but clearly distinguished from, an assessment of the project in other perspectives. This recommendation is clearly not new (e.g., Nyborg, 2014), but demonstrates the difficulties in practice.

"Broad framework" with multiple meanings

The next question is what to include in a "broad" appraisal. Alternative decision perspectives beyond value for money are suggested in Paper 7, such as equity, sustainability and the achievement of agreed goals (often referred to as target benefits). Other, holistic models are available and include a fixed set of criteria, such as the OECD DAC framework or the Five Case Model applied to UK projects by central government.

A salient observation is that different models overlap to a considerable extent, but may emphasize different aspects. This depends also on the level of analysis chosen. Whereas Paper 7 focuses on strategic success (which in turn overlaps with tactical aspects and thus covers both the *Why* and *What* questions), various other models may also incorporate the operational perspective (i.e., the *How* question) to provide an overall assessment of project success.

Thus, project success is multifaceted in more than one dimension. To be able to discuss all its elements, it is vital to define them and clarify what one is talking about. This is the main message in Paper 8, in which sustainability, itself a holistic success criterion, is defined in terms of two dimensions: the level of analysis, and the economic, social, and environmental pillars.[5] Although the purpose of the project often lies in just one of the pillars, it should ideally be sustainable in all respects.

The framework suggested in Paper 9, which has been applied to evaluate completed public projects in Norway, captures decision perspectives, and all three levels of analysis (although emphasizing the strategic level). A somewhat simpler version of the same model was presented and tested on road and railway projects by Welde and Volden (2018).

Identifying appropriate success criteria, and assessing the project according to them, is certainly not easy, and inevitably involves a number of subjective assessments, and balancing and weighing of aspects. Even the CBA, which in principle is a precise, quantitative tool used to rank projects unambiguously, is difficult in practice, and a considerable part of the analysis comprises non-monetized impacts. In this situation, the role of the evaluator, including his or her competence and impartiality, is crucial, regardless of the content of the business case.

[5] On a more general level, one could have defined different decision perspectives, areas of interest, or systems definitions, instead of the three pillars

More use should be made of ex post evaluation

Project appraisal is often a crucial part of project governance frameworks, whereas ex post evaluation is more uncommon. This is also the case in the Norwegian QA scheme. The lack of evaluation is in effect a paradox, as there is much to learn from one project to another, both within and between sectors and countries. When a project succeeds, it should be imperative to ask what was done right. Correspondingly, one should learn from mistakes. Given the poor reputation of public projects in general, the potential to improve project practices is considerable. This also applies to the potential to improve estimation tools and methods, and the quality-at-entry scheme itself.

Furthermore, accountability is another purpose of evaluation. Project managers, project owners and others with responsibility for any of the project's outputs or outcome will certainly do their best, knowing they will be held accountable in retrospect. I believe that ex post evaluation should be an essential element of any project governance scheme. Additionally, all evaluation frameworks, regardless whether they are applied ex ante or ex post, should include multiple success criteria.

The experiences relating to the six-criteria evaluation model in Norway are encouraging. Not only does it provide a holistic picture of the project's success, but it also facilitates learning and improvement. For example, using the performance of previous projects to estimate the costs and benefits in future projects is both a way to improve the accuracy of CBAs, and a way to avoid optimism bias, whether intentional or stemming from cognitive weaknesses. Furthermore, the possibility to compare, and to learn across different sectors, is considered useful. Some sectors are better at cost control, others at benefits realization, and still others at sustainability.

In having the ex post evaluation in mind already at the project appraisal stage, one may also ensure that the right elements are included in the business case to serve as baseline data. An evaluation without baseline data will clearly be more difficult and less useful. One of the findings from Papers 7 and 9 is the tendency to not save assumptions and parameter values used in the project appraisal, in which case it is almost impossible to know in retrospect whether or not performance was as expected.

7. Reflections and conclusions

Three broad research questions have guided the research process for this thesis:

> RQ1. *What are the main challenges in front-end management of projects?*
>
> RQ2. *How can project success be achieved through project governance frameworks?*
>
> RQ3. *How can project success be achieved through improved appraisal and evaluation methods?*

The scope is limited to major public projects in Norway and the Norwegian QA scheme, as discussed in Sections 1.2–1.3.

The Norwegian QA scheme provided me with an excellent opportunity to follow and learn from experiences regarding this specific initiative, which was taken by a government. This thesis has been prepared in a period during which increasing amounts of empirical data are becoming available, as the number of QA projects increase, and many of them have been completed and are now entering their operational phases.

7.1 Main findings

Each paper addresses a different issue, and together they offer a sound picture of the potential for improving project performance through the use of project governance frameworks and better appraisal and evaluation models.

The first research question, RQ1—*What are the main challenges in front-end management of projects*—is addressed in Papers 1–3. These papers are based on literature reviews and a number of Norwegian cross-case analyses of *pre-QA1* projects. Weaknesses and challenges in the front-end of public projects are framed as ten paradoxes in Paper 1. A key message is that strategic project success and front-end management ought to be given more attention. Despite the now well-known recommendations about exploring problems, needs, and alternative solutions (the *Why* and *What* questions), there is still a tendency to jump rapidly to the narrower and short-term project management issues (the *How* question). Consequently, the financing party ought to make sure that all the three questions are properly addressed. Funding should be considered part of the conceptual solution, and perverse incentives during project initiation should be avoided.

The second research question, RQ2—*How can project success be achieved through project governance frameworks*—is addressed in Papers 4–6, all of which are empirical. They discuss preliminary experiences relating to the Norwegian QA scheme and compare its characteristics with those of similar schemes in other countries. The Norwegian scheme is largely in line with key recommendations from the literature, and results are promising thus far. The use of external QA reviews has been essential. Interestingly, such a simple scheme established at the topmost level has triggered improved practices in ministries and agencies. However, some of the paradoxes reported in Paper 1 seem more difficult to overcome, and my research indicates that central government should continue to emphasize the importance of the tactical and strategic aspects. There is an opportunity for countries with similar

schemes to learn from each other in order to develop and improve their own schemes. For example, the Norwegian government may benefit from the experiences of countries that have introduced intervention points at an earlier stage.

A somewhat worrying finding is the large number of projects in which the government did not adhere to QA1 recommendations based on value for money as the overall success criterion. This gave rise to the third research question, RQ3, *How can project success be achieved through improved appraisal and evaluation methods*, which is addressed in Papers 7–9. A recurrent theme is that strategic project success is a multifaceted term, and that this ought to be reflected in the appraisal and evaluation models. In particular, "value for money" as measured by the CBA is not accepted by decision-makers as the *only* success criterion. If the importance of CBA findings is overemphasized by quality assurers, there is a risk that decision-makers will ignore their advice altogether. Instead, I recommend a framework based on the OECD DAC criteria, and present experiences of its application to 20 completed projects.

7.2 Contribution to theory

The project management community and literature has moved beyond the narrow perspective on the "iron triangle" and is increasingly recognizing the wider, strategic view on projects (e.g., Morris, 2013; Williams et al., 2019). As part of this development, front-end management and project governance have both become issues of importance. However, we still need to expand the field in various directions to cover a large number of perspectives. As an economist, I have been concerned about how economic perspectives and models, such as principal agent theory and the CBA, can contribute to a better understanding of public project success. However, I also see it as necessary to communicate the weaknesses that may be associated with such perspectives, and highly recommend that they are combined with other perspectives.

The contribution of Papers 1–3 is their exploration and synthesis of various challenges and weaknesses in the front-end of projects. The literature in this area is relatively recent and fragmented, although Williams et al. (2019) have offered an important contribution more recently. For the most part, the findings of Papers 1-3 are not fundamentally new discoveries, but some have been explored to a lesser extent, most notably the risk of perverse incentives. The framing of various weaknesses as paradoxes should be a useful step: only by understanding how counterintuitive results are possible, can we discuss how to avoid them. Paper 1, about the ten paradoxes, has been cited a considerable number of times in the international literature to date, which is itself an indication of its usefulness.

Papers 4–6 present and discuss experiences relating to the introduction of a fairly recent project governance scheme at the Cabinet level. Case studies of this type ought to be a useful contribution to project governance literature, in light of the limited amount of empirical research in this area. Only a few other similar case studies seem to have been conducted (e.g., Klakegg et al., 2015; Williams et al., 2010). The findings illustrate the potential to improve project success, not least through external reviews, but also note the difficulties in

overcoming some of the above-mentioned paradoxes. Hopefully, the number of studies of "what works" in terms of project governance frameworks in different countries and contexts will increase in the future. Paper 6 adds to the conceptual framework through the suggestion that public project governance frameworks are seen as hierarchical systems.

The main contribution to theory from Papers 7–9 is the demonstrated need for a holistic evaluation methodology. Hopefully, these findings will be absorbed in project governance research, where they may contribute to the discussion of how to assess conceptual solutions, and how to increase the use of ex post evaluation of projects. They might also contribute to evaluation literature, as they demonstrate that an evaluation framework commonly applied in development aid projects can be used on projects in a more complex context in a high-income country. Finally, the findings might contribute to the somewhat fragmented CBA literature, in which usefulness and trust in CBAs is an under-researched topic, as are assessments of the non-monetized costs and benefits. The Concept Research Programme is currently following up the latter topic, through an ongoing study to develop and improve the methodology for assessing the non-monetized impacts in a CBA.

7.3 Contribution to practitioners

As part of the work for this thesis, research findings have also been disseminated to "the industry" in various ways (the industry is defined here as ministries, agencies, quality assurers, and other project practitioners).

First, several papers include recommendations to those involved in the appraisal, governance and management of major public projects. Three papers stand out in this regard. Paper 3, which is published as a textbook chapter, includes recommendations on how to develop a project proposal in order to avoid the challenges and weaknesses often observed in the front-end. Further, Paper 7 offers a set of practical recommendations to increase CBA usefulness, and Paper 9 presents a step-by-step procedure explaining how to use the recommended six-criteria evaluation framework.

Second, as the thesis is part of a larger research initiative, namely the Concept Research Programme, the findings from each study are disseminated through the program's newsletter (reaching more than 1,000 people in the target group). The program's other channels include the biannual Concept International Symposium on Project Governance and the yearly QA Forum. The former is an international symposium organized by the Concept Research Programme, with participants from ministries and governmental agencies, as well as from academia, project organizations, and industry. Papers 1 and 9 (and briefly Papers 4 and 7) were presented at the 2018 event, Paper 5 at the 2016 event, and a preliminary version of Paper 2 was presented in 2014. The QA forum is a seminar hosted and organized by the Ministry of Finance, with participants from central government and consultancies in Norway. Most of the included studies have been presented at the forum and discussed with representatives of the industry; most recently, Paper 7 was presented in 2018 and Paper 6 the previous year. As trailing researchers, we also communicate results and improvement points

to the Ministry of Finance in yearly meetings and in their regular efforts to renew and improve the QA scheme.

Third, I have also participated in numerous other networks and seminars, presented research findings in bilateral meetings with ministries and agencies, and communicated research findings through articles in science journals in Norway.

It should be noted that any contributions to improved practices are of an indirect and long-term nature. Implementation of recommendations from the studies has *not* been an integrated part of any of the included studies. Nevertheless, it is my clear impression, based on feedback from the industry, that the findings are recognized and considered useful. Some examples are provided in the following:

- Generally, there has been a high demand for lectures on best practice concerning the decision base for the choice of concept in recent years (cf. Paper 3). My impression is that the Concept Research Programme is considered to be "the true expert" in this area nationally. When public project governance is discussed (not only the QA scheme, but also similar schemes that are currently being introduced by local governments), usually the Concepts Research Programme is cited.
- I observe that the term "perverse incentives" (introduced in Paper 1 and 2) is now well-known in the industry. Not everybody agrees that perverse incentives are (or have ever been) a major challenge, but at least the research has triggered the debate on them.
- The paper on cost-benefit analysis (Paper 7) was published very recently, but preliminary findings have been presented and discussed with representatives of the industry earlier and have attracted considerable interest. I have been invited as an expert to contribute in various projects to develop guidelines for cost-benefit analysis in government agencies. In particular, practitioners strongly agree that an improved methodology for handling non-monetized impacts is needed. The Concept Research Programme has recently initiated a new study to develop such a methodology, strongly encouraged by the industry.
- There has been great interest in the experiences of the six-criteria evaluation model (Paper 9). Several sectoral ministries have shown interest in the model and some have even started to apply the model in their own project evaluations.

7.4 Further work

The findings from this thesis are only a small step, and the knowledge generated may be developed further in various directions. Each of the nine papers includes suggestions for further work. A few overall recommendations are discussed here.

Caution should always be exercised when expanding case study findings to different contexts. The scope of this thesis has been limited to Norway and the Norwegian QA scheme, implying that the findings apply to large, state-funded projects in recent years, and in the sectors covered by the QA scheme (i.e., primarily transport infrastructure projects, defense

projects, building construction projects, major public ICT projects, and certain major sports events). I strongly recommend that more research is done on the implementation of project governance frameworks under different circumstances, such as smaller projects, projects in other countries and sectors, and under different governance schemes.

We certainly need more data on the effects of the Norwegian scheme, as well as those of other project governance frameworks. In the years to come, there will be a lot more to learn from the Norwegian QA scheme, when projects subjected to external quality assurance of the choice of conceptual solution (QA1) are completed and enter into their operational phase. The findings thus far are preliminary. Researchers should continue to follow these projects, to determine whether the encouraging results persist, but also to pursue some potentially worrying findings, such as the risk that the project content, scope and cost increase during the process that *precedes* QA2.

Hopefully, the knowledge generated through Papers 4–6 might provide a basis for formulating more specific hypotheses that could be examined with quantitative data, such as about the impact of QA1, and the choice between model types (top-down versus bottom-up), depending on contextual differences at country level. The findings could also be useful for researchers who might want to investigate the relationship between the structure-based type of governance, which has been the focus here, and other relationships-based types of governance.

Another topic that deserves further attention is the usefulness of project appraisal and evaluation from a decision-maker's perspective. The potential for improvements discussed in Papers 7–9 ought to be explored further, and any attempts to improve appraisal practices should certainly be followed up by researchers.

In some of the papers, it is explicitly assumed that decision-makers are rational and follow an instrumental decision logic. I have argued that democratic decision-making processes are inherently difficult to predict, and therefore, that the biggest potential for improvement lies in strengthening the analytical process, as well as making decision processes transparent. The study of political decision-making processes, although complex and difficult to predict, could certainly add to the understanding of why projects sometimes fail, even when there has been adherence to formal rules for planning and decision-making.

References

Adams, W.M. (2006). *The future of sustainability: Re-thinking environment and development in the twenty-first century*. Report of the IUCN Renowned Thinkers Meeting. https://cmsdata.iucn.org/downloads/iucn_future_of_sustanability.pdf (accessed July 2019).

Ahola, T., Ruuska, I., Artto, K. & Kujala, J. (2014). What is project governance and what are its origins? *International Journal of Project Management* 32(8), 1321–1332.

Annema, J.A. (2013). The use of CBA in decision-making on mega-projects: Empirical evidence. In: Priemus, H. & van Wee, B. (Eds.) *International handbook on mega-projects*, 291–313. Cheltenham: Edward Elgar.

Ansar, A., Flyvbjerg, B., Budzier, A. & Lunn, D. (2016). Does infrastructure investments lead to economic growth or economic fragility? Evidence from China. *Oxford Review on Economic Policy* 32(3), 360–390.

Aritua, B., Smith, N.J. & Bower, D. (2011). What risks are common to or amplified in programmes: evidence from UK public sector infrastructure schemes. *International Journal of Project Management* 29 (3), 303-312.

Association for Project Management (APM). (2011). *Directing change: A guide to governance of project management*. Princes Risborough: Association for Project Management.

Association for Project Management (APM). (2018). *Find a resource.* https://www.apm.org.uk/resources/ (accessed April 2018).

Baccarini, D. (1999). The logical framework method for defining project success. *Project Management Journal* 30(4), 25–32.

Bakewell, O. & Garbutt, A. (2005). *The use and abuse of the logical framework approach: A review of international development NGOs' experiences*. A report for Sida. http://pdf2.hegoa.efaber.net/entry/content/909/the_use_and_abuse_SIDA.pdf (accessed July 2019).

Berechman, J. & Wu, Q. (2006). *Cost overruns risk analysis in transportation infrastructure investments*. UBC P3 Project Working Paper 2006-5. https://webcontent.sauder.ubc.ca/Faculty/Research_Centres/Phelps_Centre_for_the_Study_of_Government_and_Business/Projects/UBC_P3_Projects/~/media/Files/Faculty%20Research/Phelps%20Centre/2006_05_berechman.ashx (accessed July 2019).

Berg, P. et al. (1999). *Styring av statlige investeringer. Sluttrapport fra styringsgruppen, Finansdepartementet*, February 10. 1999. Oslo: Norwegian Ministry of Finance.

Blumberg, B., Cooper, D. & Schindler, P. (2011). *Business research methods*. 4th ed. MacGraw Hill Higher Education.

Boardman, A., Greenberg, D., Vining, A. & Weimer, D. (2011). *Cost-benefit analysis: Concepts and Practice*. 4th ed. Pearson Education.

Cantarelli, C.C., Flyvbjerg, B., Molin, E.J.E. & van Wee, B. (2010). Cost overruns in large-scale transport infrastructure projects: Explanations and their theoretical embeddedness. *European Journal of Transport Infrastructure Research* 10(1), 5–18.

Cantarelli, C.C., van Wee, B., Molin, E.J. & Flyvbjerg, B. (2012). Different cost performance: Different determinants? The case of cost overruns in Dutch transport infrastructure projects. *Transport Policy* 22, 88–95.

Chianza, T. (2008). The OECD/DAC criteria for international development interventions: An assessment and ideas for improvement. *Journal of MultiDisciplinary Evaluation* 5(9), 41–51.

Chih, Y.Y. & Zwikael, O. (2015). Project benefit management: A conceptual framework of target benefit formulation. *International Journal of Project Management* 33, 352–362.

Christensen, T. (2009). *The Norwegian front-end governance regime of major public projects – a theoretically based analysis*. Concept Report No. 23. Trondheim: Norwegian University of Science and Technology.

Cicmil, S. & Braddon, D. (2012). Fading glory? Decision-making around the project – how and why "glory" projects fail. In: Williams, T. & Samset, K. (Eds.) *Project governance: Getting investments right*, 221–255. Basingstoke: Palgrave Macmillan.

Crawford, L., Cooke-Davies, T., Hobbs, B., Labuschagne, L., Remington, K. & Ping, C. (2008). Governance and support in the sponsoring of projects and programs. *Project Management Journal* 39, S43–S55.

Creswell, J.W. (2014). *Research design: Qualitative, quantitative and mixed methods approaches*. 4th ed. Thousand Oaks, CA: SAGE.

Dudovskiy, J. (2018). *An ultimate guide to writing a dissertation in business studies: a step-by-step assistance.* E-book. Available at: https://research-methodology.net/ (accessed July 2019).

Eisenhardt, K. (1989). Agency theory: An assessment and review. *Academy of Management Review* 14(1), 57–74.

Eliasson, J., Börjesson, M., Odeck, J. & Welde, M. (2015). Does benefit-cost efficiency influence transport investment decisions? *Journal of Transport Economics and Policy* 49(3), 377–396.

Finne, H., Levin, M. & Nilssen, T. (1995). Trailing research: A model for useful program evaluation. *Evaluation* 1(1), 11–31.

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry* 12(2), 219–245.

Flyvbjerg, B. (2009). Survival of the unfittest: Why the worst infrastructure gets built – and what we can do about it. *Oxford Review of Economic Policy* 25(3), 344–367.

Flyvbjerg, B. (2014), What you should know about megaprojects and why: An overview, *Project Management Journal* 45(2), 6–19.

Flyvbjerg, B., Bruzelius, N. & Rothengatter, W. (2003). *Megaprojects and risk: An anatomy of ambition*. Cambridge, UK: Cambridge University Press.

Flyvbjerg, B., Garbuio, M. & Lovallo, D. (2009). Delusion and deception in large infrastructure projects: Two models for explaining and preventing executive disaster. *California Management Review* 51(2), 170–193.

Flyvbjerg, B., Holm, M.S. & Buhl, S. (2002). Underestimating costs in public works projects: Error or lie? *APA Journal* 68(3), 279–295.

Gasper, D. (2000). Evaluating the 'logical framework approach' towards learning-oriented development evaluation. *Public Administration and Development* 20, 17–28.

Haanæs, S., Holte, E. & Larsen, S.V. (2006). *Beslutningsunderlag og beslutninger i store, statlige investeringsprosjekter*. Concept Rapport Nr 3. Trondheim: Norwegian University of Science and Technology.

Haavaldsen, T., Lædre, O., Volden, G.H. & Lohne, J. (2014). On the concept of sustainability, *International Journal of Sustainable Engineering 7(1)*, 2–12

Hall, P. (1981). *Great planning disasters*. Harmondsworth: Penguin Books.

Hasna, A.M. (2010). Sustainability classifications in engineering: Discipline and approach. *International Journal of Sustainable Engineering* 3(4), 258–276.

Heinzerling, L. & Ackerman, F. (2002). *Pricing the priceless: Cost benefit analysis of environmental protection*. Washington, DC: Georgetown Environmental Law and Policy Institute.

Helm, J. & Remington, K. (2005). Effective project sponsorship: An evaluation of the role of the executive sponsor in complex infrastructure projects by senior project managers. *Project Management Journal* 36(3), 51–61.

Hjelmbrekke, H., Klakegg, O.J. & Lohne, J. (2017). Governing value creation in construction project: A new model. *International Journal of Managing Projects in Business* 10(1), 60–83.

HM Treasury. (2013). *Public sector business cases using the five case model: Green book supplementary guidance on developing public value from spending proposals*. London: HM Treasury.

House, E. (2000). The limits of cost benefit evaluation. *Evaluation* 6(1), 79–86.

International Transport Forum. (2017). *Ex-post assessment of transport investments and policy interventions*. ITF Roundtable Reports. Paris: OECD Publishing.

Jenner, S. (2015). Why do projects ‘fail’ and more to the point what can we do about it? The case for disciplined, ‘fast and frugal’ decision-making. *PM World Journal* 4(3), 1-18

Kelly, C., Laird, J., Constantini, S., Richards, P., Carbajo, J. & Nellthorp, J. (2015). Ex post appraisal: What lessons can be learnt from EU Cohesion funded transport projects. *Transport Policy* 37, 83–91.

Klakegg, O.J. (2010). *Governance of major public investment projects. In pursuit of relevance and sustainability*. Doctoral theses at NTNU, 2010:15. Trondheim: NTNU.

Klakegg, O. J., & Meistad, T. (2014). “Individual relations – the core line in project based organizations”. IRSPM conference, 2014, Ottawa, Canada.

Klakegg, O. J. & Volden, G.H. (2017). Governance in public projects: The Norwegian case. In: Müller, R. (Ed.) *Governance and governmentality for projects: Enablers, practices and consequences*, 129–156. New York, NY: Routledge.

Klakegg, O.J., Williams, T. & Shiferaw, A.T. (2015). Taming the ‘trolls’: Major public projects in the making. *International Journal of Project Management* 34(3), 282–296.

Kloppenborg, T., Manolis, C. & Tesch, D. (2009). Successful project sponsor behaviour during initiation: An empirical investigation. *Journal of Managerial Issues* 21(1), 140–159.

Laursen, M. & Svejvig, P. (2016). Taking stock of project value creation: A structured literature review with future directions for research and practice. *International Journal of Project Management* 34(4), 736–747.

Lincoln, Y.S. & Guba, E.G. (1985). *Naturalistic inquiry*. Newbury Park, CA: SAGE.

Locatelli, G., Mariani, G., Sainati, T. & Greco, M. (2017). Corruption in public projects and megaprojects: There is an elephant in the room. *International Journal of Project Management* 35(3), 252–268.

Lovallo, D. & Kahneman, D. (2003). Delusions of success: How optimism undermines executives’ decisions. *Harvard Business Review* 81(7), 56–63.

Love, P., Sing, C.-P., Wang, X., Irani, Z. & Thwala, D.W. (2012). Overruns in transportation infrastructure projects. *Structure and Infrastructure Engineering* 10(2), 141–159.

Lundberg, M., Jenpanitsup, A. & Pyddoke, R. (2011). *Cost overruns in Swedish transport projects*. CTS Working Paper 2011:11. Stockholm: Centre for Transport Studies.

Mackie, P. & Preston, J. (1998). Twenty-one sources of error and bias in transport project appraisal. *Transport Policy* 5, 1–7.

Mackie, P., Worsley, T. & Eliasson, J. (2014). Transport appraisal revisited. *Research in Transport Economics* 47, 3–18.

Makovšek, D., Tominc, P. & Logožar, K. (2012). A cost performance analysis of transport infrastructure construction in Slovenia. *Transportation* 39(1), 197–2014.

McDonald, M. (2002). *Review of large public procurement in the UK*. https://www.parliament.vic.gov.au/images/stories/committees/paec/2010-11_Budget_Estimates/Extra_bits/Mott_McDonald_Flyvberg_Blake_Dawson_Waldron_studies.pdf (accessed March 2017).

Merrow, E.W. (2011). *Industrial megaprojects: Concepts, strategies and practices for success*. Hoboken, NJ: Wiley.

Meyer, W.G. (2014). The effect of optimism bias on the decision to terminate failing projects. *Project Management Journal* 45(4), 7–20.

Miller, R. & Hobbs, B. (2005). "Governance regimes for large complex projects". University of Montreal. Paper for the EURAM 2005 conference.

Miller, R. & Lessard D. (2000). *The strategic management of large engineering projects: Shaping institutions, risk and governance*. Cambridge, MA: MIT Press.

Morris, P.W.G. (2011). Managing the front-end: Back to the beginning. *Project Perspectives* 33, 4–9

Morris, P.W.G. (2013). Reconstructing project management reprised: A knowledge perspective. *Project Management Journal* 44(5), 6–23.

Morris, P.W.G. & Hough, G.H. (1987). *The anatomy of major projects: A study of the reality of project management*. Chichester: John Wiley.

Morris, P.W.G. & Geraldi, J. (2011). Managing the institutional context for projects. *Project Management Journal* 42(6), 20–32.

Mouter, N. (2017). Dutch politicians' use of cost-benefit analysis. *Transportation* 44, 1127–1145.

Musawir, A., Serra, C., Zwikael, O. & Ali, I. (2017). Project governance, benefit management, and project success: Towards a framework for supporting organizational strategy implementation. *International Journal of Project Management* 35(8), 1658–1672.

Müller, R. (2009). *Project governance: Fundamentals of project management*. New York: Gower.

Müller, R., Pemsel, S. & Shao, J. (2014). Organizational enablers for governance and governmentality of projects: A literature review. *International Journal of Project Management* 32, 1309–1320.

Müller, R., Shao, J. & Pemsel, S. (2015). *Governance and governmentality of projects*. Summary Report. April 2015. https://api.ning.com/files/BWHXRC*DO7kFFREzGKmZ*x95zAcrlm3tl3MrFkLdqncKfHfVc49MU*WJTkc0krH-pjf6*I53C1juCk5cc*wWrRtYbWhHz60g/SummaryofGovernanceofProjects.pdf (accessed July 2019).

Narayanan, V.K. & DeFillippi, R. (2012). The influence of strategic context on project management systems: A senior management perspective. In: Williams, T. & Samset, K. (Eds.) *Project governance: Getting investments right*, 3–45. New York, NY: Palgrave Macmillan.

Newby, T.M.J. (2010). *Unintended effects of development aid – a brief overview*. DIIS Working Paper 2010:06. Copenhagen: Danish Institute for International Studies.

Nicolaisen, M.S. (2012). *Forecasts: fact or fiction? Uncertainty and inaccuracy in transport project evaluation*. PhD thesis. Aalborg: Aalborg University, Department of Development and Planning.

Nicolaisen, M.S. & Driscoll, P.A. (2014). Ex-post evaluations of demand forecast accuracy: A literature review. *Transport Reviews* 34(4), 540–557.

Nyborg, K. (1998). Some politicians' use of cost-benefit analysis. *Public Choice* 95, 381–401.

Nyborg, K. (2014). Project evaluation with democratic decision-making: What does cost-benefit analysis really measure? *Ecological Economics* 196, 124–131.

Næss, P., Volden, G.H., Odeck, J. & Richardson, T. (2017). *Neglected and underestimated negative impacts of transport investments*. Concept Report No. 54. Trondheim: Ex Ante Academic.

Odeck, J. (2004). Cost overruns in road construction—What are the sizes and determinants? *Transport Policy* 11, 43–53.

Odeck, J., Welde, M. & Volden, G.H. (2015). The impact of external quality assurance of costs estimates on cost overruns: Empirical evidence from the Norwegian road sector. *European Journal of Transport and Infrastructure Research* 15(3), 286–303.

OECD. (2002). *Glossary of key terms in evaluation and results based management*. Paris: OECD.

OECD. (2010). *Guidance on sustainability impact assessment*. Paris: OECD.

OECD. (2015). *Towards a framework for the governance of infrastructure*. Paris: OECD.

Olsson, N. & Berg-Johansen, G. (2016). Aspects of project ownership in theory and practice. *Procedia Computer Science* 100, 790–795.

Ostrom, E., Gibson, C., Shivakumar, S. & Andersson, K. (2001). *Aid, incentives, and sustainability: An institutional analysis of development cooperation*. SIDA Studies and Evaluation 02/01. Stockholm: Swedish International Cooperation Agency.

Pearce, D., Atkinson, G. & Mourato, P. (2006). *Cost-benefit analysis and the environment: Recent developments*. Paris: OECD.

Picciotto, R. (2013). The logic of development effectiveness: Is it time for the broader evaluation community to take notice? *Evaluation* 19(2), 155–170.

Pinto, J.K. (2006). "Organizational governance and project success: Lessons from Boston's big dig". Paper presented at Concept Symposium, Trondheim, Norway.

Pinto, J.K. & Slevin, D.P. (1988). Project success: Definition and measurement techniques. *Project Management Journal* 19(1), 67–71.

Project Management Institute. (2017). *A guide to the project management body of knowledge (PMBOK® Guide)*. 6th ed. Newton Square, PA: Project Management Institute.

Rambøll and Agenda Kaupang. (2016). *Bruk av evaluering i statlig styring*. Rapportnr. R9392. https://evalueringsportalen.no/evaluering/bruk-av-evaluering-i-statlig-styring/Rapport%20DF%C3%98%20Bruk%20av%20evaluering%20i%20statlig%20styring.pdf/@@inline (accessed July 2019).

Rossi, P.H., Lipsey, M.W. & Freeman, H.E. (2004). *Evaluation: A systematic approach*. 7th ed. Thousand Oaks, CA, SAGE.

Samset, K. (2003). *Project evaluation: Making projects succeed*. Trondheim: Tapir Akademisk Forlag.

Samset, K. (2015). *Prosjekt i tidligfasen: Valg av konsept*. 2nd ed. Bergen: Fagbokforlaget.

Samset, K. & Christensen, T. (2017). Ex ante project evaluation and the complexity of early decision-making. *Public Organization Review* 17(1), 1–17.

Samset, K. & Volden, G.H. (2012). The proposal, book chapter in Williams, T. and Samset, K. (eds.), *Project Governance. Getting investments right*, Palgrave Macmillan, Basingstoke, 46-81

Samset, K. & Volden, G.H. (2013). Major projects up front: Analysis and decision – rationality and chance, paper presented at IRNOP Project Research Conference, Oslo, June 16–18, 2013

Samset, K. & Volden, G.H. (2014). Front end governance of major public projects – lessons with a Norwegian quality assurance scheme. *International Journal of Architecture, Engineering and Construction* 3(2), 1–10.

Samset, K. & Volden, G.H. (2015). Front-end definition of projects: Ten paradoxes and some reflections regarding project management and project governance, *International Journal of Project Management 34*, 297–313Sartorius, R.H. (1991). The logical framework approach to project design and management. *Evaluation Practice* 12(2), 139–147.

Scriven, M. (2015). *Key Evaluation Checklist (KEC)*. http://michaelscriven.info/papersandpublications.html (accessed July 2019).

Serra, C.E.M. & Kunc, M. (2015). Benefits realisation management and its influence on project success and on the execution of business strategies. *International Journal of Project Management* 33(1), 53–66.

Shenhar, A.J. (2004). Strategic Project Leadership®: Toward a strategic approach to project management. *R&D Management* 34, 569–578.

Shenhar A., Dvir D., Levy O. & Maltz, A.C. (2001). Project success: A multidimensional strategic concept. *Long Range Planning* 34(6), 699–725.

Siemiatycki, M. (2016). The making and impacts of a classic text in megaproject management: The case of cost overrun research. *International Journal of Project Management* 36(2), 362–371.

Solberg, H.A. & Preuss, H. (2007). Major sports events and long-term tourism impacts. *Journal of Sports Management* 21, 213–234.

Solberg H.A. & Preuss, H. (2015). Major sports events: The challenge of budgeting for the venues. *Event Management* 19(3), 349–363.

Standing Advisory Committee on Trunk Road Assessment (SACTRA). (1999). *Transport and the economy*. Norwich: TSO.

Swales, S. (1990). *Genre analysis: English in academic and research settings*. Cambridge, UK, Cambridge University Press.

Terlizzi, M.A., Albertin, A.L., de Oliveira, H.R. & de Moraes, C. (2017). IT benefits management in financial institutions: Practices and barriers. *International Journal of Project Management* 35(5), 763–782.

Thomas, D.R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation* 27(2), 237–246.

Too, E. & Weaver, P. (2014). The management of project management: A conceptual framework for project governance. *International Journal of Project Management* 32(8), 1382–1394.

van Wee, B. (2007). Large infrastructure projects: A review of the quality of demand forecasts and cost estimations. *Environment and Planning B: Urban Analytics and City Science* 34, 611–625.

van Wee, B. & Rietveld, P. (2013). CBA: Ex-ante evaluation of mega-projects. In: Priemus, H. & van Wee, B. (Eds.) *International handbook on mega-projects*, 269–290. Cheltenham: Edward Elgar.

Venables, A.J. (2007). Evaluating urban transport improvements: Cost-benefit analysis in the presence of agglomeration and income taxation. *Journal of Transport Economics and Policy* 41(2), 173–188.

Vickerman R. (2008). Cost-benefit analysis and the wider economic benefits from mega-projects. In: Priemus, H., Flyvbjerg, B. & van Wee, B. (Eds.) *Decision-making on mega-projects: Cost benefit analysis, planning and innovation*, 66–83. Cheltenham: Edward Elgar.

Volden, G.H. (2018a). Public funding, perverse incentives, and counterproductive outcomes, *International Journal of Managing Projects in Business 12(2)*, 466–486

Volden, G.H. (2018b). Public project success as seen in a broad perspective: Lessons from a meta-evaluation of 20 infrastructure projects in Norway, *Evaluation and Program Planning 69*, 109–117

Volden, G.H. (2019). Assessing public projects' value for money: An empirical study of the usefulness of cost–benefit analyses in decision-making, *International Journal of Project Management 37(4)*, 549–564

Volden, G.H. & Andersen, B.S. (2018). The hierarchy of public project governance schemes, *International Journal of Managing Projects in Business 11(1)*, 174–197

Volden, G.H. & Samset, K. (2017a). Quality assurance in megaproject management, the Norwegian way, book chapter in Flyvbjerg, B. (ed.) *The Oxford Handbook of Megaproject Management*, Oxford University Press, Oxford, 406-428

Volden, G.H. & Samset, K. (2017b). Governance frameworks for major public investment projects: Principles and practices in six countries, *Project Management Journal 48(3)*, 90–108

Wachs, M. (1987). Forecasts in urban transport planning: Uses, methods, and dilemmas. *Climatic Change* 11(1–2), 61–80.

Wachs, M. (1989). When planners lie with numbers. *Journal of the American Planning Association* 55(4), 476–479.

Weber, R.P. (1990). *Basic content analysis*. 2nd ed. Quantitative Applications in the Social Sciences. SAGE University Paper 49. Newbury Park, CA: SAGE.

Welde, M. & Volden, G.H. (2018). Measuring efficiency and effectiveness through ex-post evaluation: Case studies of Norwegian transport projects. *Research in Transportation Business and Management* 28, 33–39.

Whist, E. & Christensen, T. (2011). *Politisk styring, lokal rasjonalitet og komplekse koalisjoner: Tidligfaseprosessen i store offentlige investeringsprosjekter*. Concept Rapport Nr. 26. Trondheim: Concept-programmet.

Wiig, H. & Holm-Hansen, J. (2014). *Unintended effects in evaluations of Norwegian aid: A desk study*. Oslo: NORAD (Norwegian Agency for Development Cooperation).

Williams, T. & Samset, K. (2010). Issues in front-end decision making on projects. *Project Management Journal* 41, 38–49.

Williams, T., Klakegg, O.J., Magnussen, O.M. & Glasspool, H. (2010). An investigation of governance frameworks for public projects in Norway and the UK. *International Journal of Project Management* 28, 40–59.

Williams, T., Vo, H., Samset, K. & Edkins, A. (2019). The front-end of projects: a systematic literature review and structuring, *Production Planning and Control*, online-first publication doi: 10.1080/09537287.2019.1594429

World Bank. (1996). *Evaluation results 1994*. Washington DC: International Bank of Reconstruction and Development.

Yin, R.K. (2014). *Case study research: Design and methods*. 5th ed. Thousand Oaks, CA: SAGE.

Zwikael, O. & Smyrk, J. (2012). A general framework for gauging the performance of initiatives to enhance organizational value. *British Journal of Management* 23, 6–22.

Appendix 1 Terminology

Front-end management of public projects is still young as research area, and the terms may be used and interpreted differently in the literature. Unless stated otherwise in each paper, the definitions listed in Table 4 are applied throughout the thesis (to a large extent based on definitions suggested by Samset, 2003).

Table 4 Key definitions

Term	**Definition**
Appraisal	Assessment of the feasibility of a project concept in accordance with established decision criteria, such as value for money, relevance and sustainability.
Concept	Principle solution to a problem or need. Also referred to as a conceptual solution. The choice of project concept is a strategic one and is made prior to detailed planning.
Cost-benefit analysis (CBA)	A comparison of investment and operating costs with the benefits generated by the investment, normally expressed in terms of the population's willingness to pay to obtain them.
Effectiveness	A measure of the extent to which the agreed project outcome has been obtained.
Efficiency	A measure of how inputs are converted into project outputs.
Ex post evaluation	Evaluation that judges a completed project in accordance with established success criteria. Its purpose is to study how well the project performed, and to enable conclusions to be drawn for similar projects in the future.
Front-end	The project phases prior to project approval and funds appropriation. Includes not only detailed planning, but also the early idea phase and the conceptual phase.
Input	The financial, human and material resources necessary to produce the intended outputs of a project.
Other impacts	The positive and negative changes beyond the agreed outcome that can be attributed to the project.
Outcome	The likely or achieved first-order effects of the project's outputs. Also referred to as effect.
Outputs	The tangible results of the project.
Project	A temporary endeavor to create a unique product or service. A project consists of a set of planned, interrelated activities designed to achieve defined

	objectives within a given budget and a specific period of time.
Project governance	The processes and systems that the project owner (the financing party) must implement to ensure a successful project.
Project governance framework	A framework defining the governance arrangements that will apply throughout the project's life cycle. Project owners with many projects may establish a common framework to be applied to all of their projects.
Project management	The application of knowledge, skills, tools, and techniques to project activities in order to meet stakeholders' needs and expectations from a project.
Quality assurance	Encompasses any activity that is concerned with assessing and improving the merit or worth of a project or compliance with given standards. Here, primarily used in relation to independent reviews as part of a project governance framework.
Relevance	A project is relevant if there is a need for what it delivers. Project relevance is measured in relation to national political priorities, but also to stakeholders' preferences.
Sustainability	A project is sustainable if its benefits are likely to persist throughout its lifetime. This usually requires that the total impacts (financial, environmental and social) are acceptable in the long term.
Value for money	A measure of the extent to which the total willingness to pay exceeds project cost. Also referred to as benefit-cost efficiency, and measured by the CBA. (Secondarily, it is a measure of outcome in relation to cost, as measured by cost-effectiveness analysis).

Part 2 Papers

Paper I

Available online at www.sciencedirect.com

ScienceDirect

International Journal of Project Management 34 (2016) 297 – 313

International Journal of Project Management

www.elsevier.com/locate/ijproman

Front-end definition of projects: Ten paradoxes and some reflections regarding project management and project governance

Knut Samset [a,*], Gro Holst Volden [b]

[a] Norwegian University of Science and Technology, Norway
[b] SINTEF, Norway

Received 8 October 2014; received in revised form 12 January 2015; accepted 19 January 2015
Available online 16 February 2015

Abstract

The importance of the front-end decision-making phase in securing projects long-term success is being increasingly recognized. This area is underrepresented in the literature, but there are several key themes that run throughout, identifying key issues or difficulties during this stage. Clearly, a key to successful projects lies in the choice of concept. This paper presents some findings from the work of the Concept research programme on front-end management and governance of major public investment projects in Norway. It is based on studies that explore strengths and weaknesses in the processes of analysis and decision-making during the early phase before the final choice of conceptual solution is made, and the extent to which projects under study are (or are likely to be) relevant and effective in relation to needs and priorities in society. It concludes that there are frequent deficiencies in these processes, and that the potential for improvements is huge.

Keywords: Project governance; Front end management; Paradoxes

1. Introduction

A large and increasing share of the activities taking place in private as well as the public sector is organised as projects. In private sector projects, the ultimate goal is to improve the company's profitability, either directly or indirectly, through improvements in its competitiveness. In public projects, the commissioner is the government, representing the entire society and its taxpayers. In such cases, the benefits of the project must be considered in a broader societal perspective, to ensure that the project provides value for money and contributes to the desired development.

There are many challenges facing public investment projects that must be overcome to achieve project success, such as lack of competence among planners, avoidance of hidden agendas during planning, underestimation of costs and overestimation of benefits, unrealistic and inconsistent assumptions, and how to secure essential planning data and adequate contract regimes. Many of these problems can be interpreted in terms of deficiencies in the analytical or political processes preceding the final decision to go ahead. Hence, the importance of the front-end decision-making phase must be recognized to strengthen project governance.

The term governance is derived from the Latin word *gubernare*, meaning 'to steer'. It refers to the administrative and process-oriented elements of governing, whether undertaken by a government, market, or network, whether over a family, tribe, formal or informal organization, or territory, and whether

* Corresponding author.
E-mail address: knut.samset@ntnu.no (K. Samset).

http://dx.doi.org/10.1016/j.ijproman.2015.01.014
0263-7863/00/

through laws, norms, power, or language (Bevir, 2013). Governance is about processes of rule more than institutions of government. It relates to processes and decisions that seek to define actions, grant power, and verify performance. Different instruments are available to improve governance, ranging from legally binding regulations, to economic and other types of incentives, as well as information and skill development. The challenge in governance is to identify the optimal mix of different instruments.

Project governance refers to the processes, systems, and regulations that the financing party must have in place to ensure that projects are successful. This would typically include a regulatory framework to ensure adequate quality at entry, compliance with agreed objectives, management and resolution of issues that may arise during the project, and standards for quality review of key appraisal documents (Samset and Volden, forthcoming). These processes and regulations can often be described in terms of stage-gate phase models.

Project management refers to the processes established to organize and manage resources required to complete a project within defined scope, quality, time, and cost constraints. Whereas the literature on project management is substantial, project governance has only recently become an issue of importance in the project management community (e.g. Müller, 2009).

Peter Morris (1994) brought to our attention that in the early years, project management had an extremely narrow focus, reflected only in the project life cycle, and ignoring the critical front-end. He noted that as long as we only focus on the life cycle itself, we are missing the critical front-end and institutional elements (shown in his Management of Projects paradigm) that more accurately typify the responsibilities of the project owner and the project manager.

2. The present study

In the year 2000, the Norwegian Ministry of Finance introduced a governance regime for the country's largest public investment projects, the so-called Quality Assurance (QA) regime, in terms of a mandatory quality-at-entry scheme to meet such challenges. It is a simple stage-gate process with a top-down review of the quality of project proposals, which are typically the result of bottom-up processes of analysis and decision making in society. The Norwegian QA scheme includes two external reviews in the front-end: Quality Assurance of the conceptual solution (QA1) before Cabinet decision whether to start a pre-project, and Quality Assurance of the cost and steering frames (QA2) before the project is submitted to Parliament for approval and funding (see Fig. 1).

In parallel to the QA regime, the Norwegian University of Science and Technology in year 2002 initiated the so-called Concept research programme, designed to focus on the front-end management of major public projects. The governance scheme clearly would be a unique laboratory for research on longitudinal data. It has allowed researchers to follow the largest public projects in Norway since 2002. The Concept programme works to develop the research frontier in the area of project governance. This is undoubtedly an interdisciplinary field, and the programme has conducted separate studies in areas such as public management, project management, portfolio management, economic analysis, planning, decision-making, risk analysis, contract management, the theory of incentives, applied logic, and judgmental assessment.

The idea was to broaden the perspective on projects. To quote Morris (2009:60), "effective management of projects is more than just execution-oriented project management. Projects are undertaken to create value and deliver benefits. Shaping the interaction between the sponsor's goals and the way the project (or programme) is to be developed, in the best way possible, absolutely crucial — probably one of the most important aspects of managing a project".

This understanding is an underlying motivator of our research. However, the approach has been inductive rather than deductive. It has been more of a probe into new areas than a process guided by precisely formulated and theoretically founded problems. The perspective has been on projects as means to create value and deliver benefits. Some studies had a focus on decisions, others on analysis, but all of them were meant to provide insight into what is here termed project governance.

Miller and Lessard (2000) contended that the front-end phase from inception and until the budget is approved by Parliament takes 6–7 years on average in major public investment projects. This is also the case in Norway. The

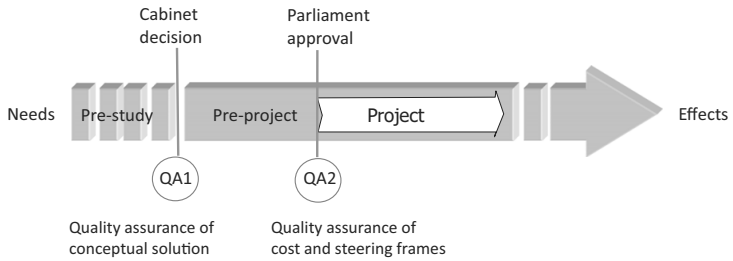


Fig. 1. The Norwegian Quality Assurance regime, a stage-gate phase model with two external reviews in the front-end of major public projects. Source: Norwegian Ministry of Finance.

subsequent implementation phase takes typically 3–5 years, and projects will have to be at least three years into their operational phase before an ex post evaluation can be undertaken. The Norwegian QA scheme has now been in operation for 14 years involving about 20 new projects each year. This means that there is a tremendous time lag as regards availability of empirical data. During the early years of the programme, researchers were first left with the option to do theoretical desk studies, and then gradually shifting into more hands-on studies of procedures and practices in planning, quality assurance and decision making during the front-end phase as more projects were added. Only recently, empirical data are becoming available. At present, the total number of projects is about 260, of which only 50 have been implemented so far. In addition, not more than 10 projects have reached a degree of maturity that allow for ex post evaluations.

Fourteen years after the quality assurance scheme was introduced it has proved to have a positive impact on cost control, since almost 80% of the first 40 projects were completed below budget, which is quite remarkable (Samset and Volden, 2013a, 2013b). Getting to grips with the choice of conceptional solution and securing the strategic performance of projects may prove to be a much more complex matter. It is challenging, but no less interesting.

The present study draws on some research findings from the Concept programme, presented below in terms of ten paradoxes, all of which have implications for the theory of project management and project governance, as well as a concluding study that demonstrates some of their implications. The term "paradox" in this paper is used to describe situations with a counter-intuitive result, some of which are based on fallacious reasoning or incomplete or faulty analysis. These are the paradoxes of:

1. How success is understood
2. The significance of front end management
3. Early information overflow
4. The opportunity space
5. Strategic alignment
6. Cost estimation
7. Disregarded analyses of costs and benefits
8. "Predict and provide"
9. Perverse incentives
10. Myoptic decisions.

Their common denominator is that they all focus on the choice of conceptual solution. Each paradox is rooted in one or more studies in the programme, but also inspired by research findings presented in the biannual international Concept symposia on project governance. According to Pinto (2014), these symposia ".. have provided much of recent theoretical and epistemological structure to the construct of project governance. Papers from these symposia have been on the leading edge of many of the insights we have on the current state of project governance".

The three first paradoxes are not rooted in empirical research, but in desk studies and literature reviews. The remaining seven (numbers 4–10) are based on cased studies involving 5–40 cases, most of them are major public projects that have been subjected to external quality assurance under the Norwegian QA scheme. It should be noted that most of these studies are written in Norwegian only. The reports can be downloaded from the programme's website www.concept.ntnu.no, where also summary reports in English are available.

Since this paper can only provide brief snapshots of the studies, most of the references will have to be found in the underlying reports, and are only to a limited degree included in this paper.

3. The success paradox: success is measured in terms of tactical performance rather than strategic performance

The term "success", used as an indicator, is a highly complex and aggregated measure. More than two decades ago Pinto and Slevin (1988) concluded that: "the concept of project success has remained ambiguously defined both in the project management literature and, indeed, often within the psyche of project managers... Until project management can arrive at a generally agreed upon determinant of success, our attempts to accurately monitor and anticipate project outcomes will be severely restricted".

"Success" may be interpreted differently by various individuals and institutions. It may be measured differently in different types of projects, and different individuals tend to assess the success of the same project differently depending on their preferences, values and to what degree they are affected by the project. In addition, the degree of success is time-dependent. For instance, Shenhar et al. (2001) offer a chronological sequence of events as a compound definition of project success: (1) meeting time, budget, and other requirements, (2) impact on the customer, (3) benefit to the performing organization, and (4) preparing the future. The project's stakeholders do not necessarily share the same view of success. The project manager typically sees his job successfully accomplished when the project is done on time, within budget, and to specifications. The users will be concerned about the immediate effects of the project, and the investor or commissioner will typically be more concerned with the long-term economic viability.

Success as a generic term means to gain advantage, superiority, accomplishment, achievement or added value. Measuring success will have to look beyond the immediate outputs of the project to assert the anticipated and wider impact in a longer-term perspective. A hospital will ultimately have to be assessed in terms of its health benefits. An industrial project might be judged essentially in financial terms, and an infrastructure project in term of its utility.

The assessment of success can be in absolute or in relative terms — that is in relation to what was agreed versus what was realistically achievable. Ambition is expressed in terms of the project's stipulated objectives. Its effectiveness is a direct measure of what has been actually achieved. Clearly, success measured in absolute terms may give a misleading conclusion if objectives are unrealistically ambitious. By

measuring in relative terms, that is in relation to what could reasonably be expected as compared with experiences in similar projects — the same project might possibly be considered a success.

The media tend to give unsuccessful projects more publicity than successful ones. However, their perspective is highly restricted. The number one criterion of failure in the media is cost overrun; number two is delay in time. Truly, a much wider view needs to be taken on the success and failure of projects. The initial choice of project concept is of critical importance. This represents the one key decision of many made during the lifetime of a project, which is likely to have the largest impact on long-term success or failure Williams (2008). Here, by "the project concept" we mean much more than just the technical solution — it includes the entire business case, all of the various organisations involved, and the various mechanisms and arrangements involved in the inter-organisational relationships, see Miller and Hobbs (2009).

Here, it is necessary to distinguish between the projects' tactical and strategic performance. Success in tactical terms typically means meeting short-term performance targets, such as producing agreed outputs within budget and on time. These are essentially project management issues. Strategic performance, however, includes the broader and longer-term considerations of whether the project would have a sustainable impact and remain relevant and effective in its operational phase, throughout its lifespan. This is essentially a question of getting the business case right, or, in short, of choosing the most viable project concept.

This is illustrated in Fig. 2. Tactical performance is a question of how the project is implemented, i.e. how inputs are converted into outputs. These are measures of its efficiency, here measured in terms of the cost, timing and quality of deliverables. Strategic performance is a question of how the project performs after the outputs have been delivered. This will have to be monitored with the more compound measures mentioned above, which would cover the broader and longer-term perspectives and to a lesser degree involve focusing on

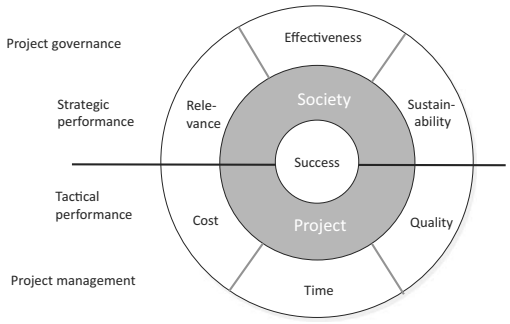


Fig. 2. Successful projects. Tactical performance is a question of delivering the project outputs as planned, while strategic performance is the worth or utility of the project as seen in a long-term perspective (Samset, 2014).

technology and management issues, but more on societal and economic aspects.

One example of tactically inefficient projects but viable in strategic terms could be the University Hospital in Oslo, Norway. Due to emerging new technologies and added responsibilities, captured during the engineering phase after the budget was decided, it was completed a year behind schedule and with considerable cost overrun, adverse newspaper reports and a public inquiry. No doubt that cost overrun was considerable in absolute terms, but in relative terms, it was equivalent to only a few months' operational costs for the hospital, and therefore insignificant in a lifetime perspective. The overall conclusion after a few years of operation was that the University Hospital was a highly successful project; and it would perhaps be unfair to suggest that initial decisions should be able to capture problems at this level of precision.

More serious by far is when a project fails in strategic terms, even if it successfully produces the intended outputs. It means that the choice of concept turns out to be the wrong one in relation to the problem at hand. In some cases, it may create more new problems than it solves, in others the initial problem no longer exists once the project is completed. One such example is an on-shore torpedo battery built inside the rocks on the northern coast of Norway in 2004 (Samset, 2008a, 2008b). The facility was huge and complex, designed to accommodate as many as 150 military personnel for up to three months at a time. It was officially opened as planned and without cost overrun. Already one week later it was closed down by Parliamentary decision, since it was obvious to all involved that a potential enemy would not expose its ships to such an obvious risk; the concept had long since been overtaken by political, technological and military development. What was quite remarkable was that this project, which can only be characterized as a strategic failure, got much less negative attention in the media than the University Hospital, possibly because it was a success in tactical terms.

Clearly, a successful project is one that delivers its outputs and significantly contributes to the fulfillment of agreed objectives. Moreover, it should have only minor negative effects, its objectives should be consistent with needs and priorities in society, and it should be viable in the sense that the intended long-term benefits resulting from the project are produced. These requirements were first formulated for US-funded international development projects by the United States Agency for International Development (USAID) in the 1960s, and subsequently endorsed by the United Nations (UN), the Organization for Economic Co-operation and Development (OECD), and the European Commission (EC) (USAID, 1980). They comprise five requirements or success factors that have to be fulfilled, i.e.: the project's efficiency, effectiveness, relevance, impact and sustainability. These are tough requirements that go far beyond the issues that usually are covered by the media or indeed by many planners and decision-makers.

Applied as standard requirements both up-front and ex post when projects are evaluated would be likely to improve project governance considerably in the future.

4. The paradox of the significance of front end management: less resources are used up front to identify the best conceptual solution (project governance), than to improve tactical performance during implementation (project management)

Projects are exposed to uncertainty in varying degrees and this is often used to explain their failures. Uncertainty characterises situations where the actual outcome of a particular event or activity is likely to deviate from the estimate or forecast value. Uncertainty may have many and various causes, related to the situation itself: the design of the project, the time perspective, available information, the implementation of the project, etc. (Marshall and Ritchie, 1993). Obviously, decision-making becomes difficult when uncertainty is high. Availability of relevant information reduces uncertainty from the decision-maker's point of view. It is widely believed that uncertainty is highest at the initial stage, when the project concept is conceived, and that it tends to reduce rapidly as information accumulates over time.

This line of thought is illustrated in Fig. 3. It follows that the utility of adding information is at its highest in the earliest stage. It is also commonly believed that the decision-maker's flexibility and the cost of making amendments are opposites. This is visualized with a similar graph. Decision-makers can juggle with different ideas and strategic solutions to a problem in the initial stages, but once decisions are being made, essential choices become locked, and it is more difficult and expensive to change the overall design. Therefore, major issues such as agreeing on the most effective solution to a problem and the choice of concept need to be dealt with as early as possible — later on is too late. Less essential issues such as avoiding major cost overrun can be handled later, for example when the final budget is agreed.

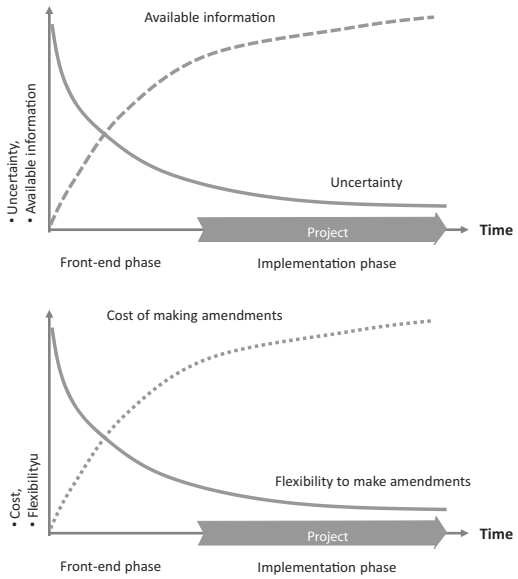


Fig. 3. People's conception of how uncertainty is affected by information and how flexibility to make amendments is restricted by cost, as time passes in a project. Source: Authors.

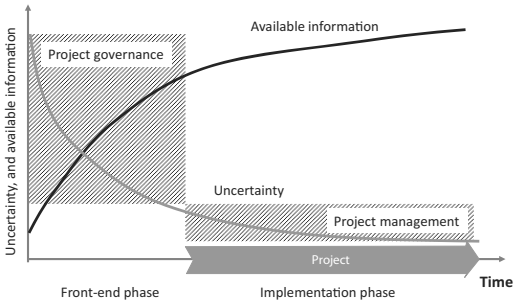


Fig. 4. The project life cycle. Uncertainty is greatest in the front-end phase and diminishes as more and better information is acquired for making decisions. Source: Samset (2010).

In Fig. 4, the distinction is made between the front-end and the implementation phase. The graph suggests that the potential to reduce uncertainty and risk is the largest up-front, and decreases substantially when the project is implemented. It is a paradox therefore that most of a project's planning resources may be spent on detailed planning and engineering, while too little is usually spent on getting the idea right from the start where the potential to reduce uncertainty by means of adding information is the largest. The paradox is that most resources are used to reduce uncertainty during the implementation phase, where the potential is much less.

Recent literature has highlighted the front-end phase including the project definition, as important for ensuring strategic project success (see for example Merrow, 2011; Morris, 2013). Where projects fail strategically, it is likely that the problem can be traced back to decisions in the earliest phases, when the initial idea was conceived and developed. What happens during the front-end phase is therefore essential for a project's success. A study by the World Bank based on a review of some 1125 projects concluded that 80% of the projects with a satisfactory "quality-at-entry"[1] were successful, while only 35% of those with an unsatisfactory quality-at-entry achieved success (World Bank, 1996). Improved front-end management is therefore likely to pay off in a wider life cycle perspective, as evinced by the IMEC study (Miller and Lessard, 2000). One way of improving quality-at-entry is by challenging initial ideas and applying simple analyses, extracting and making use of previous experience from similar undertakings, and consulting with stakeholders. Jordan et al. (1988) argued that 15% of the time and resources in projects should be spent

[1] Quality-at-entry was used as an indicator to characterize the identification, preparation and appraisal process that the projects had been subjected to upfront.

on front-end work, whereas Miller and Lessard (2000) suggested up to 35%.

In most cases the key issue at the earliest stage is to shed sufficient light on the underlying problem that provides the justification for the project, and the needs that the project is meant to satisfy. Detailed information about possible alternative solutions is less relevant. This illustrates what seems to be a major dilemma, since most projects originate as one specific solution to a problem, while the problem itself may not be analysed sufficiently, and alternative solutions may not have been considered at all. Typically, the preferred concept originates in the mind of one individual, based on intuition and experience, rather than systematic analysis of problems, needs, requirements, etc. Most of the information generated is associated only with the initially identified solution (Whist and Christensen, 2011).

A second dilemma is that this information, which may be very detailed and specific, tends to lock decisions into the initially preferred concept — to the extent that this will inevitably be the one that is finally chosen. It is all too rare that alternative concepts are identified and analysed to the extent that they get a fair trial in the subsequent decision process.

5. The paradox of early information overflow: decisions are based on masses of detailed information up front rather than carefully selected facts and judgmental information relevant to highlight the essential issues

It follows from the above that the front end phase is when fundamental choices are made, uncertainty is at its highest freedom to choose is at its optimum, and available information is most restricted. Adding information, therefore, makes sense — but only to a certain degree. The crucial issue is not the volume but what type of information is needed.

But contrary to the idea depicted in Fig. 3 the sheer amount of available information upfront might not be the issue. In the initial phase of a project the priority is to establish an overall perspective, and to analyse the problem in its context, considering the needs and priorities of stakeholders, users and affected parties, in order to come up with a sensible strategy. Opportunities and risks should be considered. Experience suggests that creativity, imagination and intuition can be more valuable at this stage than large amounts of data.

Decision making may be complex, unstructured, and affected by chance. Analysis may be biased or inadequate. Decisions may be affected more by political priorities than by rational analysis. Political priorities may change over time. Alliances and pressures from individuals or groups of stakeholders may change. The amount of information is large and may be interpreted and used differently by different parties. The possibility for disinformation is considerable.

Another aspect is that the early selection of a concept tends to survive decision-making, regardless of process, expert-driven rationalistic or more open-ended and democratic. This makes a strong case for proper research to identify the most viable concept up front. However, time factor, complexity and lack of predictability also imply that the outcome of rationalistic planning upfront tends to alter over time.

Exact quantitative information tends to be more affected by time than the choice of concept. On the one hand it is obvious that the higher the precision, the more rapidly information is outdated.[2] It is tempting to speak of the "half-life of information", see Fig. 5. For instance, exact information about the demand in a fast-developing market will have limited value after months, or even weeks. On the other hand, there are many examples to suggest that qualitative assessments tend to remain valid for much longer. Consider the assessment of users' fundamental preferences within a market segment. While it might not be possible to make a valid prediction of the actual demand three years into the future, it may be judged that demand will continue for a long time and can therefore be relied upon in strategic planning up front.

This suggests that restricted quality of information upfront may not be a major problem, since the need for precise information is low. It increases as the time for detailed planning approaches. In other words, the utility of exact information tends to reduce with the time-span. The opposite seems to be more of a problem: when decision-makers are confronted with an abundance of detailed information at an early point in time it may result in what is referred to as "analysis paralysis". This problem is discussed by Williams (2008). And besides, the cost of collecting information on a specific topic usually increases progressively with the amount of information collected. This is because more information requires more in-depth studies or more wide-ranging information searches. On the other hand, the gain in utility of additional information tends to decrease. This is because there is usually a critical amount of information that is needed to get the necessary insight in a situation: Additional information will be of limited use. Maximizing the utility/cost-ratio will set a limit to the amount of information that is useful (Jessen, 2012).

This emphasizes the need to invest in relevant information at the earliest stage of a project, while at the same time limits the search to what is useful for decision-making at this stage. A targeted search for information regarding the main uncertainties likely to affect the project is more cost-effective than an unguided search, since it makes it possible to increase the share of relevant information and reduce the total amount.

6. The paradox of the opportunity space: the choice of conceptual solution is made without systematically scrutinizing the opportunity space up front

Every project is initiated to solve some problem or meet some needs. And every project faces a choice of concept in terms of how to solve this problem. Consequently, a key task in the early phase of a project is to identify possible ways to solve the problem it has been mandated to solve (setting up the

[2] We need of course to make a distinction between lasting information, for example physical data on the one hand, and less durable information such as economic estimates on the other hand.

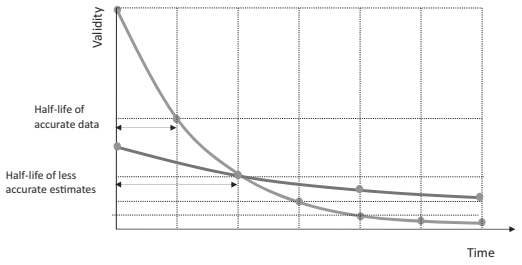


Fig. 5. Half-life of information. Validity tends to decrease over time during the front-end phase. More rapidly for accurate data than for less accurate estimates. Source: Samset (2010).

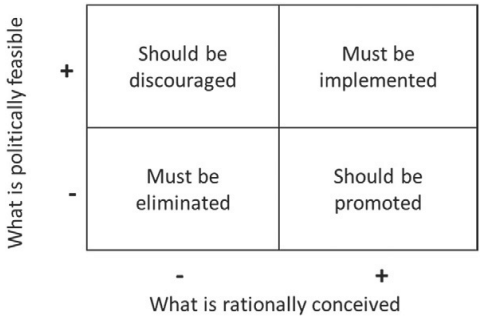


Fig. 6. Categories of projects within the opportunity space. Source: Whist and Christensen (2011).

opportunity space), furthermore to evaluate alternative concepts (limiting the opportunity space), and decide on the one best suited. There is much evidence to suggest that this is not always how things are done.

One problem is that planners are discipline experts with an inherent tendency to emphasize some aspects of the matter and downplaying others. The same may apply to organization undertaking the planning; its rules, procedures, etc. This is the reason for path dependency (Dosi, 1997; Margolis and Liebowitz, 2000); systematically choosing some solutions while avoiding others, even if these conflict with rational choices.

The situation become even more complex since these decisions are made at the intersection between the professional and political, in other words in-between what is rationally sound and politically possible. In the end, the complexity of the decision situation depends very much on whether there is an agreement about what one wants to achieve and what are the best means to this end (Christensen, 1985).

A case study of 17 major public projects was carried out to explore the use of the opportunity space, i.e. how it was defined, the type of conceptual alternatives identified and the effect on decisions (Andersen et al., 2014). It was found that in 11 cases the choice of concept had in reality already been made when the front-end process started, only in six cases, truly unique alternatives were identified. In most projects the analytic focus was narrowed to detailed project-specific issues at the expense of overall societal aspects. In half of the projects, the opportunity space was restricted to such a degree that real alternatives were excluded. There was a strong degree of path dependency where the alternatives represented a continuation of the current solution or variations over a theme.

It was emphasized that these processes take place on the borderline between the professional and political spheres, especially since the political backdrop is what exerts the most restricting effect on the opportunity space.

While the analytical process is largely within the realm of the professional constituency where the intention is to expand the opportunity space to allow identifying the best alternatives, the decision still remains with the political level. And the processes and decisions at this level are not always rational, as illustrated in Fig. 6. The two dimensions of rationally derived and politically feasible span four categories:

- The win/win projects score well on both dimensions and "must be implemented" (hydro power plant with no environmental downsides)
- Rational projects, but which are not politically mature, where a quality-at-entry approach, such as the Norwegian Ministry of Finance QA regime can aid in the decision process to get these promoted (close down nuclear power plants)
- Politically acceptable, even desired, but poorly conceived projects, these should be stopped, and the QA regime can help clarify the financial realities and thus kill such initiatives (Olympic games in a small country)
- The lose/lose projects have no support in either direction and should never go further (private exploration of space).

A separate case study of 23 major public investment projects (Whist and Christensen, 2011) went deeply into how the analytical and political processes interacted during the front-end phase, in order to understand how this affected the outcome of the projects. It was found that the majority of projects started out with a predetermined solution. In about half the cases an unambiguous problem analysis was nevertheless carried out, and in one third of the cases new problems were introduced during the front end phase, Fig. 7. The result was that two third of the projects were initiated with the same conceptual solution as the initial one, while in one third of the cases the conceptual solution was a different or changed substantially. Ten of the projects were considered relevant in relation to needs in society. Nine of these had a comprehensive problems analysis up front, and the Government had been a central actor in seven of them, while only in two of the thirteen projects were considered less relevant.

These studies, and the examples mentioned, first and foremost illustrate the unpredictability of the political system in a mature democracy; a well developed, rational decision basis is no guarantee for a rational choice of concept. It was concluded that a scheme with external quality assurance of the decision basis provided to the political level had proved to have

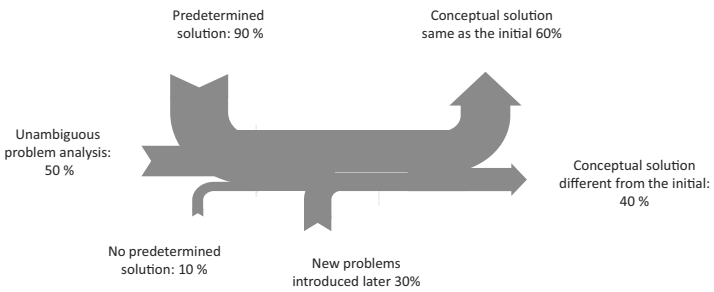


Fig. 7. Path dependency in defining and agreeing on conceptual solutions up front.
Source: Authors.

some positive effect in terms of helping make some choices more rational.

While the analytical part of the decision-making processes overall was rather weak, the participation of and control with the participating actors was considerable in these projects. From experience we know that a bad starting point may be adjusted through a successful decision-making process, even when the original idea was quite wrong. We also know that in many cases this does not happen.

This study demonstrated that there are many hurdles for any project. Democratic decision-making processes, particularly those which take long time, are complex and difficult to predict, and many will claim that this is a necessary part of democracy. If this is taken as a premise, the study suggested that the biggest potential for improvement lies in strengthening the analytical process. What would seem to be a reasonable compromise in front end analysis and quality assurance of major projects would be that the first step should be to identify and eliminate the worst alternatives. These are low hanging fruits and proper action can give a high reward with little effort. The next step should be to seek for good alternative concepts, but within reasonable limits, and not necessarily crave for the best, since the case will nevertheless be handed over to decision makers to conclude.

7. The paradox of strategic alignment: strategy and alignment of objectives are highlighted as essential concerns, but in most cases the internal logic of causalities and the probabilities of realization are erroneous

Alignment of objectives is the exercise to define the basic logical structure outlining the project by following the causal link from the basic needs of users and society, through defined goals to the delivery of project results (outputs), their outcome (effects) and long-term benefits after the project is terminated (purpose). This needs to be done before starting significant work on a project or programme. Unfortunately, this is not always done and can result in significant underperformance compared to expectations (Cooke-Davies, 2011).

Any large projects, and particularly major public investment projects, are initiated in order to produce benefits for their owners (society). Many authors have studied success factors and predictors of failure, notably Morris and Hough (1987), Pinto and Slevin (1988), Miller and Lessard (2000), Flyvbjerg et al. (2003), and Hopkinson (2007). The available literature provides several different answers to why things go wrong and what could bring success in projects. Earlier literature tended to focus narrowly on the outputs in terms of cost, time and quality, whereas more recent literature may offer a wider perspective. For example, Morris and Jamieson (2005) study the processes, practices and people issues involved in moving from corporate strategy to projects. Their results are promising, though only based on four case studies. A common feature is that when projects' strategic success is low, the problem possibly lies in the early phases of the project and at the governance level (the owner perspective).

Some studies on international development projects have provided insight in this area. An analysis of a large sample of such projects concluded that most of the uncertainties affecting these projects were internal and not contextual, for a large part associated with aspects of management and the fundamental project design (Samset and Haavaldsen, 1998). Consequently, the suggestion was that most of the problems ought to be met early, i.e. in the pre-study phase.

Youker (1999) concluded that the lack of shared objectives and agreement on the objectives of a project was one of the biggest problems facing international development projects. A study of alignment of objectives in development projects based on a sample of 30 international aid projects, concluded that most of the projects had design faults at all levels, and no projects were without faults. Typical problems identified were insufficient resources, and too many and unrealistically ambitious goals (Samset, 2006).

The same analysis was repeated on a sample of 17 large public investment projects in Norway (Andersen et al., 2014). A project strategy will always be a hierarchy of goals that are interlinked in cause-and-effect chains that illustrate the ambition levels for a project, as well as their realism. Objectives were analysed in terms of their internal causality, and ambition. Complex statements were broken down in several single objectives.

The study found that in most of the Norwegian projects the goals are consistent with the needs, but there were shortcomings when it comes to achieving reasonable levels of clarity and ambition, as shown in Fig. 8. For instance, when a project to

acquire defence equipment presents "*stability within the international legal system"* as a societal goal and a limited road construction project expects to result in "*increased settlement"*, we intuitively understand that the distance between cause and effect is too large and that the goals are too ambitious for the given project.

Fig. 8 compiles the findings from the study and breaks down the percentages of goals across the different goal levels. Of the total 152 goals presented by the 17 projects, by far most of these were defined as project outcomes with the majority of the remaining goals being societal goals. About a quarter of the project outcomes were in reality societal goals, while two thirds of the presented project outcomes actually were project outcomes. Also, a small portion of the social goals are completely unrealistic, while a small set of the project outcomes were in reality project outputs, i.e., specifying aspects of the project's deliverables. In total, none of the projects avoided erroneous definitions of goals, but they performed better than the international development studies mentioned above. But clarity seemed to be the largest problem. Five of the projects had in reality no societal goals whatsoever, while others had too many. One project had as many as seven societal goals. In such a case, the strategy is of little help to focus the efforts and clarify the purpose of the project. Regarding project outcomes, the majority of projects had 3–9 project outcomes, two projects even more than 10.

The purpose of formulating an objective is principally to clarify the direction for that which is sought. The scope also needs to be stated so one may know when an objective is attained. Multiple objectives may confuse if they all don't point in the same direction. This is particularly evident if the objectives also conflict with each other. Objectives should give rise to common understanding among and motivation of all parties involved in or affected by a project. On one hand, this means that objectives should be unambiguous and realistic. On the other hand, to motivate, they also have to be well founded, to the degree that they are accepted. Moreover, the objectives should limit the enterprise or the strategy. This means that the resources allocated and the results anticipated should correspond.

In looking at customary practice in planning projects, the threshold for improvement seemingly is very low and the possibilities of marked improvement accordingly are great.

Regardless, practice indicates a need for more concise formulation of objectives in the front end phases of projects, at any rate to establish common understanding of where a project is going and how it will get there.

8. The cost estimation paradox: the focus is on the final cost estimate (the budget), while early cost estimates are overlooked

We have already discussed how planners devote less attention to identifying the best conceptual solution than to improving tactical project success. This is understandable to some extent because planners find it easier to relate to tangible and quantified success criteria such as cost and time, than to multidimensional and qualitative assessments of societal benefits. However, the investment cost is tangible and concrete, and crucial both to the choice of concept and to tactical success. Although cost uncertainty is higher in the early stages, it too is tangible and manageable (e.g. Austeng et al., 2005). Planners should therefore be strongly committed to establishing a rough but realistic cost estimate in the early phase, for comparison with project benefits.

Under the auspices of the Concept research programme a study of cost estimates in projects' initial phase has been conducted (Welde et al., 2014). The study explored a sample of

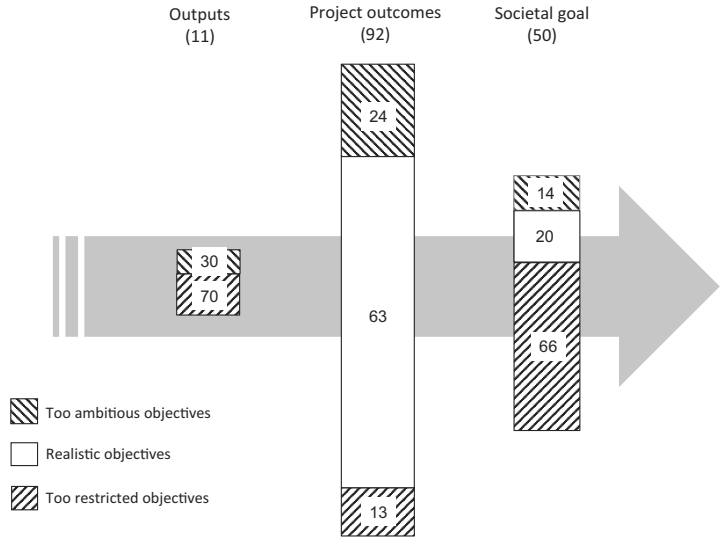


Fig. 8. Assessment of the goals in the sample of projects in terms of location in the goal hierarchy and their level of ambition (Samset et al., 2014a, 2014b).

12 projects to determine the basis for and how the first cost estimates came about and developed during the whole period from the first initiative that was taken until the project was approved by Parliament. As shown in Fig. 9 the first cost estimate in all 12 cases was far below of what was ultimately approved as the projects' final budget. The increase in cost estimates during the front-end phase ranged from +70% to almost +1300%, with an average of +650%. By comparison, the cost increase during the implementation phase was much less, and some projects were even completed below budget; the variation ranged from −19% to +186%.

The study is a first probe into the matter of early cost estimation. More research is needed to determine the extent of the problem and its implications. However, it indicates that initial underestimation may be significant and result in the approval of projects that otherwise should have been rejected in the early stages. The authors considered it likely that at least 5 of the 12 projects would have been screened out at an early stage if the first estimate had been at a realistic level as compared with what was the final cost. The question is of course hypothetical, but there is no doubt that underestimation of costs at an early stage can have dramatic implications for project selection and is probably a far more severe problem than cost overruns in the implementation phase. Hence, it is clearly a paradox that so little attention is devoted to the initial estimate.

The report discusses possible reasons for the substantial underestimation in early phases. An often used distinction is made between political, technical, and cognitive reasons (e.g. Flyvbjerg, 2005). It may be very difficult to prove that the cause is political, but in several of the projects there were clear indications that the first estimate was deliberately low in order to increase the chance of the project idea being considered. This corresponds well with other studies that have attempted to prove that costs are underestimated deliberately to make the projects appear more attractive (e.g. Flyvbjerg, 2007; Mackie and Preston, 1998; Wachs, 1987; Welde et al., 2014). Wachs (1989) discusses how the most effective planner is sometimes the one who can cloak advocacy in the guise of scientific or technical rationality. In other cases the total investment was split between several projects that would have to be approved separately in sequence. However, in the most of the projects there were also different cognitive reasons why costs had been underestimated up front. Over-optimism is a well-known phenomenon in cognitive research literature, see e.g. Kahneman and Tversky (1979). Further, the study discusses measures to reduce the problem of early underestimation, such as systematic recording of early cost estimates, the use of reference projects, of stochastic cost-estimation techniques, increased provisions for uncertainty to account for possible scope changes, and third party review at an early stage.

9. The paradox of disregarded analyses of costs and benefits: detailed estimation of cost and benefits is commonly done up front, but disregarded by decision-makers, who tend to emphasize other aspects

A substantial amount of resources is devoted in major investment projects to establish a decision basis. Detailed Cost–Benefit Analyses are often performed, and complex models are developed to simulate traffic volumes and other inputs to these analyses. However, there are indications that

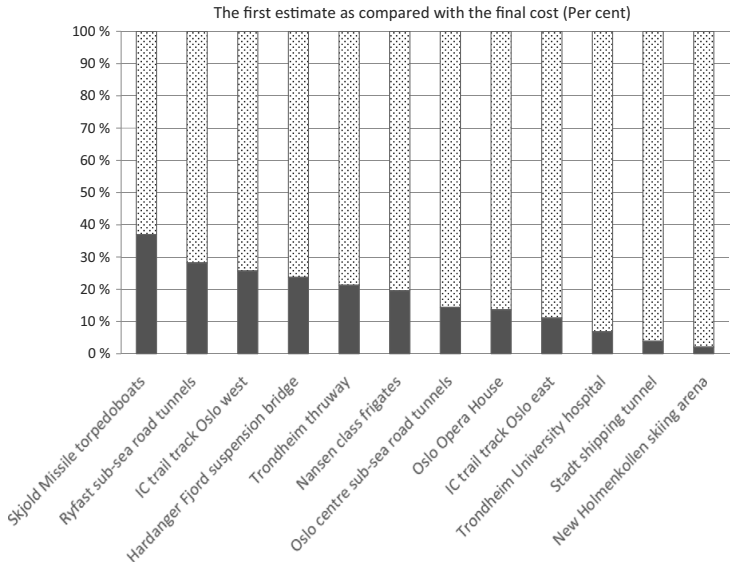


Fig. 9. The earliest cost estimate as a percentage of the final cost, for 12 Norwegian projects.
Source: Welde et al. (2014).

decision-makers have little confidence in Cost Benefit Analysis in Norway.

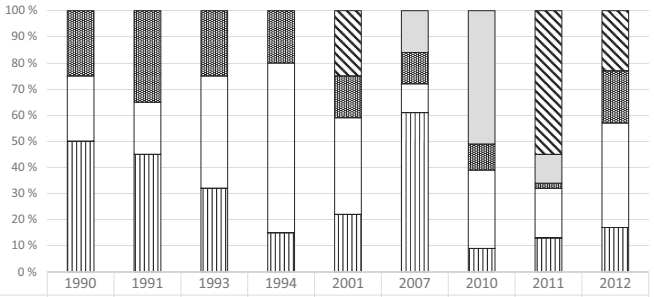
The transport sector is a special case. In this sector there is a long tradition of using Cost Benefit Analysis. A recent study conducted by the Concept research programme, Welde et al. (2013) studied the significance of Cost–Benefit Analysis in the final prioritization of road projects in Norway and Sweden, where the approaches to such analyses in the two countries are very similar and unit prices are of the same magnitudes. The study revealed that the Cost–Benefit ratio had no significant impact on the selection of projects in Norway. On the contrary, many unprofitable projects were realized, such as spectacular tunnels and bridges in sparsely populated areas. By contrast, in Sweden, the results of the Cost–Benefit Analyses had somewhat more influence on the selection of road projects. Clearly, in the case of Norway there must have been other factors that were more important but that were not included in the analyses.

One explanation for low confidence in the Cost–Benefit Analyses could be weaknesses and shortcomings in the methodology, see e.g. Næss (2006, 2012). The trend is however that more and more effects are included in the analysis, and the empirical basis for estimating realistic values is improving. See for example Vickerman (2008) on the inclusion of so-called wider economic benefits from transport infrastructure projects. Another explanation for low confidence could be strategic use of analyses to promote a desired result. One study, by Kvalheim (2014), examined a special case where nine Cost–Benefit Analyses had been made of one project, a shipping tunnel on the west coast of Norway. This study found a remarkable lack of consistency between analyses. The analyses were performed between 1990 and 2012, and the Cost–Benefit ratio varied from 0.2 (highly unprofitable) to almost 1.0, and even exceeded 2.0 (highly profitable) in an 'optimistic calculation' provided in one of the reports. The analysis reporting the most positive number was funded by local stakeholders, with no financial obligations. An interesting finding was that the relative weight put on different benefit components varied noticeably, as shown in Fig. 10. This underscores the credibility of such studies. By 2014, the tunnel project had still not been approved for funding or finally rejected.

Not all effects of an investment project may be quantified and expressed in monetary terms. Nevertheless, if they are relevant to the decision they should be systematically reviewed as much as the net present value. Norway is often regarded to be at the forefront internationally when it comes to including non-monetized impacts in Cost Benefit Analyses. However Bull-Berg et al. (2014) reviewed a practice regarding non-monetized impacts in more than 100 economic analyses in Norway. With a few important exceptions, their findings are rather discouraging. The section presenting non-monetized impacts in the economic analysis is characteristically short, and not based on transparent methodology and well-documented processes. The study concluded that there is substantial potential for improvement and a need for guidance.

The paradox in this case is that so much effort is devoted to the calculation of a net present value that decision-makers may not find useful or credible. Clearly, planners should focus more on non-monetized impacts in economic analyses, as well as other complementary analyses such as cost-effectiveness analysis, impact evaluation, and multi-target criteria analysis. In addition, competence requirements are crucial to ensure high-quality analyses.

The above situation is mirrored in the World Bank, which made wide use of Cost Benefit Analyses for decades to demonstrate its reputation as a knowledge bank and its commitment to measuring results and ensuring accountability to taxpayers. However, according to the World Bank (2010), the percentage of projects justified by a Cost Benefit Analyses has been declining, and the Cost–Benefit ratio is now rarely mentioned in policy documents. These results are explained by a decline in adherence to standards as well as increased difficulty in applying Cost Benefit Analyses in new sectors



	1990	1991	1993	1994	2001	2007	2010	2011	2012
Other benefits	0	0	0	0	25	0	0	55	23
Residual Value	0	0	0	0	0	16	51	11	0
Safety at sea	25	35	25	20	16	12	10	2	20
Coastal express line	25	20	43	65	37	11	30	19	40
Sea transport and fishery	50	45	32	15	22	61	9	13	17

Fig. 10. Percentage of the total monetized benefits in nine different Cost–Benefit Analyses of the Stad shipping tunnel, showing how much weight was placed on the various components.
Source: Kvalheim (2014).

where traditionally it has not been applied and where benefits can hardly be quantified. The situation is that economic assessments are not performed at all. The World Bank concludes that there is a need to recognize the difficulties in quantifying benefits, but at the same time quality, rigour, and objectivity must be ensured because poor data and poor analyses are misinforming and do not lead to improved results.

10. The paradox of "predict and provide": the tendency is to choose a "predict-and-provide" strategy rather than explore alternative solutions

Different perspectives can be taken when evaluating the need for an investment project. As discussed by Næss (2005), public planners tend to use a predict-and-provide approach. When confronted with capacity problems, the planners, who are often engineers, almost always recommend increased capacity based on estimates of future demand. However, unsurprisingly, there is often excess demand for public services and infrastructure offered free-of-charge to citizens. The need should not be defined narrowly as a need to increase capacity but rather as a need to solve the congestion problem. The latter allows for a variety of measures, including demand regulation, congestion pricing, and legal and informative measures, most of which are far cheaper than a construction project to expand capacity.

Our suggestion that needs should be considered in a broader perspective is supported by Odhage (2012), who studied early project planning in Swedish road projects. He found that the planners were never truly interested in finding and developing measures that would reduce the need for transport. This is obviously an example of path dependence, and Odhage asked the timely question 'Can one expect anything different from a process that is run by the transport administration and concerns transport issues?'

Further, in many cases there are political goals for a development that is quite the opposite of a predict-and-provide strategy. Næss (2005) distinguishes between (1) needs defined by national-level political objectives, (2) market-based needs as measured by demand or willingness-to-pay, and (3) the needs of different stakeholder groups. As noted, public planners tend to narrow down the identification of needs to the second demand, while ignoring the broader spectrum of needs, and even political goals to reverse the demand trend. A country with high ambitions to reduce the emission of greenhouse gases will view increased traffic (i.e. growing demands for roads) as a problem.[3] Similarly, in the university system, a purely demand-based approach probably would not necessarily lead to a distribution of graduates in line with society's need for expertise in different disciplines.

The paradox in this case occurs when needs and benefits assessments in public infrastructure projects are decoupled from overriding political priorities and goals, possibly because such overriding societal goals are conflicting and multidimensional. The result of this is that issues such as scaling and capacity of infrastructure projects, highly political choices, are left to planners, who (i) have a tendency to define the problem narrowly as absence of capacity, and (ii) use readily available estimates of demand as a reference for adjusting capacity. There is obviously a need for project owners (the government) to clarify what needs should be taken as a starting point for planners, and to express them as clear objectives for the project. Only if the development given by trend extrapolation is a clearly desired one can the predict-and-provide strategy be readily used in individual projects, as illustrated in Fig. 11.

11. The paradox of perverse incentives: public investments with no financial obligations for the target group may cause perverse incentives and result in counterproductive projects

The state often appears as a generous donor on behalf of taxpayers when financing projects that benefit specific groups or geographical regions. Such projects may be initiated either by the beneficiaries themselves or by the state out of pure altruism. There are indications that such projects often prove unsuccessful in strategic terms, and we should not be surprised by this. When a project does not entail financial obligations for recipients, there is no incentive to opt for the most socially beneficial or cost-effective alternative. Different actors may have a vested interest in certain projects being chosen.

The term perverse incentives refers to the situation where one or more actors are motivated to make choices resulting in a project that is a complete failure seen in retrospect. The theoretical basis is the principal-agent theory (e.g. Jensen and Meckling, 1976; Laffont and Martimort, 2002). There is a huge amount of literature on incentive problems in general, but less in relation to state-funded investment projects. A pivotal study in the field has been published by Ostrom et al. (2001), who demonstrate serious problems with perverse incentives in Swedish-funded aid projects that resulted in the waste of public funds and adverse side effects such as corruption.

Norway is a special case because the local government is financially weak and dependent on the state to finance local infrastructure. The State is rich due to revenues from the exploitation of petroleum resources. Whist and Christensen (2011) demonstrate how the early phase of state-funded

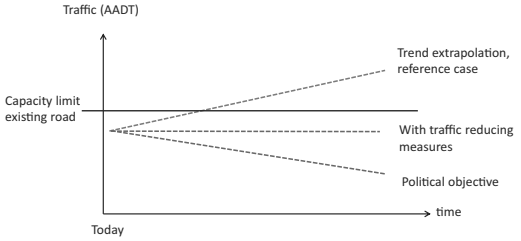


Fig. 11. Traffic development in different scenarios, illustrating that the need for an infrastructure project follows from the assumption that capacity should adapt to demand.
Source: Authors.

[3] In a separate study, Hagen (2010) discusses economic measures as accounting for external effects on the environment.

investment projects in Norway is often characterized by 'local rationality' and complex coalitions. Samset et al. (2014a, 2014b) explored the phenomenon of perverse incentives in nine Norwegian state-funded projects, to illustrate how perverse incentives might occur, what the causes and consequences might be, and what could be done to avoid them. One aid project served as a reference case to demonstrate how wrong things can go. The study revealed that half of the Norwegian projects scored very poorly, particularly in a strategic perspective (Fig. 12). Some of these projects clearly would not have been prioritized had the recipient been required to contribute to the funding. Several projects were classed as supersized because they were 'free-of-charge'. Moreover, Samset et al. (2014a, 2014b) found that costs were being underestimated and benefits overestimated in advance.

The problem of perverse incentives is twofold: (1) actors who act out of self-interest, and (2) a financing party that fails to reveal that. Measures to solve or mitigate the problem should therefore also be twofold: (1) aligning recipients' objectives with national objectives, through requirements such as co-financing and local risk taking, and (2) reducing the information asymmetry by introducing, for example, by information control, external review, and public hearings. The Norwegian quality assurance regime is thus a measure that is expected to reduce the problem of perverse incentives.

12. The paradox of myopic decisions: long-term viability is the intention but the planning horizon is too short, resulting in sub-optimal choices that one will regret later

Probably the most crucial strategic success criterion for an investment project is that it is viable and sustainable, i.e. that project net benefits are likely to continue in the long run (OECD, 2000).

Viability can only be determined in the very long run. Samset (2012) studied 10 projects from history, and found that only a few were still considered highly successful and thus viable more than 100 years after completion, whereas others had been closed down after a short time. Needs and priorities in society may change over the years, and therefore a project's viability is contingent upon its ability to adapt to changing needs. Ironically, one of the most viable projects in the study was the Eiffel Tower, which was built for no other purpose than to be an exhibition object to showcase France as a leader in science and technology, but which later became one of the greatest tourist attractions in the world.

Since viability can only be determined in the long run, an assessment of viability ex ante must have a long-term perspective and the planner must be able to think creatively about possible future scenarios. It is not sufficient that the project is feasible and relevant on the opening day; planners must consider whether it will continue to be so throughout its lifetime. Lædre et al. (2012) studied 24 appraisal reports of major public projects from the periods 2005–2011 with respect to their assessments of viability. The results were rather disappointing: needs and benefits were most often assessed in a short-sighted and static perspective; trends were extrapolated without discussing alternative scenarios; most attention was devoted to tangible effects, ignoring non-monetized impacts; and significant risk factors, such as political risk, were not identified and discussed. Such practice may lead to myopic decisions, which we are likely to regret in the future, as illustrated in Fig. 13.

However, Lædre et al. (2012) also noted that no single analytical tool is able to comprehend all aspects of a project's viability ex ante. In particular, a Cost–Benefit Analysis, although intending to capture all economic impacts of a project, cannot provide sufficient analysis of viability, one important reason being the use of a discount rate. Therefore, in order to assess viability properly, several complementary tools combining quantitative and qualitative approaches are necessary. A separate study by Hagen (2011) goes further into the question of how the Cost–Benefit Analysis, through the use of a discount rate, leads to short-termism and neglect of future generations. However, Hagen also shows that it may be appropriate to use a decreasing discount rate over time. This would in fact increase the planning horizon and thus mitigate the problem.

The paradox in this case is that the emphasis on viability as a success criterion is far from reflected in project appraisals. Projects that are meant to last for decades and sometimes centuries may have significant impact on economic, environmental, and social development, yet they are still assessed in a short-term and static perspective. Lædre et al. (2012) offer

Project	Type of project	Year completed	Total cost (mill. NOK)	State-funded without liabilities for beneficiaries	Operational success	Tactical success	Strategic success
Hvaler-tunnelen	Subsea road tunnel	1989	200	No	Yes	Limited	Limited
Linesøya	Bridge	2011	250	Minimal	No	No	No
St. Olavs Hospital	Hospital	2014	13 000	Minimal	No	Limited	Limited
Turkana Fisheries (Kenya)	Development aid	1990	1 500	Yes	No	No	No
OL Lillehammer	Sports event	1994	7 500	Yes	Limited	Yes	No
E16 Lærdalstunnelen	Road tunnel	2000	1 050	Yes	Yes	Limited	No
Lofast	Subsea road tunnel	2007	1 367	Yes	Yes	Limited	No
Rock city	Cultural building	2013	50	Yes	Limited	No	No
Stad skipstunnel	Shipping tunnel	Not yet	1 800	Yes	n/a	n/a	Not likely

Fig. 12. Selected findings from Samset et al. (2014a, 2014b), one aid project and eight Norwegian state-funded investment projects without liabilities for the target group.

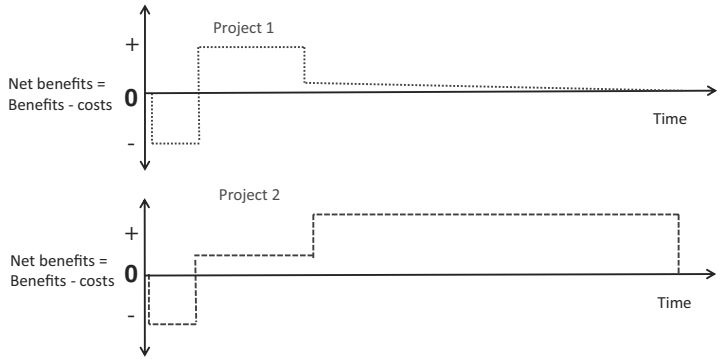


Fig. 13. Illustration of myopic decisions. Two projects with identical investment cost have different net benefit flows throughout their life-time. In a long-term perspective it is clear that project 2 is more viable, but a myopic planner would emphasize short-term effects and choose project 1. For example, investments in preparedness and prevention capacities are often very low, something that one regrets later when a disaster strikes.
Source: Authors.

some recommendations for how to obtain a broader and more long-term perspective in project appraisals. They involve shifting the analyst's attention away from detailed estimations of investment cost to estimating future benefit flows and corresponding risk. Undoubtedly, evaluating a project's viability ex ante can be challenging, but the alternative of finding out about its unviability too late is worse.

13. Discussion

Governance regimes for major investment projects comprise the processes and systems that need to be in place on behalf of the financing party to ensure successful investments. What happens during the front end phase is essential. Peter Morris (2011:7) writes that "It is evident from an extensive amount of research that management of the front-end definitional stages of projects is of overwhelming importance to their ultimate outcome yet we have little empirical data to suggest how best management competencies here should be improved."

Project governance has only recently become an issue in the project management community. In order to move forward in this field we have to find answers to what would be the optimal mix of regulations, economic means and information in improved governance regimes. What seems to be an issue for the project management community is to lift their perspective beyond the delivery of the project itself and onto the broader issues of the project's utility and effects. It is obviously not only about the quality of analyses up front but also about decision processes. To arrive at the optimal conceptual solution based on rational analysis is of little worth if it is not the one chosen.

The Concept programme did a pilot project on a sample of cases to illustrate this (Samset, 2008a, 2008b), which was followed up with a more in depth study to explore the quality and interaction between analysis and decisions during the front end phase (Whist and Christensen, 2011) and a broader follow-up of the pilot (Samset and Volden, 2013a, 2013b). The result is displayed in Fig. 14, where the flaws for the individual projects are plotted with "X". The summary row at the bottom are marked to signify whether they are considered relevant as seen in relation to needs and priorities in society (white colour) or not (black). Each project is represented with one column. The columns are sorted from left to right according to the observed number of flaws. The resulting pattern suggests that the least relevant projects have a lot of flaws in their analytic and decision making processes (between five and ten). The ones that are regarded relevant on the other hand have much less flaws (between one and four).

The studies concluded that there is a strong tendency to choose the initial concept and stick to it, almost regardless of how bad it is. Also, there is an overwhelming inertia. Once the train has been set in motion — it is always impossible to stop. This goes a long way to explain the red projects on the left hand side. Further there is a third common tendency, that incremental improvements of an inferior solution are preferred rather than fundamental change.

On the other hand experience also suggests that the opportunity space is usually larger than envisioned — and it is often largely unexplored. What was evident, however, was that the green projects seemed to have been exposed to more vigorous analyses and decision processes that were less affected by disagreements, political preferences, lengthy processes and repeated playoffs in the political decision processes.

14. Conclusions

This paper reports from several in-depth case studies of major public projects, and identifies a number of paradoxes that could guide further research. In various ways the paradoxes point to two types of problems, i) problems of efficiency in terms of delays and cost overrun, and ii) more fundamental problems that have to do with the project's strategic success (choosing the wrong concept). Project management as a discipline should be concerned with both problems. To quote Peter Morris: "The discipline needs to be less inward looking: more relevant, not just to the sponsor's needs but to society's challenges in general. We can foresee several changes in the

	Project number																							Sum
Analysis	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
Alternative concepts have not been scrutinized	X	X	X	X	X	X	X	X	X			X		X	X	X			X		X			15
Strategic underestimation of expected costs	X	X	X	X		X	X	X	X	X	X		X				X							12
Inadequate/limited analysis of problems and needs	X	X	X	X		X	X		X	X	X		X				X							11
Lack of realistic objectives and justification	X	X	X	X		X	X	X	X			X				X			X					11
Tactical splitting up and sequencisng of project	X				X		X		X	X				X	X			X	X		X			10
Predictable surprises not taken into account	X			X		X				X		X		X										6
Decisions																								
Disagreement regarding objectives and justification	X	X	X	X	X	X		X	X		X	X	X			X		X						13
Expert advice overruled by political preferences	X	X	X	X	X	X		X		X	X		X			X	X	X						13
Long lasting front end phase with shifting priorities	X	X	X		X		X	X		X		X		X			X			X		X		12
Repeated playoff in political decision process	X	X			X								X		X								X	6
Perverse incentives - benefits without liability				X			X	X							X					X				5
Political horse-trading between competing parties			X		X						X							X						4
Sum	10	8	8	8	7	7	7	7	6	6	5	5	5	4	4	4	4	4	3	2	2	1	1	
Relevance of the project																								

Fig. 14. There is a consistent tendency that projects that are considered relevant have less flaws in the analysis and decision making processes up front (Samset, 2008a, 2008b).

years ahead in the ways projects and programs will be managed, but the obvious immediate needs are to focus more on improving sponsor value and on shaping the context in which projects and programs are formed and implemented" (Morris, 2013:23).

Many of the problems facing major public investment projects can be interpreted in terms of deficiencies in the analytic or the political processes preceding the final decision to go ahead, and the complexity and uncertainties affecting these processes. In particular, the fundamental problems with strategic success could typically be traced back to deficiencies in the earliest preparatory phases of the project. The role of the front-end phase in ensuring project success is therefore crucial, as highlighted in the literature (Merrow, 2011; Morris, 2013).

Project governance is the processes, systems, and regulations that the financing party must have in place to ensure that projects are successful, strategically as well as tactically. Many organisations have introduced stage-gate phase models, also the Norwegian Ministry of Finance, who introduced a QA scheme to ensure the best choice of concept (QA1) and efficient project implementation (QA2) in year 2000. Our research indicates that QA2 has already led to improved cost control. It is still too early to conclude that QA1 has improved the choice of conceptual solutions and projects' strategic success, but there is evidence to suggest that an independent review of the project appraisal documents at a very early stage has a positive effect. There are many fundamental challenges that will have to be dealt with, such as tactical budgeting in local communities and responsible agencies at various levels, which is done in order to increase the chance to obtain government funding for a project. Another challenge is to ensure a transparent and democratic process and avoid adverse effects of stakeholder's involvement and political bargaining. But also to make the process predictable is a major challenge. The QA regime attempts to remedy these problems.

One salient conclusion from the research is that ex post evaluation should be an essential element in any project governance scheme. When a project succeeds at all levels, it should be imperative to ask what was done right. Correspondingly, one should learn from mistakes. However, experience shows that the use of evaluations for learning purposes is limited, and this is particularly true in the public sector (Samset and Christensen, 2012). The tendency is to look ahead with the concern of how to spend next year's budget, rather that look in the rear mirror to learn from experience.

As a lead part of the current trailing research on Norwegian public projects, the Concept research programme has since its inception been concerned with project evaluation and evaluation methodology as evidenced in several studies, including those by Olsson (2005), Andersen et al. (2007), and Volden and Samset (2013). The latter is a summing-up of four pilot evaluations of so called QA projects. It recommends that systematic ex post evaluations of public investment projects should be carried out to learn from experience, not least how they perform in a strategic perspective, with the aim to improve public investment projects in the future. Under the auspices of the programme therefore, a number of the major investment projects are now being evaluated, and this will continue in the years to come. Fig. 15 shows some main results for the first nine projects.

Clearly, projects may fail even when formal rules for planning and decision making have been adhered to. Democratic decision-making processes, particularly the long lasting ones, are complex and the outcome difficult to predict. Many will claim that this is a necessary part of democracy. If this is taken as a premise, one could conclude that the biggest potential for improvement lies in strengthening the analytical process, as well as making decision processes transparent.

Conflict of interest

None.

Project	Efficiency (tactical project success)	Effectiveness (strategic project success)
Svinesund national border control facility	☆☆☆	☆☆☆
Sandvika-Asker inter-city rail line	☆	☆☆
Momarken-Sekkelsten, section of a highway	☆☆	☆☆
Skjold class missile torpedo vessels	☆☆	☆
Eiksund road system	☆☆☆	☆☆☆
Lofast road system	☆☆☆	☆☆
E6 Riksgrensen-Sv.skogen, section of a highway	☆☆☆	☆☆☆
NAV ICT basic project	☆☆☆	☆☆
Svalbard Science Center	☆☆☆	☆☆

Fig. 15. Main results from ex post evaluations of nine Norwegian investment projects (three stars = high success, two = medium success, one = unsuccessful). For more detailed results, see the evaluation reports, available on www.ntnu.no/concept.

References

Andersen, B., Bråthen, S., Fagerhaug, T., Nafstad, O., Næss, P., Olsson, N., 2007. Effektvurdering av store statlige investeringsprosjekter (Impact assessment of major public investment projects). Concept report no. 19.

Andersen, B., Samset, K., Austeng, K., 2014. To which extent do projects explore the opportunity space? A study of conceptual appraisals and the choice of conceptual solutions. Int. J. Manag. Proj. Bus. Emerald 7 (3), 473–492.

Austeng, K., Torp, O., Midtbø, J.T., Helland, V., Jordanger, I., 2005. Usikkerhetsanalyse — Metoder (Uncertianty analysis — methodology). Concept report no. 12.

Bevir, M., 2013. Governance: A Very Short Introduction. Oxford University Press, Oxford, UK.

Bull-Berg, H., Volden, G.H., Grindvoll, I.L.T., 2014. Ikke-prissatte virkninger i samfunnsøkonomisk analyse. Praksis og erfaringer i statlige investeringsprosjekter (Non-monetized impacts in economic analysis. Practice and lessons from public investment projects). Concept report no. 38.

Christensen, K., 1985. Coping with uncertainty in planning. J. Am. Plan. Assoc. 51 (1), 63–73.

Cooke-Davies, T., 2011. Front-end alignment of projects — doing the right project (Making Essential Choices with Scant Information. Front-End Decision Making in Major Projects). In: Williams, T (Ed.), Palgrave Macmillan, UK, pp. 68–85.

Dosi, G., 1997. Opportunities, incentives and the collective patterns of technological change. Econ. J. 107 (444), 1530–1547.

Flyvbjerg, B., 2005. Policy and planning for large infrastructure projects: problems, causes, cures. World Bank Policy Research Working Paper 3781 (December).

Flyvbjerg, B., 2007. Cost overrun and demand shortfalls in urban rail and other infrastructure. Transp. Plann. Technol. 30 (1), 9–30.

Flyvbjerg, B., Bruzelius, N., Rothengatter, W., 2003. Megaprojects and Risk; An Anatomy of Ambition. Cambridge University Press, Cambridge.

Hagen, K.P., 2010. Markedsorienterte styringsmetoder i miljøpolitikken (Market oriented approaches to environmental policy). Concept report no 24.

Hagen, K.P., 2011. Verdsetting av fremtiden. Tidshorisont og diskonteringsrenter (Valuing the future. Time horizon and discount rates). Concept report no. 27.

Hopkinson, M., 2007. Ten causes of megaproject failure. APM Conference; The Business of Projects, London, 30–31 October 2007.

Jensen, M., Meckling, W., 1976. Theory of the firm: managerial behavior, agency costs and ownership structure. J. Financ. Econ. 3 (1976), 305–360.

Jessen S.A., Project Leadership, Bookboon.com, 2012, Ventus Publishing ApS.

Jordan, G., Lee, I., Cawsey, G. (Eds.), 1988. Learning From Experience. A Report on the Arrangements for Managing Major Projects in the Procurement Executive. Ministry of Defence, London, HMSO.

Kahneman, D., Tversky, A., 1979. Prospect Theory: An Analysis of Decision under Risk. Econometrica 47 (2), 263–291.

Kvalheim, E.V., 2014. Nyttekostnadsanalyse av Stad skipstunnel. En gjennomgang og reanalyse av tidligere analyser (Cost–Benefit Analysis of Stad Shipping Tunnel. A Review and Reanalysis of Previous Analyzes). (Master thesis). Norwegian University of Science and Technology.

Lædre, O., Volden, G.H., Haavaldsen, T., 2012. Levedyktighet og investeringstiltak. Erfaringer fra kvalitetssikring av statlige investeringsprosjekter (Sustainability and public investments. Lessons from major public investment projects). Concept report no. 29.

Laffont, J.-J., Martimort, D., 2002. The Theory of Incentives: The principal-agent Model. Princeton University Press, USA (440 pp.).

Mackie, P., Preston, J., 1998. Twenty-one sources of error and bias in transport project appraisal. Transp. Policy 5 (1), 1–7.

Margolis, S.E., Liebowitz, S.J., 2000. Path dependence, lock-in and history. Encycl. Law Econ. 11 (1), 205–226.

Marshall, D., Ritchie, B., 1993. Business Risk Management. Chapman & Hall, London 0-412-43100-9.

Merrow, E.W., 2011. Industrial Megaprojects: Concepts, Strategies and Practices for Success. Wiley (384 pp.).

Miller, R., Hobbs, B., 2009. The complexity of decision-making in large projects with multiple partners: be prepared to change. In: Williams, T., et al. (Eds.), Making Essential Choices with Scant Information. Palgrave MacMillan, Basingstoke, UK.

Miller, R., Lessard, D., 2000. The Strategic Management of Large Engineering Projects: Shaping institutions, Risk and Governance. MIT Press, Cambridge Mass.

Morris, P.W.G., 1994. The Management of Projects. Thomas Telford, London, UK.

Morris, P.W.G., 2009. Implementing strategy through project management: the importance of managing the project front-end. In: Williams, T.W., Samset, K., Sunnevaag, K.J. (Eds.), Making Essential Choices with Scant Information Front-End Decision Making in Major Projects. Palgrave Macmillan, pp. 39–67.

Morris, P.W.G., 2011. Managing the front-end: back to the beginning. Project Perspectives vol. XXXIII. International Project Management Association, pp. 4–9.

Morris, P.W.G., 2013. Reconstructing project management reprised: a knowledge perspective. Proj. Manag. J. 44 (5), 6–23 (October).

Morris, P.W.G., Hough, G.H., 1987. The anatomy of major projects. A Study of the Reality of Project Management. John Wiley & Sons, Chichester.

Morris, P.W.G., Jamieson, A., 2005. Moving from corporate strategy to project strategy. Proj. Manag. J. 36 (4), 5–18.

Müller, R., 2009. Project Governance. Gower Publishing Limited.

Næss, P., 2005. Bedre behovsanalyser. Erfaringer og anbefalinger om behovsanalyser i store offentlige investeringsprosjekt (Needs analysis in major public investment projects. Lessons and recommendations). Concept report no. 5.

Næss, P., 2006. Cost–benefit analysis of transport investments: neither critical nor realistic. J. Crit. Realism 5, 32–60.

Næss, P., 2012. Critical view of cost–benefit analyses. Presentation at: Concept Symposium 2012: Valuing the Future — Public Investments and Social Return (2012), Norway.

Odhage, J., 2012. Åtgärdsvalsstudie — en ny planeringsaktivitet för bättre lösningar på transportrelaterade problem. Erfarenheter från de första testfallen. KTH Architecture and the Built Environment, Stockholm.

OECD, 2000. Glossary of key terms in evaluation and results based management. Organisation of Economic Cooperation and Development, Paris.

Olsson, N.O.E., 2005. Hvordan trur vi at det blir? Effektvurderinger av store offentlige prosjekt (Up-front conjecture of anticipated effects of major public investment projects). Concept report no. 7.

Ostrom, E., Gibson, C., Shivakumar, S., Andersson, K., 2001. Aid, insentives, and sustainability. An institutional analysis of development cooperation. SIDA studies and evaluation 02/01. Swedish International Cooperation Agency, Stockholm.

Pinto, J., 2014. Project management, governance, and the normalization of deviance. Int. J. Proj. Manag. 32 (3), 376–387.

Pinto, J.K., Slevin, D.P., 1988. Project success: Definition and measurement techniques. Proj. Manag. J. XIX (1) (February).

Samset, K., 2006. Design of high-uncertainty projects in international aid. PROMAC 2006. Australian Institute of Project Management, Sydney (September).

Samset, K., 2008a. Concepts and decision — rationality and chance. Some paradoxes and reflections. Presentation to the 3rd International Symposium on Project Governance. The Norwegian University of Science and technology/Concept, Trondheim, Norway (September).

Samset, K., 2008b. How to overcome major weaknesses in mega-projects: the Norwegian approach. In: Priemus, H., et al. (Eds.), Decision-Making on Mega-Projects Cost–Benefit Analysis, Planning and Innovation. Edward Elgar Publishing Limited, UK.

Samset, K., 2010. Early project appraisal. Making the Initial Choices. Palgrave Macmillan, UK.

Samset, K., 2012. Beforehand and long thereafter. A Look-back on the Concepts of Some Historical Projects. Ex Ante Academic Publisher.

Samset, K., 2014. Strategic and tactical performance of mega-projects — between successful failures and inefficient successes. In: Priemus, H., van Wee, Bert (Eds.), International Handbook on Mega-Projects. Edward Elgar Publishing, UK (Chapter 2).

Samset, K., Haavaldsen, T., 1998. Uncertainty in Decision Making. Factors and Strategies Affecting Uncertainty in Planning and Management of Projects. Canadian Journal of Development Studies, Ottawa.

Samset, K., Christensen, T., 2012. Evaluering av prosjekter ex ante og ex post — og beslutningsprosessenes kompleksitet og betydning. chapter in Evaluering. Tradisjoner, praksis, mangfold, Fagbokforlaget, Oslo, 2013.

Samset, K., Volden, G.H., 2013a. Investing for impact. Lessons with the Norwegian State Project Model and the first investment projects that have been subjected to external quality assurance. Concept report no. 36.

Samset, K., Volden, G.H., 2013b. Major projects up front: analysis and decision — rationality and chance. Paper presented at the IRNOP conference, Oslo 2013.

Samset, K., Volden, G.H., 2015. Front end governance of major public projects. Lessons with a Norwegian Quality Assurance scheme. Int. J. Arch. Eng. Constr. (forthcoming).

Samset, K., Volden, G.H., Welde, M., Bull-Berg, H., 2014a. Mot sin hensikt. Perverse insentiver — om offentlige investeringsprosjekter som ikke forplikter (Perverse incentives and counterproductive investments. Public funding without liabilities for the recipients). Concept report no. 40.

Samset, K., Andersen, B., Austeng, K., 2014b. To which extent do projects explore the opportunity space? A study of conceptual appraisals and the choice of conceptual solutions. Int. J. Manag. Proj. Bus. 7 (3), 473–492.

Shenhar, A., Dvir, D., Levy, O., Maltz, A.C., 2001. Project success: a multidimensional strategic concept. Long Range Plan. 34 (6), 699–725 (December).

USAID, 1980. Design & Evaluation of Aid-assisted Projects. US AID, Washington (264 pp.).

Vickerman, R., 2008. Cost–benefit analysis and the wider economic benefits from mega-projects. In: Priemus, H., Flyvbjerg, B., van Wee, B. (Eds.), Decision-making on Mega-projects: Cost Benefit Analysis, Planning and Innovation. Edward Elgar, Cheltenham, pp. 66–83.

Volden, G.H., Samset, K., 2013. Etterevaluering av statlige investeringsprosjekter. Konklusjoner, erfaringer og råd basert på pilotevaluering av fire prosjekter (Evaluating public investment projects. Lessons and advice from pilot evaluations of four projects). Concept report no. 30.

Wachs, M., 1987. Forecasts in urban transport planning: uses, methods, and dilemmas. Clim. Chang. 11 (1–2), 61–80.

Wachs, M., 1989. When planners lie with numbers. J. Am. Plan. Assoc. 55 (4), 476–479.

Welde, M., Eliasson, J., Odeck, J., Börjesson, M., 2013. Planprosesser, beregningsverktøy og bruk av nytte–kostnadsanalyser I vegsektor. En sammenligning av praksis i Norge og Sverige (Planning, analytic tools and the use of cost–benefit analysis in the transport sector in Norway and Sweden). Concept report no. 33.

Welde, M., Samset, K., Andersen, B., Austeng, K., 2014. Lav prising — store valg. En studie av underestimering av kostnader i prosjekters tidligfase (Low estimates — high stakes. A study of underestimation of costs in projects' earliest phase). Concept report no. 39.

Whist, E., Christensen, T., 2011. Politisk styring, lokal rasjonalitet og komplekse koalisjoner. Tidligfaseprosessen i store offentlige investeringsprosjekter (Political control, local rationality and complex coalitions. Focus on the front-end of large public investment projects). Concept report no. 26.

Williams, T., 2008. Decisions made on scant information: Overwiew. In: Williams, Samset, Sunnevaag (Eds.), Making Essential Choices with Scant Information. Palgrave Macmillan.

World Bank, 1996. Evaluation Results 1994. The International Bank of Reconstruction and Development, Washington DC.

World Bank, 2010. Cost–benefit analysis in World Bank projects. Independent Evaluation Group. The World Bank.

Youker, R., 1999. The difference between Different Types of Projects, paper presented at The Project Management Institute 1999 Conference in Philadelphia, PA. USA.

Paper II

Public funding, perverse incentives, and counterproductive outcomes

Gro Holst Volden
Department of Civil and Environmental Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

Received 19 December 2017
Revised 11 July 2018
Accepted 4 August 2018

Abstract

Purpose – The purpose of this paper is to explore the adverse incentives at the front end of government-funded projects with concentrated benefits and no liabilities for the privileged groups. In particular, the author discusses the risk of perverse incentives of the types typically found in the development aid sector that results in counterproductive outcomes.

Design/methodology/approach – The paper uses a simple conceptual framework based on agency theory. A qualitative, case-based approach with purposive sampling was chosen for the empirical part of the study. Eight Norwegian projects were selected because incentive problems were to be expected, and one development aid project served as a reference case.

Findings – The paper finds that low strategic project success corresponded well with the terms of financing. There were clear indications of agency problems, in three cases to the extent that the incentives turned perverse. The paper concludes with a discussion of relevant measures to prevent the emergence of perverse incentives.

Originality/value – The paper contributes to an improved understanding of the incentives related to public project initiation and selection, which is an under-researched topic and generally not included in formal project governance schemes. The research should therefore be useful to scholars as well as practitioners within the field of project governance.

Keywords Project governance, Agency theory, Public projects, Front-end management, Government funding, Perverse incentives

Paper type Research paper

Introduction

Government-funded projects, such as transportation infrastructure, public buildings and major sporting events, are normally intended to serve some overall societal goal and ultimately to benefit the whole nation. However, in many cases, the benefits are relatively concentrated in specific groups or regions (e.g. local public goods). There may be good reasons why such projects are government funded, such as to promote redistribution or provide insurance or external benefits. Nevertheless, such projects will appear as "gifts" to a privileged group that is often heavily involved in preparing the project proposal. The starting point for this study was our general impression that such projects often score low on relevance and viability when seen in retrospect. They end up being oversized and expensive, and in some cases, it is not even clear whether they fulfill a genuine need.

This paper is based on a research project funded by the Concept research program and retrieved from www.ntnu.no/concept. Preliminary results were presented in a Norwegian report, Samset *et al.* (2014) and later in a conference paper presented at IRNOP 2015 (Volden and Samset, 2015). The author would like to thank Professor Knut Samset for many interesting discussions and useful comments on an earlier draft.

International Journal of Managing Projects in Business
Emerald Publishing Limited
1753-8378
DOI 10.1108/IJMPB-12-2017-0164

Public projects have varying reputations. The problem of cost overrun is particularly well documented (Flyvbjerg *et al.,* 2003; Morris and Hough, 1991; van Wee, 2007). Equally serious, but less studied, is the problem that occurs when projects do not meet the expectations of users and society. In extreme cases, the investment is wasted. Flyvbjerg *et al.* (2003) discussed the consistent problem of benefit shortfalls in the transport sector. Moreover, they noted that environmental and other negative side effects are systematically downplayed, whereas regional economic benefits are overestimated. Solberg and Preuss (2007) described how major sporting events are often justified by tourism-related and other economic impacts but fail to realize such effects. Furthermore, some authors note that many ICT projects do not meet the expectations of users and end up being abandoned or reworked (Cicmil and Braddon, 2012; Pinto, 2006).

Cantarelli *et al.* (2010) offered four explanations for project failures, each of which may be relevant to varying degrees in specific projects, but they generally reinforce each other: technical, psychological, economic and political explanations. In this study, we focus on the latter type, which occurs when certain stakeholders deliberately present a biased business case in order to increase the chance that a specific project will be selected. We use agency theory to explain the relationship between the government as the principal and the privileged group as the agent and demonstrate that the terms of financing can create a serious conflict of interest at the front end and cause projects to fail strategically (including the wrong projects being selected).

Samset (2003) argued that to be truly successful, projects must perform well tactically and strategically, not just operationally. In recent years, several authors have highlighted the importance of taking a holistic and "big picture" perspective on projects (Morris, 2013; Shenhar, 2004; Williams and Samset, 2010; Zwikael and Smyrk, 2012). They have also highlighted the crucial role of the front-end phase of projects. This is the stage when the project is justified, and key assumptions are made. Poor project ideas can also be screened out at no cost at this stage. However, in practice, the acceptance of a project concept at this early stage may be almost impossible to reverse due to the expectations it generates. Cantarelli *et al.* (2012) used the term "lock-in" to describe the situation in which decision makers are *de facto* committed before the formal decision to build. van Wee and Rietveld (2013) found it very likely that the extent of cost overruns reported in the international literature is greatly underestimated because most studies compare the final cost with the formal budget. Instead, they should have compared the final cost with the estimate at the time of *de facto* approval. Andersen *et al.* (2016) explored 12 Norwegian projects through their earliest phases and showed that the increase from the initial estimate to the formally approved budget was significant (on average 350 percent) and many times higher than the increase from the approved budget to the final cost. These results have been confirmed by a more recent study from Norway that includes a larger sample but is limited to road projects (Welde and Odeck, 2017). However, none of the authors of the abovementioned studies asked explicitly who proposed the project or how it was financed (e.g. whether it was 100 percent government funded or had some level of co-financing).

As noted by Samset and Volden (2016), addressing the front-end phase and securing the strategic performance of public projects often proves to be a highly complex matter. In this paper, we only explore one specific factor that may explain why public projects fail strategically, namely, the terms of financing. When a group views a project as free of charge, the group's perspective will be positive as long as the benefits that accrue to them do not become negative. Even in cases when the target group is largely indifferent, there may be an intermediary party that has much to gain. The experiences gained from development aid projects are particularly useful for exploring this problem, which is also where we find the most extreme cases of perverse incentives.

The paper is organized as follows. In the next section, we review the literature relevant to incentives at the front end of public projects. Thereafter, we present the framework for analysis before describing the data that relate to our nine case projects. The main findings are presented and then further discussed in the final sections of the paper.

Extant literature

Agency theory

Agency theory originated in economics; it pertains to a situation in which a principal depends on an agent to achieve his or her goals, but the agent may pursue a different objective and thus act in his or her self-interest (Eisenhardt, 1989; Laffont and Martimort, 2002; Tirole, 1994). Due to external uncertainty, it is impossible to know for certain whether the agent is acting in the best interests of the principal. Thus, a problem is generated by the combination of conflicting goals and asymmetric information.

In general, there are two types of agency problems. One type is moral hazard problems, which typically occur when an agent is guaranteed a benefit regardless of whether he or she exerts the proper level of effort. This approach makes contractors inefficient, insured people more careless, and, we expect, a group that can freely acquire a new piece of infrastructure cares less about its value for money. The typical remedy is to ensure that the agent bears some of the costs of his or her actions (i.e. by introducing an incentives-based scheme). Alternatively, monitoring and control systems can be established to overcome the asymmetry of information. The second type of agency problem relates to adverse selection, which can occur in a situation in which a choice between alternatives must be made by the principal under uncertainty, and the agent, who knows the quality of each choice, may be motivated to offer the principal the poorer alternative. Again, the solution is to design proper incentive schemes and/or to invest in information. If the parties meet regularly, learning and reputation can also work as a disciplining factor. The optimal solution in each case depends on, *inter alia*, the seriousness of the goal conflict, the risk level, the agent's risk attitude and ability to control risk, and the cost of obtaining information.

The term perverse incentives refers to agency problems so severe that they yield outcomes in the opposite direction of the intention – that is, more negative than positive. A prominent example is described by Vann (2003) as "the great Hanoi rat massacre." In 1902, Hanoi was facing the bubonic plague from rats that had spread throughout the city. To address the problem, the government decided to pay a bounty for each rat killed. The rat's tail had to be provided as evidence. At first, the scheme was successful, but the rat hunters soon realized that they would be better off keeping the tailless rats alive to breed more rats for their tails. Rat farming became popular, and the problem went from better to worse. The authorities' use of a bounty thus had the opposite effect of what was intended.

Agency theory was originally used to describe the relationship between the owner and managers of firms, but it can be applied to a variety of situations within and between organizations. Tirole (1994) discussed how the theory of incentives can be helpful to understand the public sector. In this sector, the incentive problem is partly related to the risk of "capture," which stems from officials' discretionary power. Therefore, monitoring and control is crucial to ensure accountability.

Agency and the project governance literature

Agency theory is a logical starting point for studies of modern corporate governance, of which project governance is often seen as a subset (Müller, 2009; Müller and Turner, 2005). A key issue is to ensure that the implementing agent will act in conformity with the interests of the owner. However, the literature on project governance is still fragmented, and different perspectives have been used in different studies (Ahola *et al.,* 2014). Williams *et al.* (2010) distinguished between governance of projects, which aims for efficient delivery, and governance through projects, which aims to choose the right concepts and ensure that the intended effects are realized. In practice, the focus in the literature as well as in practice has been more on the former than the latter (Volden and Andersen, 2018), and to date, this is where agency theory has had an influence. An exception is Zwikael and Smyrk (2015), who showed that there are principal–agent relationships at multiple levels, with the "funder" on

top, who hires a project owner to be accountable for benefits realization, and the project owner, in turn, hires a project manager to be accountable for efficient output.

Opportunistic behavior has been particularly studied in relation to the choice of contract strategy and the relationship between commissioner and contractor. For example, a number of authors have discussed the optimal design of public–private partnership contracts from a principal–agent perspective (Boardman and Vining, 2012; Ho *et al.,* 2015; Iossa and Martimort, 2015; Liu *et al.,* 2016). The key is to make the contractor accountable, both to ensure efficient implementation and for the operational and maintenance phase.

Biesenthal and Wilden (2014) found that principal–agent theory has been somewhat less influential in relation to the organizational level (corporate governance). Although agency theory can also be useful to understand organizations (Eisenhardt, 1989), it can be argued that it provides a somewhat narrow perspective, with its often strong focus on "hard incentives" (Joslin and Müller, 2016). In the study of organizations, the theory should therefore not be used alone but rather in combination with other theories, such as stakeholder and stewardship theory, transaction cost economics and resource dependence theory (Biesenthal and Wilden, 2014).

Our focus is not on the project-based organization but rather on the whole nation and the government as principal, who receives project proposals from various agents. To date, the project governance literature has largely disregarded this perspective. Admittedly, there is a wide body of literature on stakeholder involvement in which it is noted that stakeholder inclusiveness involves the risk of expectation escalation (Eskerod *et al.,* 2015). Further, Morris and Hough's (1991) study is pivotal because the authors examine factors "beyond project management," such as political and organizational aspects and community involvement. However, there have not been any follow-up studies.

A few more recent studies have analyzed project governance frameworks for major public projects (Volden and Andersen, 2018; Volden and Samset, 2017a; Williams *et al.*, 2010). They all document an increasing focus on the front end, including quality assurance of the business case. However, Volden and Andersen (2018) noted that the earliest idea phase is generally not included in the governance schemes, with the argument being that this phase addresses "strategic and political issues beyond the project." Project ideas are often initiated or "picked up" from below, but procedures and roles in this phase are nonexistent.

Studies on over-optimism and deception

Another relevant group of studies is those that address deception and similar phenomena in public projects. Bent Flyvbjerg and his colleagues stand out, with their many publications on cost overruns and benefit shortfalls, especially in transport projects, which they explain as largely due to deliberate miscalculations on the part of key stakeholders and project promoters (Flyvbjerg, 2009; Flyvbjerg *et al.*, 2002, 2003, 2009; Cantarelli *et al.*, 2010). The issue has also been raised by, among others, Wachs (1987, 1989) and Mackie and Preston (1998). Solberg and Preuss (2007, 2015) discussed how major sporting events are often funded by the national government based on the argument that they are public goods, but this approach often leads groups of "free riders" (at the host destination) to exaggerate the economic value of the event and downplay the cost.

However, as noted by Siemiatycki (2016), these studies have not been very influential in the engineering/project management field, in which cost overruns are still largely explained by "honest errors." Deliberate manipulation is difficult to prove and especially to distinguish from over-optimism stemming from cognitive biases (Lovallo and Kahneman, 2003; Meyer, 2014).

Lefley (2006) discussed the role of the project champion, and how this person may bias project selection. Through a single case, Lefley demonstrated that the project champion, who was also a member of the appraisal team, clearly gave more optimistic scores than the others on the team. Pinto and Patanakul (2015) examined the situation in which project champions turn into narcissists; the authors argued that champions tend to select riskier,

more high-profile projects and are more likely to escalate commitment in the face of clear evidence of poor performance.

Other publications have addressed related topics. Kvalnes (2014) explored the concept of dishonesty and used it to explain misreporting issues in projects. He argued that truth telling is not necessarily an inherent characteristic of a person but may depend on the situation. Locatelli *et al.* (2017) introduced the term corrupt project context and highlighted it as a factor that can undermine the performance of projects. They also argued that public megaprojects hold a special risk of corruption due to high economic rents, public officials' discretionary power and often weak institutions.

Experiences from development aid projects

A branch of the literature that is particularly relevant is studies of the effects of development aid. There has been a growing awareness of the adverse incentives created by access to "free funding" and how it can affect outcomes negatively. For a literature review, see Newby (2010) or Wiig and Holm-Hansen (2014). It is also useful to visit Merton's (1936) classic text on unanticipated effects on purposive social action and Boulding's (1981) textbook on grants as an economic phenomenon. An important message is that it is naïve to believe that a scheme or project meant for the common good will be perceived that way by everyone involved. There will always be some who win and some who lose, and all those involved will adapt in a way that is best for themselves. This must be well understood in order to design an effective scheme.

A pivotal study was conducted by Ostrom *et al.* (2001), who used agency theory to explain what happens. Moral hazard problems imply, for example, that recipient countries take less responsibility for investing in infrastructure. They exchange their own funding for aid, implying that in effect, the money does not finance new projects, but rather something else that was not intended by the donors. Adverse selection problems are very common and can be observed in the form of unviable projects being approved by the donor country. Generally, information asymmetry makes these problems occur. Furthermore, the problems increase with the number of layers in the hierarchy. Contractors are crucial because they may have an interest in particular projects being selected and prolonged. The contractor is often the initiator and serves as the link between the recipient group and the funder; thus, he controls the flow of information in both directions.

The phenomenon perverse incentives is frequently observed in this sector. A typical example is when aid intended to make the recipient community more robust actually makes it more dependent on long-term foreign aid. Ostrom *et al.* (2001) indicated serious problems with perverse incentives at many levels that result in unsuccessful projects, waste of public funds and corruption.

The key message is that donors must be aware of the incentives that they create with their aid. The multifaceted set of relationships should be properly analyzed to see how they will be affected by new projects. Local beneficiaries' ownership should be strengthened by making them enunciate a demand for aid, allocate at least some of their own assets to the project, obtain benefits and have clear-cut responsibilities. The prospects for solving the agency problem are, however, not very encouraging due to the weak institutions found in many of these countries.

Fiscal federalism

Incentive problems related to funding between the central and local level have also been studied in developed countries. Fiscal federalism is a field in public economics that discusses how revenues and costs should be allocated across the vertical layers of administration (e.g. Oates, 1999). According to welfare economic principles, local public goods, such as local transportation infrastructure, should be provided and funded at the local level which knows the local preferences best. When local taxes reflect the benefits of local services, anyone can move to the municipality of their choice (Tiebout, 1956).

There may be good reasons for the national government to transfer money to the local level. One is local risk aversion and the need for the government to act as an insurance company. Another is the positive external effects of investment in one region on neighboring regions and even for the nation as a whole. A third argument concerns, in the same way as in development aid, preferences for redistribution from rich to poor regions. However, regardless of which argument is used, federal grants on a large scale create an imbalance between geographically concentrated benefits and dispersed costs and give rise to incentive problems, such as a lack of economic discipline and recurrent problems with bailouts. Game theory has been used to show that the national government lacks credibility when announcing "hard budget constraints" because it will be better off saving an irresponsible municipality than allowing it go bankrupt (Goodspeed, 2002; Rodden *et al.,* 2003; Wildasin, 2004). de Rus and Socorro (2010) discussed a similar problem in relation to infrastructure with supranational (EU) funding and the incentives that occur in national governments.

How we intend to fill the gap

We have searched rather broadly for relevant studies that use theory of incentives to explain what occurs at the front end of public projects, such as over-optimistic appraisals and their effects on project selection and strategic success. Agency theory has definitely inspired the project management and project governance literature. However, it has primarily been used in discussions of how to motivate project managers and contractors and less so in relation to the front-end phase, in which the project promoters are typically external parties. Flyvbjerg is one of a few researchers who have discussed miscalculations and deception on the part of local communities and others in the front-end phase.

However, in the development aid sector we find a long tradition of studying adverse and even perverse incentives in relation to project selection, which may be explained by "free funding." There are some important distinctions between infrastructure projects in a developed country such as Norway and projects funded by an external donor in a developing country. After all, the allocation of a common tax pool is a democratic issue that concerns all groups in society, not a gift from one party to another. However, there are clear similarities: in a developed country, just as in a developing country, the group that gains from the project will (if small enough) consider the project to be practically free-of-charge. Moreover, in both cases, there may be several layers of principal–agent relationships from the funder to the privileged group, where only the top level (at best) is concerned about the common good for the larger society. As part of our study, we explore how far this comparison with aid projects can be taken. We also find some inspiration from the literature on fiscal federalism, although it takes more of a macroeconomic perspective than a project perspective.

In this study, we explore the financial incentives and principal–agent relations at the front end of a sample of Norwegian Government-funded projects. In particular, we look for cases of perverse incentives of the types found in the development aid sector. We also discuss what can be done to avoid the problem and whether any measures taken in the studied projects either mitigated or avoided the most perverse outcome.

We hope that our contribution will provide an improved understanding of the incentives related to public project initiation. We currently know little about the earliest idea phase of public projects, which is often not included in governance schemes. Our research should therefore be useful to scholars as well as practitioners within the field of project governance.

Our framework of analysis

In this section, we present our framework of analysis, which is based on agency theory.

Simple model

In its simplest form, the model includes two parties: the national government as the funding party and the group or community that receives the major portion of the benefits. This concentration of benefits implies that we focus on local public goods (or even private goods) as opposed to national public goods such as defense acquisitions and national highways, for which the benefits would be almost uniformly distributed across the population.

Project approval and funding is discretionary; hence, the alternative for the privileged group is to receive nothing. The privileged group is more or less well organized. Its role is to provide the government with information about local conditions and to propose new projects when needed. Project implementation, which is not the issue here, occurs under the auspices of the government or of an agency whose perception on project selection does not differ from the government's perception.

Agency problems arise when two preconditions are present: an underlying conflict of interest and information asymmetry. This is illustrated in Figure 1 in terms of a flow of unconditional funding on the right-hand side and a restricted information flow in the opposite direction (hence the dotted line) on the left-hand side.

The conflict of interest is introduced by the combination of distributed costs and concentrated benefits. If we assume that costs and benefits can be measured in monetary terms and that the government's decision criterion is the benefit–cost ratio, abbreviated as BC ratio = B/C, it can easily be shown that the privileged group's BC ratio always exceeds the government's BC ratio as long as the privileged group's share of the benefits exceeds its share of the costs. The other precondition for agency problems is that the government cannot verify the information coming from below.

Agency problems at the front end can be expected to materialize in three main ways:

(1) Positively skewed appraisal: the privileged group presents a biased appraisal in order to pass the point of *de facto* approval. A variant is strategic split-up (i.e. only presenting the main project component at first and then later disclosing the remaining components and adjoining projects). Lobbying activities to affect the government's "demand" for the project are also included in this category.

(2) Expansion after *de facto* approval: a related group of problems occurs when the level of ambition is kept low until the decision makers are *de facto* committed. Then, a restricted scope is turned into an oversized project.

(3) Moral hazard: whereas the first two are adverse selection problems, cases with moral hazard problems can also occur. This is when the privileged group's promises to contribute for complementary local projects or to sufficient maintenance, and to take responsibility for benefits realization and sustainability in the long run, is not followed up.

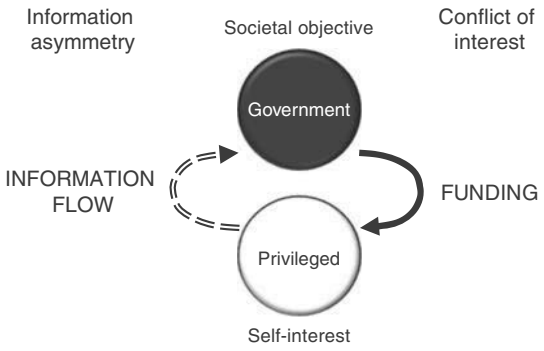


Figure 1.
Incentive problems in public investment projects – simple presentation

In extreme cases, incentives may turn perverse, resulting in counterproductive outcome. It is not easy to provide a precise definition of perverse incentives, for example, in terms of a threshold level for the national BC ratio, but here we define perverse incentives as "agency problems that are so severe that they lead to the selection of projects that are highly unsuccessful in strategic terms and a waste of public funds."

Extended model

A more realistic model should take into account that a whole chain of principal–agent relationships may be involved. Between the government and the ultimate beneficiaries, there may be one or more intermediaries such as a local government, or a self-interested public agency or consultancy firm. The intermediary will typically receive funding on behalf of the privileged group, which adds to the information asymmetry. Intermediaries have their own objectives, such as to maximize their budget. In development assistance, consultancy firms are often major contributors to agency problems because they convince donor governments to choose particular projects that benefit themselves (Ostrom *et al.*, 2001). However, an intermediary's private objectives may not be visible in the project appraisal, where this actor pretends to be concerned about the societal impact.

We could extend this model further by, for example, including additional layers in the upper part to take into account the fact that bureaucrats are agents for politicians, and politicians are, in turn, agents for the people. However, such relationships are less formalized; thus, adverse incentives would be difficult to document empirically. Another extension would be to include stakeholders who are not officially involved in the transaction but enter the process as "freeloaders" and try to influence the selection process, such as neighboring landowners or possible suppliers for a future project.

We present a simplified illustration of our model with only one intermediary in Figure 2.

Methodology and data

This study is only meant to be an initial probe into the phenomenon of perverse incentives at the front end of public projects. We have therefore chosen a qualitative, case-based approach with purposive sampling, where the aim is not to draw conclusions about scope and frequencies but rather to identify and understand the mechanisms and provide some

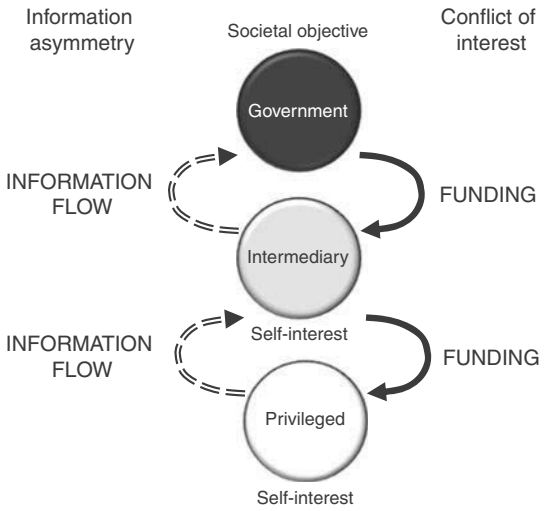


Figure 2.
Perverse incentives in public investment projects – extended model

insights into the subject matter. In line with Yin (2014), we find that the concrete, context-dependent knowledge that can be obtained from case studies is highly valuable and that precisely because of the detailed understanding of the causal relationships that are obtained, the results can be applied to other contexts. Furthermore, as noted by Flyvbjerg (2006), extreme cases can often reveal more information and clarify deeper causes better than average cases.

We address the following research questions:

RQ1. Describe the principal–agent relationships at the front end of the case projects, in terms of degree of conflict of interest and information asymmetry.

RQ2. What types of agency problems, if any, materialize in the case projects? (cf. the three types listed above).

RQ3. How do these problems seem to affect the projects' strategic success?

RQ4. What occurs in the extreme cases when agency problems turn perverse?

RQ5. Describe any steps taken to try to avoid or mitigate the problems.

Our sample consists of one aid project, which serves as a reference case, and eight government-funded investment projects from different sectors in Norway. The nine case projects were not selected randomly but rather because they were assumed to represent cases of perverse incentives. Our selection criteria were that each project was large in terms of investment cost; benefitted a limited group or community; was funded by the state, with few or no obligations for the privileged group; used discretionary funding; and had data available on the actors, costs and benefit estimates from the project's front-end history. Details of the selected projects are provided in Table I.

We mostly used secondary sources, namely, document studies from the projects' front-end phases, such as needs analyses, business cases, risk assessments and cost estimates, and we supplemented these data with interview data (two to three semi-structured interviews per case) and information from the public debate retrieved from the Norwegian digital media archive Retriever. All the case projects were large, and most of them were widely debated, both before and after they were implemented. Therefore, there are large amounts of publicly available information.

For each project, we identified the parties involved, their preferences and their roles in the front-end phase (privileged group, intermediary or others). In addition, we described the flows of funding and information in order to identify indications of agency problems and, in extreme cases, perverse incentives. We also assessed the projects' strategic success, defined as the extent to which they appeared relevant and feasible even in the long term, and whether the total benefits were worth the cost. Any statements from the government or by independent analysts concerning this issue were registered retrospectively and supplemented by our own analyses. Finally, we registered any attempts to avoid or mitigate emerging problems or discussions about such measures.

Findings and analysis

The main findings from our analysis are summarized in Table II and explained further below.

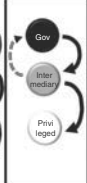
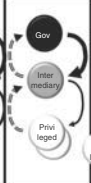
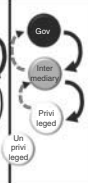
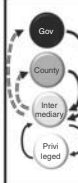
Principal–agent relationships

All nine case projects had in common that they were largely funded by the Norwegian Government while they benefitted specific groups. The thickness of the right arrow in each figure in Table II indicates the seriousness of the conflict of interest, which depends on the combination of the privileged group's share of benefits and costs. Any imbalance was normally in line with the intention, but in some cases, the original objective as defined by the

Table I.
The projects studied ($n = 9$), sorted by co-financing

Name of project	Description	Status as of 2017	Target group/ privileged	Co-financing	Total cost (NOK million)	Year completed
Turkana Fisheries	Aid project in Kenya, including a refrigeration plant, fishing boats, trucks and roads, to generate employment and improve living conditions for nomads	Many unforeseen problems occurred. The fish resources were limited, and the plant's operational costs were unacceptable. The nomads' situation worsened, and the project was terminated many years later	Nomads in Kenya	None	1,500	1990
Linesøya Bridge	Bridge to a remote island in central Norway. The goal was to stop depopulation and boost the local economy	Investment cost was twice as high as budgeted. Depopulation continued, and the level of traffic is very low today	Islanders from Linesøya	None	250	2011
Stad Shipping Tunnel	1.7 km tunnel for boats on the west coast of Norway to avoid the coast in dangerous seas	After decades of planning, the local expectations are high, but the cost estimate has increased rapidly. The government has signaled that funding will be provided	Local fishermen and other inhabitants in the Stad area	None	2,500[a]	n/a
E16 Laerdal Road Tunnel	Road tunnel along one of the five roads (not the shortest) between the cities of Oslo and Bergen through the small town of Laerdal	It was the world's longest road tunnel when it was built (24.5 km). The tunnel is still in daily use, though with little traffic, primarily local	Inhabitants of the Laerdal region	None	1,050	2000
Winter Olympics in Lillehammer	Venue and organization of the 1994 Winter Olympics, with the intention to "sell" Norway as a tourist destination and stimulate growth in the eastern part of the country	The event was highly successful as a sporting event, but the project became very expensive, and the long-term economic impacts were limited	Inhabitants of the Lillehammer area/ sports enthusiasts in the nation	None	7,500	1994
Lofast Link Road	Roads and tunnels connecting the Lofoten region to the mainland	There was considerable disagreement about the choice of route. The chosen route contributed to the isolation of the neighboring Vesterålen region	Inhabitants of the Lofoten region	None	1,367	2007
Rock City	Museum in the town of Namsos in central Norway. Was partly funded by the government to be a national center for pop and rock music	The music industry found the location remote. In practice, it was run as a center for local music. Was eventually taken over by Namsos municipality but is currently being closed down	Inhabitants of the town of Namsos	Some	50	2013
St Olavs Hospital	Renewal and expansion of the main hospital in central Norway	The hospital has received prizes for its architecture, but the investment cost was high, and it is considered an expensive way to operate a hospital	Sick people/ inhabitants of the Trondheim region	Some	13,000	2014
Hvaler Road tunnel	Subsea road tunnel to a small island in south-east Norway. Partly initiated to reduce depopulation	Was partly funded by motorists in the larger region through toll roads, which led to conflicts for some time. Depopulation seems to have been avoided	Inhabitants of the Hvaler area	Some	200	1989

Note: [a]Estimate, since the project has not yet been implemented

Name of project		Turkana Fisheries	Linesøya Bridge	Stad shipping tunnel	E16 Laerdal tunnel	Winter Olympics	Lofast Link Road	Rock City	St Olavs Hospital	Hvaler Road tunnel
Principal–agent relationships										
Who initiated the project		Intermediary	Privileged	Privileged	Parlamentarians	Privileged	Privileged	Privileged	Intermediary	Funding party
Risk of agency problems	Conflict of interest[a]	Yes	Yes	Yes	Yes	Yes (different groups)	Yes (different groups)	Partly	Partly	Partly (different groups)
	Info asymmetry	Yes	Yes	Some	Some	Yes	Some	Yes	Yes	Some
Agency problems materialized	Positively skewed appraisal	Yes	Yes, probably deliberate	Yes, probably deliberate	n/a, widespread lobbying	Yes, and widespread lobbying	No	Yes	Yes, and widespread lobbying	Partly
	Expansion after de facto approval	Yes, took 20 years to terminate	Partly	Yes, the scope increased over time	Uncertain	Yes, until reorganization was required	Uncertain	Uncertain	Yes, to a very large extent	No
	Moral hazard	Yes, infra structure not maintained	Partly	Uncertain	Uncertain	Partly	Uncertain	Partly, until more local funding required	Yes	Uncertain
Strategic success	Value to privileged group	No	Yes, useful to locals, but they are few	No	Yes, useful to locals, but they are few	Yes	Yes, high for one group, but low for another	Yes, until more local funding required	Yes, until more local funding required	Yes, high for one group, but low for another
	Value to society	No	No, other projects displaced	No, low value for money	Partly	Partly	No, low for the region as a whole	No	Partly, low to medium value for money	Yes, relevant for the region as a whole
Perverse incentives?		Yes	Yes	Yes	Partly	Partly	No	No	Partly	No

Notes: $n=9$. [a]Agent with a combination of high share of benefits and low share of costs

Table II.
Main findings summarized

government was broader and related to national goals. For example, the E16 Laerdal Road Tunnel was intended to improve the main road between two major cities in Norway, and Rock City was established as a national center for pop and rock music, but in both cases, the local perspective became dominant. Together with the information asymmetry, which was present to a greater or lesser extent in all projects, this indicated an inherent risk of agency problems in all the projects.

A state agency was normally involved, but in most cases was not considered self-interested; therefore, such agencies are considered part of the "government" in the figures. On the other hand, all the projects also had intermediaries that represented an additional layer of agency problems, such as the county and/or the affected municipality, and in some cases, specific organizations were established to represent the privileged groups.

In contrast to the aid project, the intermediaries in the Norwegian projects did not dominate the privileged groups, which were profoundly involved in project initiation, design and planning (but not implementation). Notably, there was not always a single privileged group, but we found examples of conflicts of interest between groups at the local level.

Agency problems materialized

Clear signs of agency problems materialized in several projects. Generally, adverse selection problems were the easiest to identify, whereas moral hazard problems, such as those seen in the aid project, were less prominent in the Norwegian cases.

In particular, the initial cost estimate presented by the privileged groups was underestimated in most cases. We observed indications that this was done deliberately, but it was difficult to obtain confirmation from the interviewees. For example, the initial cost estimate for the venue and organization of the Winter Olympics Project was clearly unrealistic at less than 25 percent of the final cost. Similarly, the benefits for users and society were overestimated. In the Linesøya Bridge case, the first cost–benefit analysis indicated that the project was marginally profitable, while the final result showed a cost 200 percent higher than the estimate and benefits at only 25 percent of the estimate, resulting in extremely low value for money. There was also a clear tendency for some of the case projects to be supersized, and the cost-driving requirements were strategically presented after passing the stage of *de facto* approval. The Stad Shipping Tunnel, the Winter Olympics and St Olavs Hospital were the three worst cases in this regard.

The government seems to have had surprisingly high confidence in the information provided by the initiating party in each case. Appraisals were either provided by the agents themselves or by the state agency based on input from the agents. Only in one project (the Stad Shipping Tunnel) did the government, admittedly very late in the process, demand an external review of the business case. In two of the projects, the incentive problems were clearly amplified when key politicians and other stakeholders on the funding party's side either had their own agenda (campaigning) or their loyalty was with the privileged group rather than with society at large.

Strategic project success

We conducted a rough assessment of the projects' value to the privileged group, as measured by user benefits, and to society as a whole, as measured by alignment with national policy and value for money. It is clear that the user benefits were often not at the anticipated level. However, in most of the projects the privileged group was either indifferent or satisfied with the situation. Only the aid project generated impacts on the target group that were clearly negative. In the case of the Shipping Tunnel, it is unlikely that the project will improve the lives of the target group at all, whereas the other seven projects did bring some (often limited) user benefits. As will be discussed further below, in three cases, the terms for funding changed during the process, and the privileged group eventually had to take responsibility for a considerable share of the cost as well. In one of these projects, this led to bankruptcy.

The value to the broader society was considered meager in all projects but one. Our analysis and assessments show that at least four of the projects (Turkana Fisheries, Linesøya Bridge, Stad Shipping Tunnel and Rock City) brought so little value and their relevance to society was so low that they should never have been approved, from a rational, economic perspective. In these cases, a different project would have resulted in more efficient and effective solutions to the problems at hand, or there was no need for an investment at all.

Perverse incentives?

We also looked for perverse incentives, leading to counterproductive outcome. We have already argued that four of the projects were outright failures in strategic terms. But were they selected because the government was misled by self-interested agents? In the following, we argue that the answer is yes in three out of the four cases, with the exception of Rock City.

Admittedly, the project appraisal for the Rock City project was also biased, and the government was misled to believe that the museum would be operated as a national center for pop and rock music, not just as a local center. However, the seriousness of the agency problem was more limited in this case. The government's share of the funding was only approximately one-third, and the initiators must have known that in the longer run, the museum would have to be self-sustaining. Therefore, we conclude that the initiators mainly misled themselves into believing that the museum would be financially viable and would boost the local economy. The case is an example of over-optimism more than perverse incentives.

The first project with perverse incentives was the reference project from development aid aimed to generate employment for nomads in a remote area of Kenya. It was originally initiated to mitigate a short-term crisis, but became a permanent and extremely costly commitment encouraged by the Norwegian entrepreneurs. The result was a complete failure from all perspectives except that of the entrepreneurs'. The implicit assumption that the nomads supported the development was not verified, and there were large unintended effects on the environment.

Two of the Norwegian projects were on par with the aid project: Linesøya Bridge and the Shipping Tunnel. The former was a fully government-funded project to build a bridge connecting a small population on an island to the mainland. The privileged group, as represented by local politicians and landowners, was heavily involved in lobbying upfront and presenting visions of economic growth and wider benefits. This lobbying was supported by a cost–benefit analysis that later turned out to be flawed. The decrease in population continued, and there has hardly been any commercial activity on the island. Only some local landowners have benefitted from selling plots of land to tourists. There were also indications that the municipality never followed up and expended the effort that it should have to stop the depopulation (i.e. a moral hazard).

In the case of the Shipping Tunnel, local authorities promoted the project for decades, and a lobbying organization was established with the sole purpose of ensuring that the project would be realized. At first, the main argument was that ships in the area faced high risks from dangerous seas. Some national politicians signaled a positive attitude toward the project, which encouraged the lobbying group. However, over the years, vessels became larger and safer, and wave detection technology was implemented, such that eventually there was no longer any need for the project. Still, the proponents did not want to let go of the idea. They brought new arguments to the table, including that the project would provide local employment opportunities and that the tunnel would be a tourist attraction, or simply "it has already been promised to us."

The project has still not been implemented, and over time the government has turned suspicious about the project. The government has commissioned several independent analyses that demonstrate its low value for money. But surprisingly, it might be too late to stop it now due to the effects of perverse incentives over more than 20 years. The fundamental problem seems to be that the government never refuted the premise that the funding should come entirely from the state. Therefore, the trend has been for local proponents to come forward with new proposals that steadily increase in the scale, scope and ultimately the cost of the tunnel. Figure 3 shows the development of the project through ten project appraisals over a period of 25 years. The scope certainly increased before stabilizing at approximately NOK2.5bn. The net present value was marginally negative in the first part of the front-end phase, when the scope was moderate. In the second part, three cost–benefit analyses commissioned by the government showed highly negative results, whereas two analyses commissioned by the privileged group concluded that the project would be profitable.

Changes in the terms of financing

An interesting observation is that in three of the case projects, the rules of the game were changed mid-way through the project. Measures to ensure accountability were introduced,

and the financial incentives were brought in line with government's objectives. In all three cases, there are clear indications that this was done to reduce the agency problems and that it actually managed to prevent or mitigate the most serious consequences.

In the Rock City case, the financial scheme was unclear from the start. The center and the municipality were hoping that the government and/or the county would reimburse any cost overruns related to the investment project and yearly deficits. The organizers of the center therefore chose a high level of activity, even after the revenues failed to materialize. Eventually the government terminated its funding. The municipality had no other choice than to take over, but after a few years, it decided to close down the center. Without this change, the center would probably have continued its activities with annual deficits covered by the state and the county.

A similar story can be told in the St Olavs Hospital case. Based on the experiences of other state-funded hospitals in previous years, the initiators at the county level (the owner of the hospital at the time), simply assumed that the national government would take responsibility for funding, and accept any cost increases after the *de facto* approval of the project. In the beginning they were correct. From an initial NOK1bn, the cost estimate increased to NOK12bn within a few years and would most likely have increased further, if the government had not, coincidentally, introduced an extensive hospital reform measure at the time. The reform implied that hospitals in Norway would no longer be managed by underfinanced counties but by regional health companies owned by the government with results-based funding. The St Olavs Hospital project had already been approved at the time, and the government therefore agreed to grant the NOK12bn as promised. To avoid further overruns, however, the new health company had to introduce considerable reductions in scope, and finally managed to complete the project at the budgeted cost.

The last case was the 1994 Winter Olympics. After the government's decision to guarantee the cost in 1987 based on an NOK1.8bn estimate, an organization that comprised the municipality, the Norwegian Olympic Committee and the Ministry of Culture was established to be responsible for further planning. The Ministry signaled that the government would fund the necessary improvements in the national transportation infrastructure and telecommunications/TV but that the other parties would have to fund the sports facilities and the various local infrastructures. However, this statement was not credible because the government had guaranteed to fund the whole event. The scope increased rapidly based on new needs and requirements identified by the municipality and the Norwegian Olympics Committee. Additionally, the International Olympic Committee

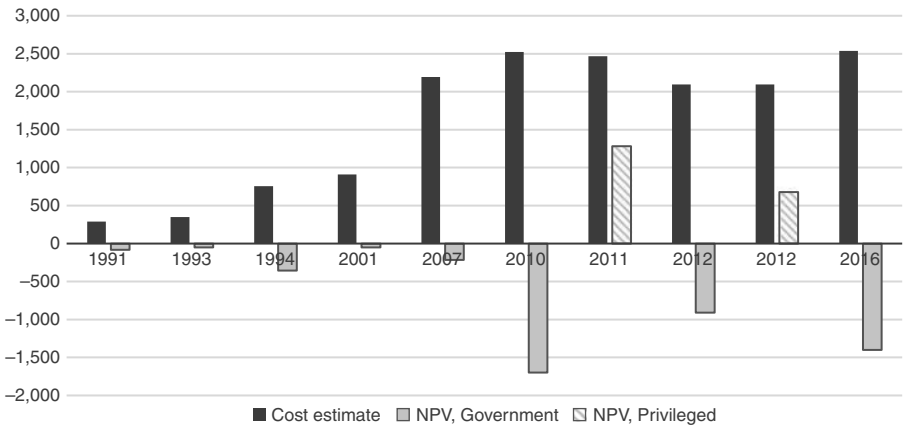


Figure 3.
Stad Shipping Tunnel, ten project appraisals from 1991–2016; the first nine are based on Kvalheim (2015), and the last one is from the most recent National Transport Plan (million NOK, 2011)

had extensive detailed requirements and also acted an agent with adverse incentives. By 1989, the Ministry of Culture decided to reorganize the project and transfer responsibility to a new company owned by the state and funded by a block grant and economic incentives to comply with it. Thereafter, there were no further cost increases.

Discussion and conclusions

Findings and implications

Our nine case projects all differed with respect to, for example, size, complexity, the parties involved and sector. However, they were included in the sample because they had some characteristics in common: they were largely funded by the government, on a discretionary basis, with concentrated benefits for privileged groups. Moreover, all of the projects were large – not necessarily in absolute terms but certainly in relation to the privileged groups.

The following points summarize our answers to the research questions:

(1) All the case projects were characterized to various degrees by conflicts of interest at the front end, often with the municipality in a key role. In most cases, the information about local conditions was clearly asymmetric. Thus, the risk of agency problems was clearly present in most cases.

(2) We observed several signs of agency problems materialized. Early cost estimates were typically unrealistic, and the benefits to users and society were overestimated. It was surprising how much confidence the government had in the information received from below. Further, project initiatives that might have been sensible in the first place grew out of proportion and became over-dimensioned as a result of requirements the privileged groups introduced after the *de facto* approval of the project. Substantial resources were spent on lobbying.

(3) The value for money was often meager and much lower than anticipated. The development aid project was extreme, wherein the gross benefits for the target group were negative. In the Norwegian projects, the value to the privileged group was generally positive but often low. In one case, the privileged group accidentally ended up with the bill and was forced into bankruptcy.

(4) We conclude that perverse incentives were present in three cases. These projects were selected by the government because of the adverse incentives and represented a waste of public money.

(5) The introduction of liabilities and incentives for cost control seems to have had a disciplining effect in three cases. Although they were introduced at a late stage when it was too late to stop the project, they probably led to the prevention of more severe problems.

The purpose of this study was to show how a simple framework of analysis based on agency theory can be used to explore the risk of perverse incentives at the front end of public projects. The comparison with development aid may seem inappropriate, but it worked well for the purposes of our study. In line with, for example, Eisenhardt (1989) and Biesenthal and Wilden (2014), we do not claim that agency theory is sufficient or that it should be used alone. However, it provides a useful perspective on project initiation that has rarely been applied in the project management literature. The absence of liabilities, such as co-funding, is a simple and obvious risk factor of which all project funders should be aware.

A timely question is whether these findings are relevant to the larger group of public projects in developed countries. This study does not answer that question, but we do know that in Norway, as in many other countries, there is a gap between the need for local infrastructure and the availability of local capital. This explains why, for example, local

roads, hospitals, universities and major sporting events are largely funded at the national level. Admittedly, hospital projects are no longer included among this group of projects in Norway due to the reform in 2002 that led to the establishment of health companies. Other projects remain in the high-risk category. Many transportation projects are partly funded by user fees, but the local share is often low, and as noted by Ostrom *et al.* (2001), for co-funding to provide the right incentives, there must be a link between the decision to initiate a project and the liability to pay. When a municipality or landowner initiates a project and sends the bill to motorists, this requisite is not met.

Countermeasures

The problems we have described in this paper do not necessarily have quick-fix solutions. There may be good reasons why the target group should be involved in the planning process while not putting its own money at stake. Still, much can be done in terms of improving the processes and systems related to individual projects. The potential for avoiding perverse incentives is much greater in a setting such as Norway compared with a country receiving development aid, about which Ostrom *et al.* (2001) concluded that the lack of well-functioning institutions is a serious obstacle.

Samset and Volden (2012) recommended that a thorough discussion of the terms of financing should be part of every project assessment. Agency theory can be used not only to identify the risk of perverse incentives but also to suggest how to mitigate it. From our Figure 1, it is apparent that the problem is created by the combination of conflicts of interest and information asymmetry. Thus, the solution should be twofold, as illustrated in Figure 4 and discussed below.

First, measures are needed to improve the information flow in order to ensure that the government bases project selection on high-quality information about the problems, needs, benefits, costs and risks. As noted by Klakegg and Volden (2016), the public sector depends on transparency as a means to strengthen accountability, whereas the private sector has competition. Conceivable measures would be to conduct third-party reviews of project proposals, give the general public an opportunity to express their views (e.g. through open hearings) and perform systematic *ex post* evaluations to learn about the level of impacts that can be expected.

Second, as much as possible, one should reduce the conflicts of interest between the government and privileged groups. Liabilities should apply both ways so that the privileged group has ownership of the project and is made accountable for the project's success. Solutions may be to require co-financing or local risk-taking, given that such requirements are credible, or to reward benefits realization *ex post*. Much of the literature on incentives in development aid focuses on how aid can be used as both a carrot and a stick. Ostrom *et al.* (2001) discussed the incentive problems related to different aid modalities and argued that

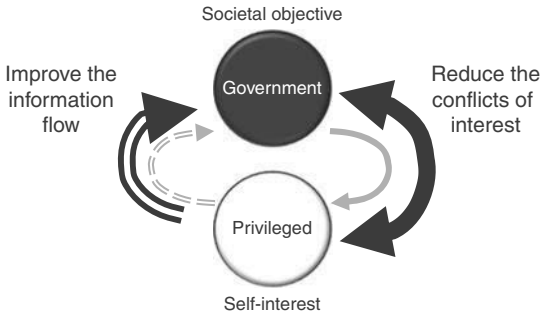


Figure 4.
Measures to solve the problem of perverse incentives

loans have advantages over pure transfers for which no repayment is required, given that ownership is sufficiently rooted.

An even better solution would be to avoid the system of discretionary project-based funding of individual projects. As noted by de Rus and Socorro (2010), the most high-powered incentive scheme would normally be fixed-price contracts. In relation to our projects, this would mean a lump sum that the privileged group could allocate freely. This approach would be in line with the recommendations of Rattsø (2003), who discussed the more general need to make municipalities accountable, and Flyvbjerg *et al.* (2003), who argued that the state should grant a general allocation (to the local administration or a state agency) and require that the project selection meet certain objective criteria, such as value for money.

However, in cases where the government does finance projects with concentrated benefits, it is crucial to have an overall project governance framework in place that takes the risk of front-end agency problems into account. In recent years, Norway and some other countries have introduced governance schemes that cover the choice of concept (Volden and Samset, 2017a). Since 2006, Norwegian Governments have required that the largest investment projects undergo an external quality assurance of the conceptual choice to ensure that they are based on real needs, that alternative solutions are considered and that their value for money is assessed (Volden and Samset, 2017b). It is too early to conclude whether the scheme has led to more successful projects, but it clearly has helped sift out some poor project ideas early in the process. For example, in recent years, the Norwegian Olympic Committee and municipal partners have twice applied for a state guarantee to again host the Winter Olympics (2018 and 2022). In both cases, the external reviewer found that the benefits were overestimated and the costs underestimated, which led to the government rejecting the proposals. By contrast, the Shipping Tunnel is now being realized despite having been exposed to an external quality assurance evaluation. The external review was conducted in 2012, but as noted above, that was probably too late since the government had already *de facto* approved the project.

The Norwegian quality assurance scheme attaches great importance to providing transparency, controlling the quality of analyses and making all project information publicly available. However, it does not require co-funding or include any other measures to reduce conflicts of interest. Some other countries have introduced stricter co-funding requirements in their project governance schemes (Volden and Samset, 2017a). For example, the Dutch scheme requires co-funding from local authorities that come forward with a project proposal, and it requires that all investment initiatives in excess of EUR60m have private co-funding. The rationale is that this will result in more weight being attached to long-term revenue flows as well as efficient project implementation.

Limitations and future research

In all case-based studies, there is a risk of subjective bias. Researchers start out with some hypotheses that they try to confirm through the cases. This is especially true for studies that apply purposive sampling. However, Flyvbjerg (2006) noted that the question of subjectivism and bias applies to all methods, including, for example, the choices of categories and variables in a quantitative study and the structure of a questionnaire. Flyvbjerg stated that the case study contains no greater bias toward verification than other methods. On the contrary, experience indicates that the case study contains a greater bias toward falsification of preconceived notions.

Perhaps the most severe limitation of our study is that we have largely referred to the national government as one internally consistent unit and downplayed the way politicians and bureaucrats enter and leave the picture. Political decisions are clearly made through processes in which agreements about the goals and fundamental assumptions cannot be taken for granted (O'Leary, 2012), and in which there are many examples of irresponsible behavior, even

at the top level (Miller and Hobbs, 2005). Future studies should explicitly extend the model presented in this paper to include principal–agent relationships at the government level.

Another question that we have largely ignored here is whether different motivations for the use of government funding (e.g. distributional, wider impacts beyond local effects or insurance) require different measures to avoid agency problems. Still another question, which will have to be answered empirically, is whether perverse incentives in a certain sector merely lead to a suboptimal project portfolio or whether the total number of projects also puts pressure on overall budgets. In the latter case, there is a common-pool problem that might require specific measures.

There is no doubt that more research is needed on perverse incentives in order to learn more about the mechanisms involved, the scope of the problem and countermeasures that may work in various contexts, as well as how agency theory can best be combined with other theories.

References

Ahola, T., Ruuska, I., Artto, K. and Kujala, J. (2014), "What is project governance and what are its origins?", *International Journal of Project Management*, Vol. 32 No. 8, pp. 1321-1332.

Andersen, B., Samset, K. and Welde, M. (2016), "Low estimates – high stakes: underestimation of costs at the front-end of projects", *International Journal of Managing Projects in Business*, Vol. 9 No. 1, pp. 171-193.

Biesenthal, C. and Wilden, R. (2014), "Multi-level project governance: trends and opportunities", *International Journal of Project Management*, Vol. 32 No. 8, pp. 1291-1308.

Boardman, A.E. and Vining, A.R. (2012), "The political economy of public-private partnerships and analyses of their social value", *Annals of Public and Cooperative Economics*, Vol. 83 No. 2, pp. 117-141.

Boulding, K. (1981), *A Preface to Grants Economics. The Economy of Love and Fear*, Praeger, New York, NY.

Cantarelli, C.C., Flyvbjerg, B., Molin, E.J.E. and van Wee, B. (2010), "Cost overruns in large-scale transport infrastructure projects: explanations and their theoretical embeddedness", *European Journal of Transport Infrastructure Research*, Vol. 10 No. 1, pp. 5-18.

Cantarelli, C.C., van Wee, B., Molin, E.J. and Flyvbjerg, B. (2012), "Different cost performance: different determinants? The case of cost overruns in Dutch transport infrastructure projects", *Transport Policy*, Vol. 22, July, pp. 88-95.

Cicmil, S. and Braddon, D. (2012), "Fading glory? Decision-making around the project – how and why 'glory' projects fail", in Williams, T. and Samset, K. (Eds), *Project Governance: Getting Investments Right*, Palgrave Macmillan, Basingstoke, pp. 221-255.

de Rus, G. and Socorro, M.P. (2010), "Infrastructure investment and incentives with supranational funding", *Transition Studies Review*, Vol. 17 No. 3, pp. 551-567.

Eisenhardt, K. (1989), "Agency theory: an assessment and review", *Academy of Management Review*, Vol. 14 No. 1, pp. 57-74.

Eskerod, P., Huemann, M. and Ringhofer, C. (2015), "Stakeholder inclusiveness: enriching project management with general stakeholder theory", *Project Management Journal*, Vol. 46 No. 6, pp. 42-53.

Flyvbjerg, B. (2006), "Five misunderstandings about case-study research", *Qualitative Inquiry*, Vol. 12 No. 2, pp. 219-245.

Flyvbjerg, B. (2009), "Survival of the unfittest: why the worst infrastructure gets built – and what we can do about it", *Oxford Review of Economic Policy*, Vol. 25 No. 3, pp. 344-367.

Flyvbjerg, B., Bruzelius, N. and Rothengatter, W. (2003), *Megaprojects and Risk: An Anatomy of Ambition*, Cambridge University Press, Cambridge.

Flyvbjerg, B., Garbuio, M. and Lovallo, D. (2009), "Delusion and deception in large infrastructure projects: two models for explaining and preventing executive disaster", *California Management Review*, Vol. 51 No. 2, pp. 170-193.

Flyvbjerg, B., Holm, M.S. and Buhl, S. (2002), "Underestimating costs in public works projects: error or lie?", *APA Journal*, Vol. 68 No. 3, pp. 279-295.

Goodspeed, T.J. (2002), "Bailouts in a federation", *International Tax and Public Finance*, Vol. 9 No. 4, pp. 409-421.

Ho, S.P., Levitt, R., Tsui, C.-W. and Hsu, Y. (2015), "Opportunism-focused transaction cost analysis of public-private partnerships", *Journal of Management in Engineering*, Vol. 31 No. 6.

Iossa, E. and Martimort, D. (2015), "The simple microeconomics of public-private partnerships", *Journal of Public Economic Theory*, Vol. 17 No. 1, pp. 4-48.

Joslin, R. and Müller, R. (2016), "The relationship between project governance and project success", *International Journal of Project Management*, Vol. 34 No. 4, pp. 613-626.

Klakegg, O.J. and Volden, G.H. (2016), "Governance in public projects: the Norwegian case", in Müller, R. (Ed.), *Governance and Governmentality for Projects: Enablers, Practices and Consequences*, Routledge, New York, NY, pp. 129-156.

Kvalheim, E.V. (2015), "Kan man stole på samfunnsøkonomiske analyser? En gjennomgang av elleve analyser av prosjektet Stad skipstunnel", Concept Working Paper, NTNU, Trondheim.

Kvalnes, Ø. (2014), "Honesty in projects", *International Journal of Managing Projects in Business*, Vol. 7 No. 4, pp. 590-600.

Laffont, J.-J. and Martimort, D. (2002), *The Theory of Incentives: The Principal-Agent Model*, Princeton University Press, Princeton, NJ.

Lefley, F. (2006), "Can a project champion bias project selection and, if so, how can we avoid it?", *Management Research News*, Vol. 29 No. 4, pp. 174-183.

Liu, J., Gao, R., Cheah, C. and Luo, J. (2016), "Incentive mechanism for inhibiting investors' opportunistic behavior in PPP projects", *International Journal of Project Management*, Vol. 34 No. 7, pp. 1102-1111.

Locatelli, G., Mariani, G., Sainati, T. and Greco, M. (2017), "Corruption in public projects and megaprojects: there is an elephant in the room", *International Journal of Project Management*, Vol. 35 No. 3, pp. 252-268.

Lovallo, D. and Kahneman, D. (2003), "Delusions of success: how optimism undermines executives' decisions", *Harvard Business Review*, Vol. 81 No. 7, pp. 56-63.

Mackie, P. and Preston, J. (1998), "Twenty-one sources of error and bias in transport project appraisal", *Transport Policy*, Vol. 5 No. 1, pp. 1-7.

Merton, R. (1936), "The unanticipated consequences of purposive social action", *American Sociological Review*, Vol. 1 No. 6, pp. 894-904.

Meyer, W.G. (2014), "The effect of optimism bias on the decision to terminate failing projects", *Project Management Journal*, Vol. 45 No. 4, pp. 7-20.

Miller, R. and Hobbs, B. (2005), "Governance regimes for large complex projects", *Project Management Journal*, Vol. 36 No. 3, pp. 42-50.

Morris, P.W. (2013), *Reconstructing Project Management*, John Wiley, Chichester.

Morris, P.W.G. and Hough, G.H. (1991), *The Anatomy of Major Projects: A Study of the Reality of Project Management*, John Wiley, Chichester.

Müller, R. (2009), *Project Governance: Fundamentals of Project Management*, Gower, New York, NY.

Müller, R. and Turner, J.R. (2005), "The impact on principal-agent relationship and contract type on communication between project owner and manager", *International Journal of Project Management*, Vol. 23 No. 5, pp. 398-403.

Newby, T.M.J. (2010), "Unintended effects of development aid – a brief overview", DIIS Working Paper No. 2010:06, Danish Institute for International Studies, Copenhagen.

Oates, W.E. (1999), "An essay on fiscal federalism", *Journal of Economic Literature*, Vol. 37 No. 3, pp. 1120-1149.

O'Leary, T. (2012), "Decision-making in organizations", in Williams, T. and Samset, K. (Eds), *Project Governance: Getting Investments Right*, Palgrave Macmillan, Basingstoke, pp. 175-220.

Ostrom, E., Gibson, C., Shivakumar, S. and Andersson, K. (2001), *Aid, Incentives, and Sustainability: An Institutional Analysis of Development Cooperation*, SIDA Studies and Evaluation 02/01, Swedish International Cooperation Agency, Stockholm.

Pinto, J. and Patanakul, P. (2015), "When narcissism drives project champions: a review and research agenda", *International Journal of Project Management*, Vol. 33 No. 5, pp. 1180-1190.

Pinto, J.K. (2006), "Organizational governance and project success: lessons from Boston's big dig", paper presented at Concept Symposium, NTNU, Trondheim.

Rattsø, J. (2003), "Vertical imbalance and fiscal behaviour in a welfare state: Norway", in Rodden, J.A., Eskeland, G.S. and Litvack, J. (Eds), *Fiscal Decentralization and the Challenge of Hard Budget Constraints*, MIT Press, Cambridge, MA and London, pp. 133-161.

Rodden, J.A., Eskeland, G.S. and Litvack, J. (2003), "Introduction and overview", in Rodden, J.A., Eskeland, G.S. and Litvack, J. (Eds), *Fiscal Decentralization and the Challenge of Hard Budget Constraints*, MIT Press, Cambridge, MA and London, pp. 3-33.

Samset, K. (2003), *Project Evaluation: Making Projects Succeed*, Tapir Akademisk Forlag, Trondheim.

Samset, K. and Volden, G.H. (2012), "The proposal", in Williams, T. and Samset, K. (Eds), *Project Governance: Getting Investments Right*, Palgrave Macmillan, Basingstoke, pp. 46-80.

Samset, K. and Volden, G.H. (2016), "Front-end definition of projects: ten paradoxes and some reflections regarding project management and project governance", *International Journal of Project Management*, Vol. 34 No. 2, pp. 297-313.

Samset, K., Volden, G.H., Welde, M. and Bull-Berg, H. (2014), "Mot sin hensikt. Perverse insentiver – om offentlige investeringsprosjekter som ikke forplikter", Concept report no 40, Ex Ante Academic Publisher.

Shenhar, A.J. (2004), "Strategic project leadership®: toward a strategic approach to project management", *R&D Management*, Vol. 34 No. 5, pp. 569-578.

Siemiatycki, M. (2016), "The making and impacts of a classic text in megaproject management: the case of cost overrun research", *International Journal of Project Management*, Vol. 36 No. 2, pp. 362-371.

Solberg, H.A. and Preuss, H. (2007), "Major sports events and long-term tourism impacts", *Journal of Sports Management*, Vol. 21 No. 2, pp. 213-234.

Solberg, H.A. and Preuss, H. (2015), "Major sports events: the challenge of budgeting for the venues", *Event Management*, Vol. 19 No. 3, pp. 349-363.

Tiebout, C. (1956), "A pure theory of local expenditures", *Journal of Political Economy*, Vol. 64, pp. 416-424.

Tirole, J. (1994), "The internal organization of government", *Oxford Economic Papers*, Vol. 46 No. 1, pp. 1-29.

van Wee, B. (2007), "Large infrastructure projects: a review of the quality of demand forecasts and cost estimations", *Environment and Planning B: Urban Analytics and City Science*, Vol. 34 No. 4, pp. 611-625.

van Wee, B. and Rietveld, P. (2013), "CBA: ex-ante evaluation of mega-projects", in Priemus, H. and van Wee, B. (Eds), *International Handbook on Mega-Projects*, Edward Elgar, Cheltenham, pp. 269-290.

Vann, M.G. (2003), "Of rats, rice and race: the great Hanoi rat massacre, an episode in French colonial history", *French Colonial History*, Vol. 4, pp. 191-203.

Volden, G.H. and Andersen, B. (2018), "The hierarchy of public project governance schemes: an empirical study of principles and practices in Norwegian ministries and agencies", *International Journal of Managing Projects in Business*, Vol. 11 No. 1, pp. 174-197.

Volden, G.H. and Samset, K. (2015), "Perverse incentives in the front-end: public funding and counterproductive projects", paper presented at the IRNOP Conference, London, June 22-24.

Volden, G.H. and Samset, K. (2017a), "Governance of major public investment projects: principles and practices in six countries", *Project Management Journal*, Vol. 48 No. 3, pp. 90-109.

Volden, G.H. and Samset, K. (2017b), "Quality assurance in megaproject management: the Norwegian way", in Flyvbjerg, B. (Ed.), *The Oxford Handbook of Megaproject Management*, Oxford University Press, Oxford, pp. 406-427.

Wachs, M. (1987), "Forecasts in urban transport planning: uses, methods, and dilemmas", *Climatic Change*, Vol. 11 Nos 1/2, pp. 61-80.

Wachs, M. (1989), "When planners lie with numbers", *Journal of the American Planning Association*, Vol. 55 No. 4, pp. 476-479.

Welde, M. and Odeck, J. (2017), "Cost escalations in the front-end of projects – empirical evidence from Norwegian road projects", *Transport Reviews*, Vol. 37 No. 5, pp. 612-630.

Wiig, H. and Holm-Hansen, J. (2014), "Unintended effects in evaluations of Norwegian aid: a desk study", Norwegian Agency for Development Cooperation, Oslo.

Wildasin, D.E. (2004), "The institutions of federalism: towards an analytical framework", *National Tax Journal*, Vol. 57 No. 2, pp. 247-272.

Williams, T. and Samset, K. (2010), "Issues in front-end decision making on projects", *Project Management Journal*, Vol. 41 No. 2, pp. 38-49.

Williams, T., Klakegg, O.J., Magnussen, O.M. and Glasspool, H. (2010), "An investigation of governance frameworks for public projects in Norway and the UK", *International Journal of Project Management*, Vol. 28 No. 1, pp. 40-59.

Yin, R. (2014), *Case Study Research. Design and Methods*, 5th ed., SAGE.

Zwikael, O. and Smyrk, J. (2012), "A general framework for gauging the performance of initiatives to enhance organizational value", *British Journal of Management*, Vol. 23 No. S1, pp. 6-22.

Zwikael, O. and Smyrk, J. (2015), "Project governance: balancing control and trust in dealing with risk", *International Journal of Project Management*, Vol. 33 No. 4, pp. 852-862.

About the author

Gro Holst Volden holds the position of Research Director of the Concept Research Program on Front-end Management of Major Investment Projects at the Norwegian University of Science and Technology in Trondheim, Norway. Her research is within project governance, public decision processes, and assessment and evaluation of major public investments. She is Economist from the Norwegian School of Economics, and has a prior career as Senior Advisor in the consulting industry and in the government bureaucracy, aiming to increase the value for money of investments, regulations and other public measures. Gro Holst Volden can be contacted at: gro.holst.volden@ntnu.no

Paper III

Section 2. The Proposal

Knut Samset and Gro Holst Volden

Contents

Abstract

This section discusses different views and perspectives on success, and how projects can be defined in terms of their tactical and strategic performance. Further how quality-at-entry can help to assure project performance and viability. It takes a look at what a proposal for a project is and what it needs, with a particular view on the alignment of needs, objectives and anticipated effects. It presents an approach to defining possible project concepts, and how to identify systemic borders to guide the selection of alternatives. Strength and weaknesses of making essential analyses up front with very limited information is highlighted, and how to apply probability assessment and logic in judgmental assessments. There is a special focus on estimating cost and benefit throughout the front-end phase of the project. Also, some essential lessons and common pitfalls in projects that can be associated with analyses and decisions in the earliest phases are discussed.

Finally, financing mechanisms for projects and their characteristics are presented, their features and strength and weaknesses under different circumstances: Public-private partnership, central government funding, Public and private loans, soft loans, Government guaranteed funding, central/local government cost sharing, etc. The issue of providing incentives with or without liability for the users (perverse incentive) is also discussed.

1. Different views and perspectives on success

Measuring success in projects is not a simple and straightforward undertaking. That is because the term "success", used as an indicator, is a highly complex and aggregated measure. For one, the degree of success is time-dependent. Secondly, it may be interpreted differently by different individuals and institutions. Thirdly, it tends to be measured differently in different types of projects, depending on the nature of their immediate outputs and long-term outcome. And finally, different individuals tend to assess the success of the same project differently depending on their preferences, values and to what degree they are affected by the project.

To illustrate how success is affected by time the track-record of the Empire State Building in New York can serve as an example. It was commissioned 1929 by General Motors, who wanted to exceed the height of rival car manufacturer Chrysler's building. It was completed one year ahead of schedule, almost 50 per cent below budget (helped by the depression), and to the specifications as designed. In the immediate perspective the project would therefore be a complete success. However, only 20 per cent of the building space was rented at the building's opening, so it was nick-named the Empty State Building. It took 17 years for the building to have enough tenants to turn a profit. It has been a success ever since and is again the tallest building in New York with almost 100 per cent tenancy. The project went from success, to failure and then success again.

Success is measured differently in different types of projects, depending of the nature of their immediate output and more long-term outcome. A hospital is assessed in terms of its health benefits, an industrial project might be judged essentially in financial terms, and an infrastructure project in term of its utility.

The assessment of success can be in absolute or in relative terms - that is in relation to what was agreed versus what was realistically achievable. *Ambition* is expressed in terms of the project's stipulated objectives. Its *outcome* is a direct measure of what has been actually achieved. Clearly, success measured in absolute terms may give a misleading conclusion if objectives are unrealistically ambitious. By measuring in relative terms, that is in relation to what could reasonably be expected as compared with experiences in similar projects - the same project might possibly be considered a success.

Williams (2008) suggests that the archetypical "man in the street" would be likely to think of projects as generally unsuccessful. A key word often associated with them in the public's mind is the English colloquialism "white elephant" (something whose cost and subsequent upkeep is much greater to the owner than its value, deriving from the reputed practice of monarchs giving sacred white elephants as gifts), Morris and Hough (1987), concluded that "the track record of projects is fundamentally poor, particularly for the larger and more difficult ones.… Projects are often completed late or over budget, do not perform in the way expected, involve severe strain on participating institutions or are cancelled prior to their completion after the expenditure of considerable sums of money."

More than two decades ago Pinto and Slevin (1988), concluded that "the concept of project success has remained ambiguously defined both in the project management literature and,

indeed, often within the psyche of project managers. Projects are often rated as successful because they have come in on or near budget and schedule and achieved an acceptable level of performance. Other project organisations have begun to include the client satisfaction variable in their assessment of project success. Until project management can arrive at a generally agreed upon determinant of success, our attempts to accurately monitor and anticipate project outcomes will be severely restricted".

Success as a generic term means to gain advantage, superiority, accomplishment, achievement or added value. One interpretation of project success is that the stakeholders that are part of or affected by a project are satisfied. Being such a compound measure, success will have to be translated into a hierarchy of indicators that would enable measuring. Wideman (1996, p3-4) describes a sequential set of four success measures, all of them time dependent: (1) "internal project objectives (efficiency during the project), (2) benefit to customer (effectiveness in the short term), (3) direct contribution (in the medium term) and (4) future opportunity (in the long term)". Three of these measures go beyond the project's immediate outputs. There are many examples of projects that score highly on efficiency, but subsequently prove to be disastrous in terms of their effect and benefit. There are also numerous projects that failed to pass the efficiency test but still prove to be tremendously successful both in the short and long run.

Clearly, a successful project is one that delivers its outputs and significantly contributes to the fulfilment of agreed objectives. Moreover, it should have only minor negative effects, its objectives should be consistent with needs and priorities in society, and it should be viable in the sense that the intended long-term benefits resulting from the project are produced. These requirements were first formulated for US-funded international development projects by the United States Agency for International Development (USAID) in the 1960s, and subsequently endorsed by the United Nations (UN), the Organisation for Economic Co-operation and Development (OECD), and the European Commission (EC). They comprise five requirements or success factors that have to be fulfilled: the project's efficiency, effectiveness, relevance, impact and sustainability. These are tough requirements that go far beyond the issues that usually are covered by the media or indeed many planners and decision-makers.

1.	**EFFICIENCY**	Delivery of outputs in term of scope, timing and cost in relation to what was agreed
2.	**EFFECTIVENESS**	The extent to which the objective has been achieved
3.	**IMPACT**	All other positive and negative changes and effects caused by the project
4.	**RELEVANCE**	Whether the objectives are aligned with valid priorities and users' needs
5.	**SUSTAINABILITY**	Whether the positive effects of the project will be sustained after the project has been concluded

2. Tactical and strategic performance

In applying the success criteria above, we distinguish between the projects' *tactical* and *strategic* performance. Success in tactical terms typically means meeting short-term performance targets, such as producing agreed outputs within budget and on time. These are essentially project management issues. Strategic performance, however, includes the broader and longer-term considerations of whether the project would have a sustainable impact and remain relevant and effective over its lifespan. This is essentially a question of getting the business case right, or, in short, of choosing the most viable project concept.

This is illustrated in Figure 2.1. Tactical performance is a question of how the project is implemented, i.e. how inputs are converted into outputs. These are measures of its *efficiency*, here measured in terms of the cost, timing and quality of deliverables. Strategic performance is a question of how the project performs after the outputs have been delivered. This will have to be monitored with the more compound measures mentioned above, which would cover the broader and long-term perspectives. It would to a lesser degree involve focusing on technology and management issues, but more on societal and economic aspects.

The Empire State building was a success in tactical terms from the very start. The economic depression, with low costs and abundance of cheap labour offered a golden opportunity for this type of investment. That affected the tactical performance favourably. Then for 17 years it was considered a failure in strategic terms, until it had got enough tenants to turn a profit. In times of depression the market for expensive offices in the centre of New York had collapsed, and the project was therefore not *relevant*. The building was underutilized, running at a loss, and therefore also not *effective* or *sustainable*. With time, things changed radically and the project became a huge success, not only in terms of its relevance, effectiveness and sustainability, but also its *impact*, since it rapidly became a major tourist attraction and national symbol.

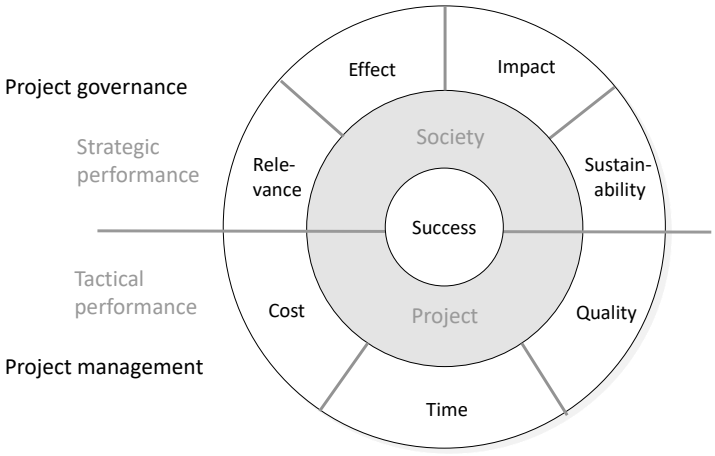


Figure 2.1 Successful projects. Tactical performance is a question of delivering the project outputs as planned, while strategic performance is the worth or utility of the project as seen in a long-term perspective

Generally speaking, tactical considerations typically are restricted in time and perspective, such as the presumed ability to meet short-term performance targets and trade-offs to keep stakeholders on board, all likely to prove ephemeral when matched against the lifespan of most projects. This includes the often proclaimed success of a project simply because it has come in "on cost and on time." Strategic performance is the key issue, but strategic success will only emerge over time in the context of the project having sustainable impact and remaining relevant and effective over its lifespan.

Projects that score highly on all the five success criteria mentioned above are those that perform successfully both tactically and strategically. Such projects may be rare cases. The tactical performance of projects is of less concern at the early project proposal stage. Strategic performance is the main issue. But to what extent is it possible to make a meaningful assessment of the proposal in relation to the above success measures? Would the amount and types of information available early on be sufficient?

In order to assess the *efficiency*, the costs of a project and the nature of its delivery are reasonably well understood at an early stage. But there's doubt as to whether cost estimates are realistic and the conditions of implementation will allow that outputs are produced as anticipated. Consequently, attempting to gauge efficiency may not be worthwhile in the front end phase. Not least the complications facing planners and decision makers in estimating realistic costs clearly indicate that the basis for rigorously evaluating efficiency usually is poor.

The same is true of *effectiveness*. Undoubtedly, the anticipated first-order effects usually are clearly known, their realization is time-dependent and relies on the fulfillment of other events outside the scope of the project. Realistic forecasting may therefore be notoriously flawed.

Early estimates of *impacts* are even more difficult. Undoubtedly, experiential knowledge may be acquired by studying similar projects. But we face conditions that are difficult to forecast and arguably require imagination and guesswork beyond our capabilities.

However, the situation for *relevance* differs. Common sense and user surveys, as well as knowledge of markets, laws and regulations permit us to form an early, accurate picture of whether an initiative is relevant. That we also are notoriously poor at this sort of early evaluation is not due to it being impossible, but rather to it not being done to a sufficient extent.

Forecasting future *sustainability* is also difficult. However, the question is closely related to whether the proposed project is relevant. Moreover, early on, we have long been able to realistically analyze cash flows.

Consequently, the answer to the question above is that with modest effort, we can gain a good picture of whether a project is relevant and sustainable. Extensive analyses of the other three criteria may not be worthwhile at an early stage. The good news is that relevance and sustainability are precisely the attributes that determine whether a project will be successful or

not in the long term. Therefore, this may be a minimalistic answer to the question raised, or a quick-and-dirty approach to ex ante evaluation of project proposals, in which the benefits are great compared to the cost.

3. Quality at entry

One of the prime goals of strategic planning is to attain structured and effective continuous management. The strategy shall conduce decision makers at various levels to pull in the same direction by providing a common long-term goal to keep in mind in making decisions. Research has shown that this is essential to attaining good results (Heijden 1996). A study of 1125 projects compared the extent and quality of pre-project studies, appraisals and design prior to project inception with whether or not they were successful. The conclusion was that 80% of the projects that scores highly in terms of their quality at entry (QaE) were successful, compared to just 35% of those that were started without proper preparation (World Bank 1996).

A survey conducted in the USA of about 600 project managers contributed to identifying which critical factors influence the level of achievement in projects. The conclusion was that planning the project was by far the most important factor. Problems that could have been avoided with a better project plan arose repeatedly during the entire project cycle (Pinto and Slevin 1988).

An international project that drew on the experience of 60 large infrastructure programmes concluded that projects with great strategic depth, that is an appreciable level and extent of strategic assessments that underlie a project, were more likely to be successful. A clear concurrence showed that the projects that attained the best results had allocated greater portions of their overall costs to their front end phases. These cost allocations varied from 3% for simple projects and as much as 35% for complex projects. The costs in the front end phase, before the decision was made to start, varied from 15 to 500 million USD. The conclusion was that such costs often were justified and resulted in considerable cost reductions in the implementation phase, more socially acceptable projects and better risk management (IMEC 1999). The study also found that in particular, three aspects characterized the most successful projects: (1) the front end phase had been long, that is, several years, (2) the concept had been revised several times, and (3) problem solving was systematic and inclusive. Moreover, it was found that the use of risk analysis was vital and that there was a decided advantage in open debate on project planning.

Typically, the less successful projects resulted from authoritative choices made by investors, public agencies or strong interest groups and often were carried out under time pressure. Little time was allocated to pre-project studies or to evaluation or appraisal of concepts. The original concept was maintained to save time, with insufficient emphasis on acquiring relevant information. Consequently, in many cases, projects had conflicting goals and were based on assumptions imposed by interest groups or the authorities.

Paradoxically, the greater portion of resources expended to ensure project success is not used up front but during the implementation phase. Moreover, the greater part of the resources expended up front is used to work out a relatively detailed strategic plan, while only a relatively smaller part of the resources are used in concept development, that is to identify and test alternatives and delineate a strategic framework for the final project.

This is perhaps one of the principal problems in project activities in general. Early on, before the project or process is initiated, there's often too little attention given to fundamental questions concerning the concept itself. There are fine-tuned, resource-intensive precision instruments for controlling processes that don't detect whether the concept is sensible or not. In many cases, the methods are used to marginally improve concepts that should have been discarded and thereby contribute to upholding them. Projects with budgets in the billions are precision controlled in time spent, costs incurred and quality delivered, while the choice of concept itself is insufficiently considered.

This is due in part to the complexity of assessments, as they depend not only on knowing, but also on foreseeing. In the initial phase, uncertainty is greatest and the amount of reliable, factual information smallest. At that stage, there is not a great diversity of methods to apply, because information is scarce, qualitative and often judgemental. Consequently, it makes little sense to use precision instruments.

4. Tactical flexibility

Strategic planning is only part of the solution, *tactical flexibility* is equally important, to allow for manoeuvring within the strategic framework delineated, as the project is implemented. Additionally, there should be latitude for changing the strategic perspective if that becomes necessary. Strategic planning is built on judgment and assumptions and don't necessarily identify the most suitable choices in situations that may arise. Requiring that a strategic plan be followed strictly can make it a straightjacket. In practice, this means that there's little sense in formulating a detailed strategic plan early on.

This line of thought is underscored by Napoleon Bonaparte's remark on planning, *"Plans are nothing, but planning is everything."* The creative, initial planning process affords decision makers the opportunities of identifying and assessing the key alternatives and of finding the way to a sensible, realistic concept. Planning helps decision makers think through alternatives and thereby become better equipped when they later are faced with situations in which they must make tactically vital choices. In some cases, these tactical choices will influence and change the strategy. Normally, the chances are limited that a precise strategy will be implemented in detail strictly as laid out.

A plan presupposes a degree of determinism, a quality of information and a clear cause-effect relationship that at best exists only in the implementation phase. It allows only cursory consideration at an early point in time of the inconceivability of foreseeing the interplay between various involved or affected parties over time, of the incompleteness of information and of the cause-effect relationship being influenced by uncertainty that can change the analytic context that comprises the base of the goals and strategic choices undertaken.
A story often cited in the project literature concerns a Swiss military troop that returned exhausted to base camp after three days in a blizzard high in the Alps. By the troop leader's recount, the men had lost their way and thought that they were doomed until one of them found an old map in his pocket. Courage renewed, the men found shelter, waited until the storm subsided and then used the map to find their way out of the area. Afterwards, they were astonished to find that the map was of the Pyrenees, not the Alps.
The story is used to show that in a situation with high uncertainty, it's not necessarily the quality of the strategic instrument that counts, but rather the tactical response chosen. Yet strategy can be useful even though it's completely wrong. It's principally an aid to point out a main direction. A detailed strategy strictly followed can be worse than any strategy.

There are differing perceptions of what a strategy is: The ideal would be the rational, i.e. that it is (1) a designated plan based on thorough preparation and designed to serve a specific purpose. In many cases however it is merely (2) a pattern, i.e. a standardized scheme which is applied over and over again in different projects. And finally, which is also quite common, it is merely a ploy which is designed to serve a different purpose, for instance to stir interest in a scheme or attract funding. Mintzberg, H. (2005)

Consequently, the question of which concept is best concerns more than the systematic, rational identification and assessment of various alternatives. In the front end phase, the interests and prioritizations of various parties become evident, intervene and lead to decisions that often are far from that which appeared logical and rational at the outset. Hence, understanding this process is as vital as questions regarding the information base and the rational analysis choice of method.

Major public projects are typically conceived as the result of politically expressed needs in dialogue between various stakeholders. This is followed by a lengthy process of developing the project and making the necessary decisions, typically involving the government at various administrative levels, but also political institutions, the public, the media, and consultants and contractors in the private sector. Such processes are often complex and unpredictable. The processes can also be deceptive and irresponsible, affected by hidden agendas rather than openness and social responsibility. Flyvbjerg, B., Bruzelius, N. and Rothengatter, W. (2003). These aspects are discussed further in sections 5. – 7. of this book.

5. The project proposal

As discussed, it is a long way from analytic results to decisions and actual project realization. The merit of the decision basis is central. Clearly, its quality cannot be assessed solely on the grounds of the methods used or the quality of the input data, but must be viewed in connection with what happens later in the process. All too often the decision basis is restricted to a detailed assessment of just one alternative concept. The basis for decision could for example essentially be a probabilistic analysis, which results in expected values of costs and time expenditure that are considered favourable. Such a decision basis is too narrow in most projects. The assessments must embrace more than the narrow implementation perspective. They should also consider the long-term consequences of the project. Moreover, they need to build on real assessments of alternative conceptual solutions. Studies of the scope and quality of decision bases of projects in general show that this often is not the case.

Decisions may be made on very simple bases. One might toss heads or tails, or, if reliable information is available, undertake a simple assessment of foreseen reward relative to costs. But the decision bases also may be comprehensive. Projects of some extent usually have a thorough, detailed pre-project study. In some cases, a pre-project study may take years and include complex analyses, simulations, pilot studies, etc. Studies of managerial uses of decision information have shown that many managers first decide on the basis of their own experience and intuition, perhaps after having conferred with persons they trust. Thereafter, available information is used to support the decision, not as a basis for making a decision.

However, in many cases, the type and extent of studies in the *initial phase* often are severely limited. The terms of the final project often are shaped more by the events of the initial phase than by the pre-project study. At that point of time, the terms of the pre-project study are determined. With a prior, top-down assessment of the concept itself, strategic guidance could be included in an initial phase that also puts the pre-project study on a sensible track. This may be extremely useful, both in the short term and the long term, not least because the costs of the initial, broad and often qualitative concept studies are relatively small.
The extent of effort in the initial phase of a project apparently is either pretty limited or relatively comprehensive. This may be ascribed to formal requirements, such as for impact assessments and quality assurance, first being imposed when the project exceeds a certain size. Hence, there is no accepted tradition for systematic front end phase appraisal of smaller projects. The same is true of systematic use of risk analyses in project activities. Today, there are no widely used method tools or standards for such analyses.

A project proposal should ideally include the following steps and elements:

1. A needs analysis mapping all stakeholders and affected parties and assessing the project's relevance in relation to needs and priorities in society.
2. A specification of all requirements that need to be fulfilled when the project is implemented (e.g. functional, aesthetic, physical, operational and economic)
3. An overall strategy defining the project's goal and purpose (first order and long-term effects) with emphasis on consistency, realism and verifiability.
4. Specification of concepts that might be considered as alternative solution to realize the identified strategy
5. An alternatives analysis including a full economic analysis and risk analysis involving at least two alternative main concepts and the zero-option (doing nothing). The analysis should evaluate the proposed alternatives with emphasis on:

    - **Relevance** in relation to
        - Needs
        - Societal priorities
        - Existing portfolio of projects under the responsible ministry/agency
    - **Feasibility** in relation to
        - Proposed budget
        - Time frame
        - Quality of outputs
        - Composition and timing of elements in the total project
    - **Sustainability** in the operational phase with emphasis on
        - Long-term economic benefit
        - Financial sustainability
        - Uncertainties
6. Rank the proposed alternatives and provide recommendations regarding decision strategy and implementation strategy for the project.

The project proposal should be prepared by the end of the pre-study phase at a time when the choice between alternative concepts is still open. It should guide decision makers by

providing a sound basis for the decision whether or not to initiate a pre-project with further investigation of alternatives.

After the pre-project phase, the project proposal should be followed up with an overall project management document with the aim to ensure the quality of the decision basis, including cost estimates and uncertainties associated with the chosen project alternative before it is submitted for final approval and funding. The document would typically include information on the following:

1. Outline of the strategy for the project (what, when, where, how and by whom, and so on).
2. Scope of activities and assumptions
3. Financial analysis including cost estimates
4. Assessment of return on investment and effort
5. Risk assessment with identification and classification of possible threats or risks during implementation and subsequent operational phase
6. Technological assessment of feasibility, technological risk, trends and outlook for relevant technologies
7. Environmental impact assessment (EPA)
8. Project management structure and administration
9. Contract strategies and associated risks

The relationship between strategic management and project-management is not one-way. Morris (2008) describes how strategy implementation is accomplished with project management, but project management can also contribute to strategic management. He points out that project management's contribution "can add value to the emerging strategy and ensure that benefits are reaped from its realisation." The strategy is or ought to be a major concern to both parties, because it lays out the direction and justification for the project in a long-term perspective. Alignment of needs and objectives is a key issue.

6. Alignment of needs, objectives and anticipated effects.

Projects in a typical management environment, public or private, can often be said to be in a "wicked mess". "Projects are complex, ambiguous, confusing phenomena wherein the idea of a single, clear goal is at odds with the reality". Linehan and Kavanagh, (2004) Engwall (2002), describes the establishment of the perfectly correct goal as a "futile dream". For projects to be aligned with organisational strategy – *and stay aligned* – it is important to recognise the turbulence of the environment, and build in the capability to cope with this turbulence at the start of the project. Miller and Hobbs (2005) suggest that this is equally important when the project is being undertaken by a heterogeneous consortium or group of organisations, where processes and structures need to be developed to deal with turbulence.

Therefore, flexibility needs to be built into the project strategy, both in the front-end concept stage, and in later stages. Olsson (2006) shows the need for tactical flexibility within a defined strategy, and Samset (2010) points out the danger in seeking predictability. He warns that "prediction [can] become a prescription…it shifts the decision-makers focus from finding the best solution to …[making] his own idea or prescription come true". Premature lock-in to an inappropriate concept can be a major danger to project success.

At the earliest phase however, and despite of remonstrances such as those mentioned above, there is a need to ensure alignment of *needs, goals* and anticipated *effects*. Strategies are designed in response to certain needs. The phenomena of needs, goals and effects are closely related, and they need to be compatible in the sense that the causality or logic between them is right.

- The goal specifies the need formally in terms of scope, time and quality
- The effect should correspond at least to the anticipated results specified by the goal
- The gross effect should be such that the needs are satisfied

For example, a hydroelectric power project is initiated and planned to meet a need for electric power in the market. The project is to build a facility with a stated capacity. The goal is to attain stable delivery to the grid at that level. Need and effect often are expressed indirectly in derived units. For example, the triggering need and effect can both be expressed in economic terms, in this case respectively in production and consumption. As mentioned, the design of the project shall include the basic requirement of a connection between needs and effect. The goal shall be derived from the needs, and the effect shall at least correspond to the goal set for the enterprise. The needs must be real to attain the anticipated effect. Basic user and market research may be used to ascertain whether this is the case. The lack of user or market adaptation lowers the chances of success.

Alignment in this context would imply the following requirements:

- Needs are expressions of a future desired situation, and should not be expressed as one specific solution to the problem at hand. Needs should be expressed in a way that allow for alternative solutions or concepts to be considered
- The goal should specify what to be achieved as the result of the project and expressed in terms that can be measured. An intention becomes a goal if and only if measures are made to fulfill it. The goal should be realistically achievable as compared with the time and resources available, and uncertainties that might affect project implementation
- The effect expresses the degree to which the goal was achieved. The effect can only be established in retrospect. The combined or gross effect should also include any side effects that might be attributed to the project

As they are formulated and agreed upon, objectives are a project's prime success criteria. Formally viewed, success is ensured when a project is implemented as efficiently as possible and causes effects that concur with its objectives and correspond to the needs that triggered it. So formulating, furthering and following up objectives are a management function.
Large investment projects are complex and usually have several objectives that are more or less mutually dependent. Customarily, a hierarchy of objectives is defined to clarify how the various objectives relate to and support each other. The location of an objective in the hierarchy indicates how general or concrete it may be but doesn't necessarily indicate its importance. The hierarchy displays cause-effect relationships. So, to a degree, it indicates realizability, in other words, the ambitiousness of the individual objectives.

Studies of projects have shown that ambitious objectives motivate better performance, but also that performance drops when objectives are overly ambitious or completely unrealistic. Næss (2004) contends that in American literature, this is used to argue that objectives should

be formulated so that they are realistic, that is, they can be achieved with the means available. In Scandinavian literature, it's asserted that entirely realistic objectives (that we are certain that we can achieve) are insufficiently challenging in a continually changing world. Visionary objectives are needed to bring out the best performance. This means that overall objectives should be sufficiently ambitious to motivate yet be realistically attainable in time. Of course, impossible objectives are purposeless.

The formulation of an objective should indicate what's needed to attain it. This is what strategy sets forth. An objective may be expressed at the project or process level, such as building and furnishing a new opera, or at the organizational level, such as attaining a target market share or membership. Or it may be at the national level, such as in keeping inflation at a specific level.

The purpose of formulating an objective is principally to clarify the direction for that which is sought. The scope of that which is sought also needs to be stated so one may know when an objective is attained. Multiple objectives may confuse that which is sought if they point in different direction. This is particularly evident if the objectives also conflict with each other. The development of a new oil field hardly can be justified with an environmental objective, as the investment undeniably will result in increased emission of atmospheric pollutants. Here there's a conflict of objectives. Using an environmental objective for a hydroelectric project will not give rise to such conflict, disregarding other environmental aspects, such as those associated with the damming of watercourses.

Objectives should give rise to common understanding among and motivation of all parties involved in or affected by a project. On one hand, this means that objectives should be unambiguous and realistic. On the other hand, to motivate, they also have to be well founded, to the degree that they are accepted. Often, this isn't possible, simply because there are differing prioritizations and needs and because some parties simply may be opponents of the project.

Moreover, the objectives should limit the enterprise or the strategy. This means that the resources allocated and the results anticipated should correspond. Inadequate allocation of resources leads to insufficient conditions for realizing an output. If the objective is overly ambitious, the anticipated effect isn't achieved. Finally, objectives should be expressed in ways that permit assessing performance and results. This means that objectives are verifiable and measurable. Such requirements often are expressed in terms of SMART, a mnemonic for Specific, Measurable, Attainable, Realistic and Time-bound.

Practice often differs considerably from this ideal. A study of major Norwegian governmental investment projects conducted by the Ministry of Finance in 1999 found that the formulations of objectives were vague and overly ambitions, unrealistic and little suited to overriding management. The objectives stated mostly were activities or tasks, while there was no hierarchy of objectives between these extremes (Berg et.al. 1999). The finding was hardly unique. Rather, it seems to be commonplace practice, as corroborated by several studies, including (Samset 1998).

7. Defining project concepts, and systemic borders to guide the selection of alternatives

The generic notion of a concept designates an abstract idea or model that corresponds to something concrete in reality or in language. As used in the context of project definition, a concept is a construct of thought that is meant to solve a problem or satisfy specific needs. The concept should be of principal nature in the sense that several different concepts might be identified as solutions to the same problem. Further, in each specific case, all concepts ought to be real alternatives in the sense that they are mutually exclusive. This would imply that the concepts should have certain common features that make them suitable as solutions to the same problem. Finally, the quality of being principled means that the concepts are not just variations of a particular solution. This is illustrated in Figure 2.2, in which the investment case is distinguished from the project. The investment case is an abstract construction or an instrument used by the financer or the commissioner as a basis for appropriating funds, subsequently to be implemented in a project.

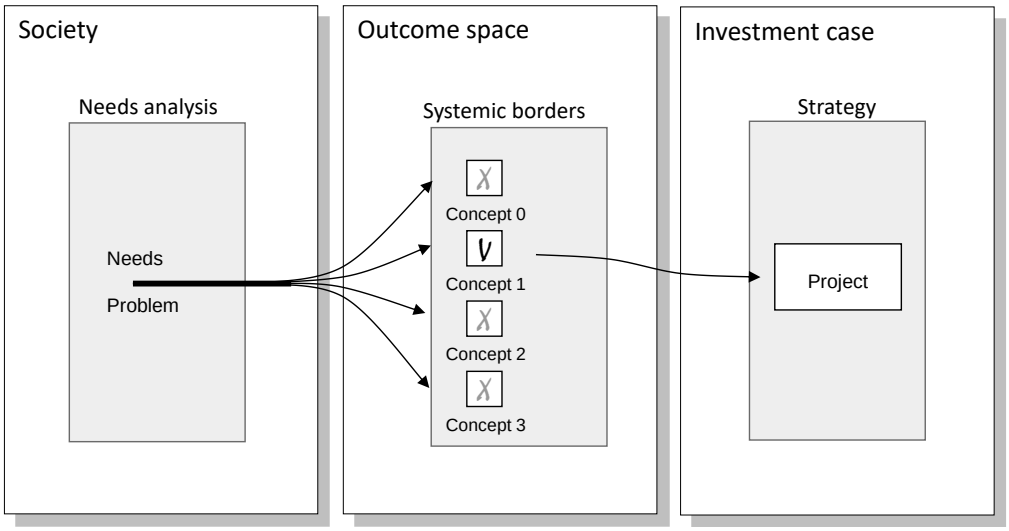


Figure 2.2 An investment case is implemented as a project after prior assessment of alternative concepts

As mentioned above, the needs, goals and effects are expressions of the same phenomenon that would be appear at three subsequent stages: up front, during implementation and in the operational phase. The point of departure is an undesirable condition in society, here called the problem, which is the cause that gives rise to a need. To satisfy the need, there must be a positive change, here called the goal. If the goal is realized, an effect is achieved so that the undesirable condition ceases. The original problem then is solved.

Needs
Intervention is necessary to make the cause-effect chain process work. This is called the concept. It comprises the actions that enable realization of the goal. The choice of concept then is guided by the original problem and the expected effect. Needs, goals and effects may

be defined at various levels. The less general the definitions of needs, the more it will provide guidance in the direction of specific types of solutions. That brings in the risk of the project not being suitable to attaining the overriding goals. There are numerous examples of needs analyses identifying one particular technical solution as a need, and that goals and impact assessments being constrained to concern implementation of a given main concept.

What this means in practice is illustrated by the following example of the planning of a transport project in an urban area suffering congestion of its main streets (Næss 2005). At the concept level, for example, the needs may concern reducing travel time between sectors of the urban area, prompting a more environmentally-friendly transport mode distribution and the furthering less travel-generating, car-dependent urban development patterns. The goals at this level must reflect these needs, and the effects of various solution concepts (and their relevant combinations must be assessed).

When a main concept, such as an urban tramway system, is chosen, demand analyses, goal setting and impact assessment will focus on ensuring that it is designed and implemented in the most socially acceptable manner. Needs and goals at this level may, for example, be concerned with attaining high passenger volumes, financially favourable and environmentally-friendly routing, and with contributing (through the locations of stations) to urban development in targeted areas.

Whenever demand analyses, goal setting and impact assessment at the strategic level are skipped, and instead the project level is initiated within the framework of a given solution, the initiators' needs can easily be confused with those of the society. Hence, the wishes of special interests for financial gains, prestige or ideologically preferable solutions may take precedence over top-down political goals and the needs of broader social groups. Such constraint of planning at a premature stage is a commonplace weakness in the planning of large, public investment projects.

Problems

The assumed effect is decisive to the choice of concept. But often the starting point is an undesired condition or a problem that initiates a search for a solution. In such cases, there are different aspects that need to be considered in determining a concept.

First, it's essential to focus on existing problems, not assumed, probable or future ones. Second, problems should not be expressed as absence of a particular solution. For example, the farmers' problem is not that they don't use pesticides, but that their crops are infested by pests. So there are considerable differences in the way the problem can be approached. There are many alternatives in addition to spraying crops. The problem therefore ought to express an existing undesirable condition, and it needs to be concrete. If the problem concerns traffic congestion, stating it in terms of too few traffic lanes points to just one solution.

Expressing the problem in more general terms gives latitude for several alternative solutions. Instead of directly dealing with the problem of too few traffic lanes, one but may seek other indirect solutions, such as by routing some traffic on other streets or by using other means of transport. All are solutions to the overriding problem, which in this case deals with traffic flow.

This example underscores another aspect, namely that the concepts chosen should be dissimilar. Nonetheless, they would have to share common characteristics suited to solving

the same problem. If that is not the case, they are merely variants of one set solution. Of course, the final choice of solution also needs to be assessed. But that should happen not at the concept level, but at the project level after the concept has been chosen.

The alternatives also have to be genuine, in the sense that they exclude each other. An oversimplified example is that if you want to start a family and have found two potential spouses, you normally are faced with two mutually exclusive alternatives, unless you wish to be a bigamist. If at the same time, you have three job offers, each in a different city, you have 2 x 3 = 6 mutually exclusive alternatives (Løwendahl and Wenstøp 2002).
We have no solid tradition for identifying truly alternative concepts as bases for designing projects. Most often, the choice is made at the starting point, and assessment is mainly at the project level. For example, in a study of a new national museum of art, architecture and design in Oslo, the choice was between alternatives that all featured co-location of the museums on the same site. The alternatives differed in distribution of space above and below ground, remote or central storage, and the like. So, obviously the concepts merely were variations on the same solution. Genuine alternatives would, for instance, look more closely at which museums should be co-located and where in the city or in the country they should be located. These aspects could then be weighed against the increased benefit envisioned. In this case, the problem, the anticipated effect and the benefit were all vague and gave no clear guide for choice of alternatives. Consequently, there was no substantive discussion of the reality of the proposal put forth.

The reason for the requirement of genuine alternatives is that it would stimulate creative thinking and thereby increase the chances of a good choice. Experience suggests that this is worthwhile. At the same time, we know that innovative thinking is no guarantee that it will happen. So there's a need to assess several alternatives. Moreover, these alternatives ought to be assessed against the zero option to avoid ending up with something that turns out to be worse than what existed.

There are no commonly agreed guidelines for a best practice for systematic identification and selection of unique and different solution to a problem, what is here termed concepts. Also, there are not a great many studies that offer a systematic inquiry into how this is done in practice, the range of alternative concepts that are identified, and which ones are chosen. One such study, which is not conclusive but that might offer some clues on the state of affairs in the Norwegian setting, concluded that (Minken et.al, 2009):

- The alternatives being considered were merely different technical solutions to the same problem than mutually exclusive concepts.
- The tendency was that the preferred technical solutions were used to guide the choice of concept rather than vice versa.
- The link between the choice of concept and the underlying societal need or problem was often not made explicit.
- The project-triggering need or problem would frequently be confused with other perceived needs or problems
- The anticipated, desired effect of the project was often confused with various positive or negative anticipated side effects.
- The zero-option, or the low-investment alternative solution, was often not identified, formulated or considered in relation to the alternative concepts being analyzed.

On other words, there is a strong tendency to choose the initial concept and stick to it. Experience also suggests that we tend to prefer incremental improvements of an inferior solution - rather than fundamental change. Also that there is an overwhelming inertia: once set in motion – a project being implemented is almost impossible to stop. On the positive side there is much to suggest that the window of opportunities is usually larger than envisioned - and largely unexplored.

The window of opportunities is exactly the same as what is termed the outcome space in Figure 2.2. It is delineated by the systemic borders that will define what can be identified as possible concepts. These borders would to a considerable degree translate into what would subsequently be the investment case's strategy or the strategic frame for the project. Laying out the systemic borders at an early stage is therefore much more essential than formulating objectives according to the SMART requirement which would eventually have to be done later, when the project is implemented.

One challenge would be to apply different perspectives in the quest for sensible conceptual solutions, both (1) the retrospective, looking at trends in the past, (2) the normative, identifying the desirable and useful, (3) the explorative, using projections to identify what is possible, (4) the interdisciplinary, to identify opportunities, uncertainty and risk, and (5) the counterfactual, the take a second look at the zero option, which would usually be the lowest cost alternative.

8. Front end analyses with limited information - strengths and weaknesses

When projects fail strategically, it is likely that the problem can be traced back to decisions in the earliest phases, when the initial idea was conceived and developed. What happens during the front-end phase is therefore essential for a project's success. There are different ways to improve quality-at-entry, for example by challenging initial ideas, extracting and making use of previous experience from similar undertakings, and consulting with stakeholders.

In most cases the key issue at the earliest stage is to shed sufficient light on the underlying problem that would provide the justification for the project, and the needs that the project is meant to satisfy. Detailed information about possible alternative solutions is less relevant. This illustrates what seems to be a major dilemma, since most projects originate as one specific solution to a problem, while the problem itself may not be analysed sufficiently, and alternative solutions may not have been considered at all. Typically, the preferred concept originates in the mind of one individual, based on intuition and experience, rather than systematic analysis of problems, needs, requirements, etc. Most of the information generated is associated only with the initially identified solution. A second dilemma is that this information, which may be very detailed and specific, tends to lock decisions into the initially preferred concept – to the extent that this will inevitably be the one that is finally chosen. It is all too rare that alternative concepts are identified and analysed to the extent that they get a fair trial in the subsequent decision process.

The gravity of this is obvious because this is exactly the stage when the fundamental choices are made, when uncertainty is at its highest, freedom to choose is at its optimum, and also when available information is most restricted. Adding information, therefore, makes sense - but only to a certain degree. However, some available information might not be relevant in the

decision-making process, and information that would seem necessary will not be available until later.

The crucial issue is not the volume but what type of information is needed. In the initial phase of a project the priority is to establish an overall perspective, and to analyse the problem in its context, considering the needs and priorities of stakeholders, users and affected parties, in order to come up with a sensible strategy. Opportunities and risks should also be considered. Experience suggests that creativity, imagination and intuition can be more valuable at this stage than large amounts of data. Therefore, lack of information in the earliest phase may not necessarily be a problem: it can even be to our advantage. Many planners have learnt that in the early phase of a project it can be of considerable help to operate primarily with qualitative expressions and only to a very limited degree with quantitative data.

Scheibehenne and von Helversen (2008), concludes that “less can be more”, and that having less information can actually help decision-makers. A restricted, but carefully selected, sample of relevant facts and judgemental information may be an advantage in the effort to establish a broad overall perspective, and identify and test alternative strategies. Omitting details and less relevant information helps avoid “analysis paralysis”, when decision-makers are presented with large amounts of detailed information too early in the decision-making process. Furthermore, accurate quantitative information tends to quickly become out of date. This is a problem, since the front-end phase in major projects may last for years, even decades. The phenomenon can be coined the “half-life of information” (Samset, 2008). For instance, exact information about the demand in a fast developing market will have limited value after months, even weeks. We cannot make a valid prediction of the actual demand three years into the future, but might be pretty certain that the demand will remain for a long time and therefore rely on it in strategic planning up front. In other words, carefully extracted qualitative information about a well thought-out project concept could provide reliable and valid input to the decision for the whole of the front-end phase.

What is of interest here is the principle that decisions need to be based on a foundation of assessment. The solidity of assessment depends on the selection of decision criteria, and the underlying information used to substantiate these. Each decision criterion needs to be substantiated with a number of parameters or indicators, producing an information hierarchy. The principle is valid, however, regardless of the type of information used. It can be factual or judgemental, quantitative or qualitative. For the assessment to be useful and trustworthy, the selected decision criteria need to both capture the essential aspects that ought to be considered, and be sufficiently comprehensive. Underlying supporting information needs to be valid and reliable. Reliability is a question of whether you can trust the information, this being determined by the quality of sources and the way it is collected. Validity is a term used to express the extent to which an indicator provides information that corresponds to what is to be measured. The type of indicators chosen will determine the validity of the assessment. Using several indicators at the disaggregate level helps improve validity at the aggregate level, providing that each indicator is valid. In most projects the five success criteria mentioned in the beginning of this section could be applied.

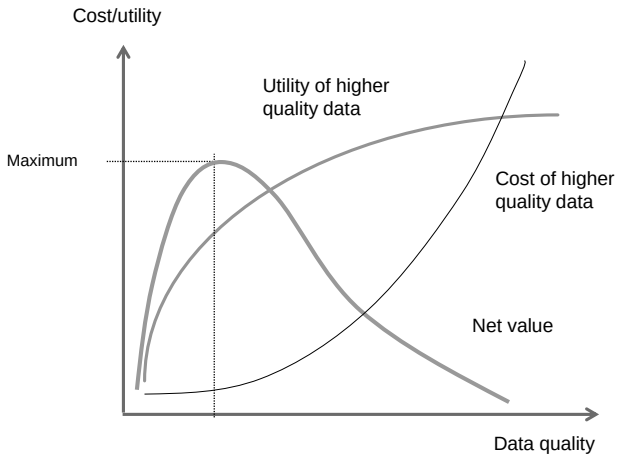


Figure 2.3 Trade-off between the amount/quality of information and the acquisition cost. Restricting precision to rather rough information giver the highest benefit/cost ratio.

Clearly, up-front decision making is not simply an issue of adding masses of information. As is illustrated in Figure 2.3, the cost of collecting information on a specific topic usually increases progressively with the amount of information collected. This is because more information requires more in-depth studies or more wide-ranging information searches. On the other hand, the gain in utility of additional information tends to decrease. This is because there is usually a critical amount of information that is needed to get the necessary insight in a situation. Additional information will be of limited use. Maximising the utility/cost-ratio will therefore set a limit to the amount of information that is useful. The maximum would typically be quite a bit to the left in the diagram, which would come as a surprise to many planners and decision makers alike.

Adding to this it is useful to point out that decisions may be affected in different ways:

- More by subjective or political priorities than by rational analysis.
- By priorities that may change over time.
- By changing alliances and pressures from stakeholders
- How information is interpreted and used by different parties.
- By the existence of disinformation, etc.

Still, when taking the above reservations into account, the fact remains that the soundness of the documentation that constitutes the basis for decisions, or the quality-at-entry, has proved beyond doubt to be of vital importance for the outcome of investments.

9. Front-end estimation of cost and benefit

In project management, cost is the management parameter that attracts the most attention during the front-end phase and implementation. Some would argue that other parameters, such as project relevance deserve more attention. In many projects, even large cost overruns

have had little effect on long-term profitability. Yet in other cases, cost overruns may comprise a death blow.

Cost is eminently suitable as a management parameter, because it is expressed quantitatively with great precision and is continuously updated as a part of all transactions in a society. Costs are suited to making participants accountable, to gauging progress and result attainment and to comparing expenses with income to assess economic viability over time.

The prime focus is on cost overruns related to budgets. Major cost overruns can be serious, not least because they may trigger prolonged conflicts between the responsible parties on who shall pay the bills or how costs shall be divided. But the type of costs involved in budget overruns often is only top of an iceberg. In innumerable cases, the budget increase in the front end phase, from the first cost estimate to the adopted budget, is much greater. An interesting observation is that in projects in general, the initial cost estimate almost without exception is lower, not higher than what eventually is decides as the final budget.

In principle, there are four causes of cost overruns. They occur successively in the course of the front phase and the implementation of a project:

- Initially, planners and decision makers wilfully estimate costs low to increase the chances of a project being considered.
- The information base and the cost estimation methods are unsatisfactory.
- Unforeseen situations necessitate changes, for instance regulations imposed by the public that will increase cost.
- Inadequate cost management when the project is implemented.

Of the four, the first often has the greatest effect in terms of increased estimates. In many cases, the reason is deliberate underestimation to gain consideration. The principal point is obvious: get on the agenda, because the longer a project has been in the budget process and the further it has been studied, the greater the chances that it will be approved and implemented. Hence, underbidding price in the first round can be decisive. Moreover, even very large budget increases in the front end phase seldom have consequences for the responsible parties. Of course, it's the final cost estimate that's applicable. So, what's the problem? Evidence pretexts including *'we only wanted to start the discussion'* or *'a better estimate wasn't possible because we lacked information'*. Decision makers are surprisingly tolerant of what gets by early on, in spite of it arguably being the most decisive part of the entire project process. The same is true of the cost estimates of projects that have passed the first enquiry and are on the agenda. It has become so commonplace that one no longer speaks of systematic underestimation, but rather on normalization of deviation (Pinto 2006). In other words, a culture has evolved with lax views of honesty and compliance, to the extent that decision makers no longer see reason to trust the figures put forth in the front end phase. Hence, the possibilities of controlling and influencing are going down the drain.

This is serious. It means that poor projects slip through, though they should have been rejected had a realistic estimate put forward up front. Needless to say, this is a far greater problem than marginal budget overruns in the implementation phases of projects.

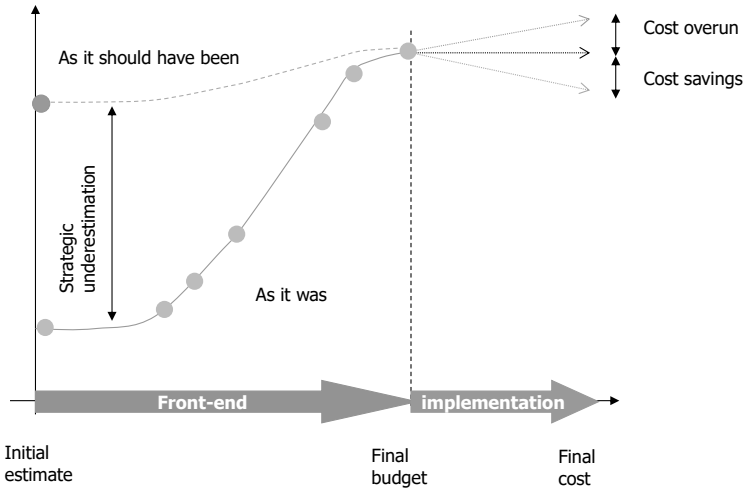


Figure 2.4 Early underestimation relative to what is the finally approved budget often is far greater than the cost overrun. Improving cost estimation in the front-end phase could result in fewer poor projects being chosen.

Systematic underestimation appears to be greatest in public projects, particularly so in local projects put forth for national financing. Hence, the phenomenon has become known as *strategic underestimation*. The principle of it is shown in Figure 2.4. The dots indicate cost estimates in the front end phase. The plot often ends up in some sort of S shape. Cost estimates are low in the initial period before the first systematic estimates of costs are undertaken. With time, the information basis improves, and the first surprises come to light. In turn, that triggers greater focus on the effort, demands for greater openness and realistic estimates, often by independent appraisals, and the cost estimate rises rapidly to the level at which it should have been at the outset. Thereafter, there are minor modifications until the final budget is approved.

The dashed line uppermost illustrates the development of cost in the front end phase as it should have been had the process started with an estimate at a realistic level. The difference between the dashed and solid lines is called strategic underestimation. In many cases, this is called tactical budgeting, which is a misunderstanding since what is at stake here is the choice of the project concept, which is a strategic choice.

The development of cost in the implementation phase is indicated by two dots at the upper right, designating cost overrun or cost savings. Strategic underestimation, as it is used here, often is large and in many cases many times the cost overrun. Cost overrun in relation to budget is typically in the range 10 – 100 per cent. The final budget is often several times as high as the first estimate, in some cases as high as 10 – 20 times.

A disproportionate amount of research has focused on the problem of cost overruns in projects. The difference between budget and final cost is erroneously designated by some as strategic misrepresentation (Flyvbjerg et al. 2003). In light of the discussion above, this might be called tactical cost estimation. Here we may distinguish between two phenomena. Strategic underestimation in the front end phase influences the actual choice of project. Improving cost

estimation in the front end phase conceivably is far more important than gaining control of cost overruns in implementation, as it may lead to fewer poor projects being chosen and thereby to increasing the overall benefit of investments.

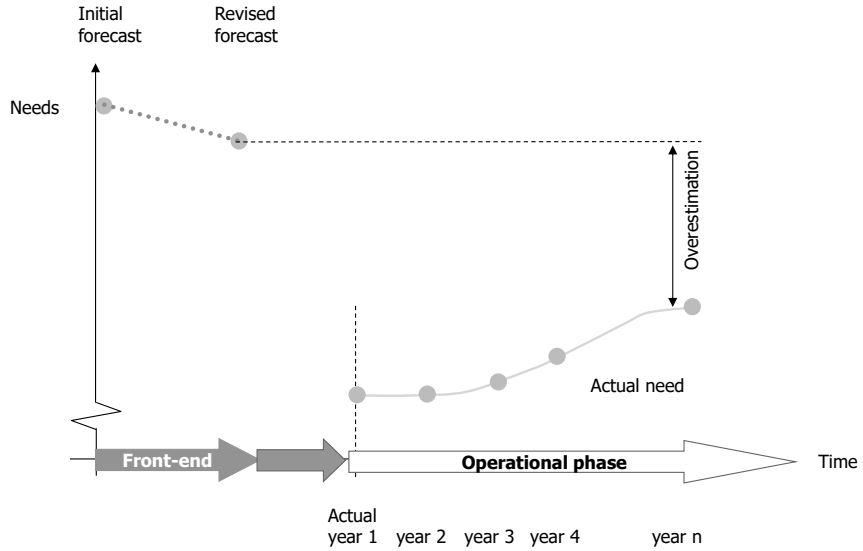


Figure 2.5 Strategic overestimation of benefits. The moment of truth is some time in the operational phase. The actual demand often deviates considerably from that assumed.

Adding to this picture, much the same problem as can be seen in cost estimation also applies in estimation of the anticipated utility or benefits of projects, as illustrated in Figure 2.5. Strategic overestimation of utility is common, much for the same reasons as mentioned earlier.

It's reasonable to assume that project utility often is more difficult to foresee than cost. The final effect of a project may be assessed only at some time after it has been handed over, and often many aspects difficult to predict affect user and market responses. In the front end phase, utility is estimated on the bases of parameters such as traffic volume, turnover, market response and the like. In some cases, the estimates are revised in the front end phase as more information is acquired. The moment of truth arrives when the project has been implemented and user response is evident. Initial response often is much lower than forecast response. Thereafter, response goes up and perhaps flattens out during the first few years, indicating an S-curve. The gap between the actual response curve and the prognosis amounts to what is here called strategic overestimation of utility.

The combined result of overestimated benefits and underestimated costs when expressed in terms of a benefit/cost-ratio could obviously be exceedingly misguiding for decision makers and their possibility to decide on a sound choice of project concept. In numerous infrastructure projects in the USA and Great Britain, the actual benefit/cost-ratio turned out to be 15% to 25% of that assumed at the time funding was approved (Flyvbjerg et.al. 2003). This

implies that the benefit-cost-ratio was prospectively overestimated by a factor of four to seven. That said, it is worthwhile to give heed to the possibility that going backward in time in each of these projects, to the earliest cost estimates and the earliest prognoses on which utility assessments were based, would reveal far greater exaggerations of benefit/cost-ratio or economic viability.

10. Financing mechanisms for projects

A necessary part of every project proposal is a plan for how to meet the project's financial obligations during every period of its life cycle. Firstly, the project needs capital as an input to production and to finance the high and negative cash flow during the construction phase. Secondly, the subsequent payback throughout the operational phase is not necessarily in financial terms. For example, a public road is expected to generate benefits in terms of time savings and improved safety conditions, which are benefits that are not easily sold in a market. This is illustrated in Figure 2.6, in which the dashed line is the *private* (financial) benefit and the solid line is the *social* benefit of the investment.

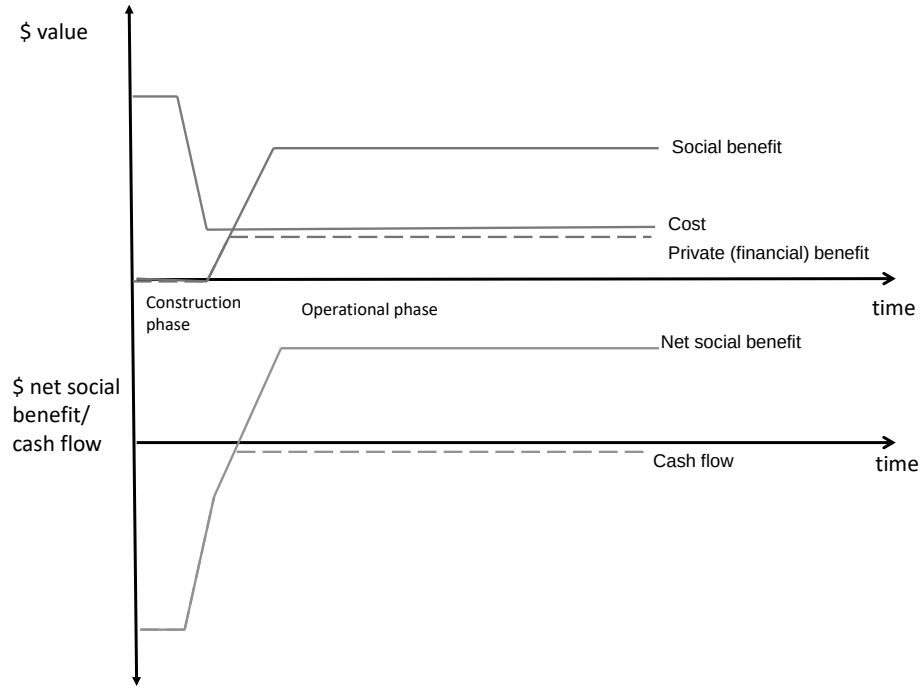


Figure 2.6 Costs and benefits over the project's life cycle

The Asian Development Bank (1997) suggests that the financial analysis focus on three questions:

1) Are adequate funds available to finance the project's expenditure? The cost of the funds needs to be taken into account.

2) Is it possible and desirable to seek recovery of some of the project's costs from the beneficiaries?
3) Is there an incentive to ensure a continued participation from the central stakeholders in the project?

In this section, the focus is primarily on public projects, particularly on infrastructure projects. State government funding is often the principal source of funds to meet investment and operating expenditures. Assuming that these funds mainly come from *extra taxes*, the marginal cost of taxation should be estimated. Moreover, a choice must be made as to whether it is possible and desirable to let users bear some of the financial burden in the form of *user fees*. If the local community as a whole is the beneficiary, one option is to seek *co-funding from local governments* (which in turn will be paid by local taxes). Another funding option for public projects is to bring in *private capital* of some form. Arguments for the latter are that public budgets are tight, and that private capital may be considered more cost-efficient. These alternatives will be discussed in more detail below, and we will argue that they may have an impact on the project's overall performance, both in strategic and tactical terms.

11. General taxes versus user fees

The cost of public funds
The collection of taxes, whether by central or local government, creates an efficiency loss in the economy through the taxes' impact on relative prices, and on people's behaviour. In addition, there are administrative costs associated with tax collection systems, which is called 'the cost of public funds' or 'the shadow price of taxation' (Grønn, 2003; NOU 1997:27). Some taxes are more distortionary than others. For example, personal income tax normally induces a high efficiency loss because it leads to a decrease in production. Sales taxes are somewhat less distortionary, whereas property tax is as close as we can get to a 'lump sum tax', which exhibits virtually no effect on people's behaviour. Some taxes even *increase* economic efficiency. For instance, 'green taxes' are introduced to correct a market failure. It is possible to estimate the efficiency loss from the mix of all taxes in an economy, thereby estimating the marginal cost of increased taxation for financing new projects. In Norway, for instance, which is a country known for a rather high income tax level, a 'conservative' estimate of the cost of public funds is approximately 20% (Finansdepartementet, 2005; NOU 1997:27).

The deterring effect of user fees
The main alternative to general taxes is a 'tax' linked directly to the use of the good, i.e. a user fee. However, for this to be a realistic option it must be possible to identify and charge users, and to reject those who do not pay. Most infrastructure projects have some elements of a 'public good'.[1] Public goods are characterized as being: (1) *non-rival in consumption*, implying that once provided, the marginal cost of giving another consumer access to it is zero, and (2) *non-excludable in consumption*, implying that it is impossible or very expensive to prevent anyone from consuming it. Examples of pure public goods are defence infrastructure, lighthouses and to some degree, transport infrastructure.[2] It follows that public goods are associated with a 'free-rider problem'. Potential investors will not be able to obtain the necessary return on their investment, which is why free and uncoordinated markets are not

[1] See any textbook on the subject of economics, particular in relation to public finance, e.g. Rosen (1995).
[2] Classification of a good as excludable or not is not absolute, as it depends on the state of technology, on costs and on legal arrangements.

able to provide public goods. This is a paradox for cases in which the project is highly socially desirable. In such cases, public funding through general taxes is therefore normally the only viable option.

When goods are excludable, user charging schemes may be established. The basic principle from welfare economics is that users should pay the marginal cost of providing the good or service to them. In the case of 'private goods' such as health care and education in which marginal costs are substantial,[3] user fees may provide an efficient way to prevent overconsumption. A public infrastructure is typically characterized by high costs upfront and negligible marginal costs, which implies that the optimal fee is close to zero. Any attempt to set a fee that exceeds the marginal cost will have a distortionary effect similar to those of other taxes (NOU 1997:27). This is illustrated in Figure 2.7 below; the socially optimal allocation is one where the price is zero and all users are being served. The impact of a tax is twofold: 1) area 'a' illustrates the tax revenue, i.e. a transfer from users to the infrastructure owner, while (2) area 'b' represents the social loss in terms of foregone benefits for users who exit the market.[4]

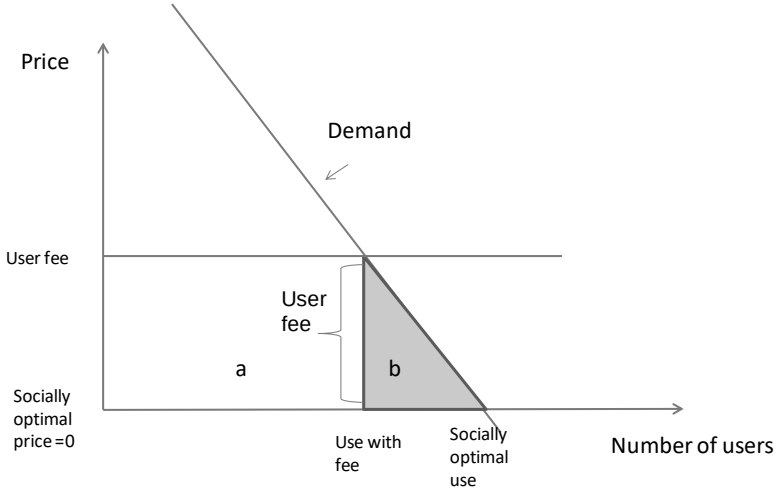


Figure 1.7 Deterring effects of user fees

Which are less inefficient, user's fees or general taxes?

Based on the above discussion, we realize that neither taxation nor user fees is a perfect choice as long as the marginal cost is zero. Below, we briefly discuss two situations in which user fees could be preferable (and vice versa):

- *Inelastic demand:* The size of the efficiency loss of a user fee depends on demand elasticity with respect to the fee. If demand elasticity is low (illustrated by a demand curve with a steep slope), users have few alternative options except to pay the fee and continue to use the infrastructure. In such cases, the efficiency loss will be low, as shown in Figure 2.8.

[3] The use of the labels 'public' and 'private' does not necessarily mean that it must be provided by those sectors, respectively.

[4] We have assumed a uniform user fee. If price discrimination is possible, less distortionary solutions may in theory be found (Rosen, 1995).

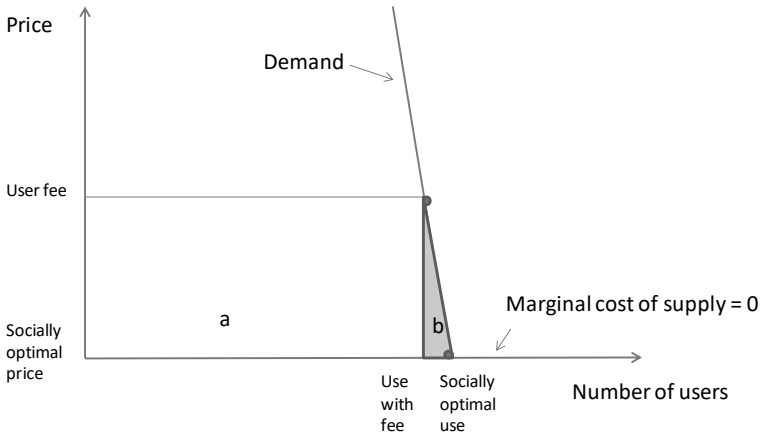
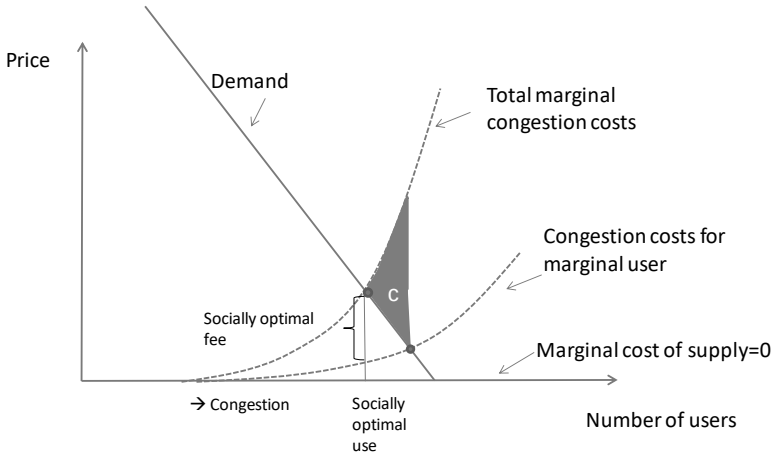


Figure 2.8 Inelastic demand

- *Negative external effects*: In some cases, using the infrastructure entails negative external effects not considered by users that can lead to overconsumption. For example, transport may have negative impacts on health, safety and the environment. Not least, transport systems impose substantial costs on society due to congestion (see for example Goodwin, 2004). The marginal congestion cost typically increases with the total number of users in the system, which is illustrated in Figure 2.9 below (based on Grønn, 2003). Area 'c' shows the social loss without pricing. Introducing a user fee during peak periods may yield a considerable social surplus, and according to NOU 1997:27, congestion taxes should *always* be implemented even before the need for increased capacity is considered.[5]



[5] However, as pointed out by, e.g. Parry and Bento (2001), this conclusion may be questioned in the presence of pre-existing distortions *outside* the transport sector.

Figure 2.9 Pricing of congestion

Administrative costs
User fee collection requires a collection system. Amdal et al (2007) studied the operational costs of toll road companies in Norway, and concluded that some projects are unsuited for private finance, as the operational costs comprise a too high proportion of revenues, up to as much as 40%. Welde (2011) found that operational costs are difficult to predict ex ante, and often turn out to be higher than estimated. On the other hand, general taxes are collected through existing collection systems and the marginal increase in their costs is negligible. While this has been an argument against toll roads hitherto, toll collection costs are expected to decrease when electronic collection systems are adopted.

The benchmark for user fees should be the general costs of taxation, which, and as explained above, have been estimated to be up to 20% of revenues in Norway. This means that if the costs of user fees (i.e. the sum of the efficiency loss and administrative costs) could be kept below 20%, then user fees can be a good solution.

12. What is 'fair' funding?

Society cares not only about efficiency, but also about how resources are distributed. This explains why governments often provide not only public goods, but also education and health care, and why tax systems are used for redistribution purposes.

Different financing mechanisms have different distributional effects. Conventional theory has suggested that user fees have a tendency to be regressive, i.e. to comprise a larger proportion of the income of the poor compared to the rich. Recent research on congestion charging in the transport sector has indicated that the opposite could also be true. In general, equity effects depend on the choice of charging scheme, including revenue recycling (Levinson, 2010). Furthermore, there are different views on what is 'fair' or 'reasonable' with regard to cost sharing. Is it reasonable that users pay? Or are infrastructure goods to be considered 'necessities' that everyone should have a right to consume regardless of their income? We will leave this discussion to the politicians. Nevertheless, we realize that it is important to consider these issues before making a choice of financing mechanism. The project's relevance and sustainability depends on it being accepted throughout society, including the sharing of the financial burden.

13. Local beneficiaries and perverse incentives

Infrastructure projects are often 'local public goods' and initiated by stakeholder groups in a local community. There could be many reasons why local infrastructure projects should be funded in part by state government budgets such as free rider problems, optimal risk sharing, distributional concerns and so forth. Different types of subsidies are available, e.g. central/local government cost sharing, 'soft loans' (an interest rate below the market rate), government guarantees, etc.

However, this induces a risk of a market failure known as the principal-agent problem in economics and contract theory.[6] This situation is characterized by: 1) a *conflict of interest* between the state government (the principal) and local agents, and 2) *asymmetric information*, in which the local agents know better about their real needs and their future behaviour. The state government is assumed to act on behalf of the entire nation, seeking the best projects within a national perspective. By contrast, local promoters only consider the benefits and costs accruing to them. A region with a low share of total taxes (for example, due to a small workforce) will regard a nationally funded project as being practically free of charge (Helland and Sorensen, 2007). Privileged *individuals* are particularly eager, e.g. a landowner who will benefit from increased land prices, a local politician who will increase his popularity among local voters and so on. Their arguments are typically formulated in terms of 'societal needs', although privileged groups have incentives to overestimate benefits, while underestimating costs and risk. It is not only local agents, but also planners in public organizations who may have a personal interest in the result. This leads to the 'survival of the unfittest', in which it is not the best projects that are built, but the most misrepresented ones (Flyvbjerg, 2007). Moreover, a project's success ex post often depends to a certain degree on local agents' efforts, although the state government cannot know whether these agents will make an effort once the funding is raised. These challenges are well known from development aid, and often increase within the layers of a hierarchy (see Ostrom et al., 2001).[7]

This problem arises because the state government does not have enough information to separate good project proposals from bad ones, and because agents are protected from the consequences of their own actions. One obvious solution is to therefore demand co-financing from local communities, the thinking with this being that agents will not promote bad projects if they have to bear the costs. Additionally, transparency and broad involvement in the pre-study phase is essential to ensure that local agents really do represent their entire community. An alternative explanation for the disproportionate distribution of national funding of local projects is the political parties' desire to maximize their numbers of seats in the national assembly (Helland and Sorensen, 2007). Parties will allocate more funds to districts with high voter mobility, and with many voters on the ideological cut-point. Furthermore, districts with a high ratio of parliamentary seats to voters will be favoured. Even so, in this case as well, local (co-)financing may be an appropriate measure. Local communities will not accept a useless 'gift' if it comes with an invoice.

14. 'Project-based funding' – with private capital?

The discussion thus far has primarily been in relation to the potential of a project in terms of a social surplus ('strategic potential'). Yet, another special challenge in public projects more related to tactical performance is the following: The lifetime of most infrastructure projects is several decades, with an implementation phase alone that amounts to five years or more. For an infrastructure administrator to develop and operate major projects in a time and cost-efficient way, a high degree of predictability and flexibility is essential.

In the real 'public budgeting' world, however, there are impediments to the optimal planning of major projects. In particular, government budgets are tight, and grants are only available on a year-to-year basis. The result of this is often a prolonged construction phase, less returns to

[6] See, e.g. Rosen (1995).

[7] A related challenge is the Samaritan's Dilemma, which was first identified by Buchanan (1975).

scale and a later realization of benefits. According to Vista Analyse (2010), the economic cost to society of this year-to-year principle amounts to 22-25% of construction costs. Mechanisms exist to overcome the problem by providing all the necessary funding at start-up, which is to be spent by the project administrator in a time-optimal way. These mechanisms are categorized as 'project-based funding', and normally involve the use of private capital (see Vista Analyse, 2010 and Econ Pöyry, 2008). Some examples of this are:

- *Multi-year budgeting* (politically binding but normally not legally);
- *State government loans* to public agencies, or allowing them a right to enter into *loan agreements in the financial market* (normally requires a change in organization, e.g. to a state-owned enterprise);
- Separating out *the projects as a legal unit*, with a right to take up loans;
- Leaving the development and operation of the infrastructure entirely in private hands (*privatization*);
- Creating a *partnership with a private contractor* who takes a major share of the responsibility for finance, development, as well as often maintenance and operation.

All of these models have been tested, and the results are mostly positive. Britain was a pioneer with Public-Private Partnerships for infrastructure projects such as roads and rail. One important experience is that such contracts work best when the entire project life cycle is taken into account in the contract (so-called life cycle models) and when the transfer of risk to the private party is 'real'.

As pointed out by the OECD (2008), there is an increasing gap between the demand for infrastructure and the available public finances in most OECD countries, due to ageing populations, an increasing health expenditure, etc. Hence, most countries have no other choice than to make more intensive use of private capital to help ensure the provision of infrastructure in the future. However, private investors will only consider their own payback and not the total social benefit, so in the end we are left with taxation and user fees as the ultimate (amortization) alternatives.

15. Conclusions

It all starts with the project proposal, and challenges in developing the proposal are abundant and complex. One is to avoid problems such as tactical budgeting, whereby responsible parties tend to underestimate costs in order to increase the chance to obtain funding for a project. Another challenge is to increase the chance that the most relevant project concept is identified. It is also crucial to ensure a transparent and democratic process and to avoid adverse effects of stakeholder's involvement and political bargaining. A major challenge is to make the process predictable, when the front-end phase will last for years. Many of the strategic performance problems facing investment projects can be interpreted in terms of deficiencies in the interaction between analysts and decision-makers in the front-end process.

Although we appreciate the rational decision model as an ideal, we are fully aware of the limitations facing planners and decision makers in real life: Time is limited, information is sparse, and stakeholder preferences vary and often conflict. But above all, we all live in a

political reality that isn't rational or even reasonable and only to a limited degree predictable. What can be achieved by rational analysis and planning is accordingly limited.

The bounded rationality model (Simon 1979) holds that problems and decisions should be reduced to a level at which they will be understood. In other words, the model suggests that we should interpret information and extract essential features and then make rational decisions within these boundaries. We can hope not for a perfect solution but for one that is "good enough" based on the limited abilities of the analysts to handle the complexity of the situation, ambiguity and information.

Then we must take into account whether or not the analysts' advice is applied by decision makers. In the ideal model for decision making, decision and analysis follow in a logical, chronological sequence that eventually leads to the selection and go-ahead of the preferred project without unforeseen interventions or conflicts. In reality, the process is complex, less structured, and affected by chance. Analysis may be biased or inadequate. Decisions may be affected more by stakeholders' priorities than by rational analysis. Priorities may change over time. Alliances and pressures from individuals or groups of stakeholders may change. Information may be interpreted and used differently by different parties. The possibility for disinformation is considerable, etc.

Under any circumstances, starting with a well formulated strategy may be an advantage, but is no guarantee for the best choice when the final decision is made. In some cases, the result may be entirely different from the initial choice. In other cases, the lengthy and unpredictable decision process may result in an optimal decision, even though the initial choice was entirely wrong.

A financial analysis is a crucial part of every project proposal, not least in public infrastructure projects. Different financing mechanisms can be applied under different circumstances and their features, strength and weaknesses vary.

Private capital is often expected to improve project flexibility and thereby improve the project's tactical performance. Ultimately however, a public project must be financed either by taxes or by user fees. Both create an efficiency loss in the economy as well as administrative costs. User fees will normally reduce demand ex post and thereby benefits realization. This option should therefore be considered only in cases of inelastic demand, negative external effects and/or highly efficient (electronic) collection systems.

Cost sharing between government and users/beneficiaries is also an issue. The risk of local agents' 'perverse incentives' could be an argument for local co-funding. The local share will ultimately be paid by local taxes or user fees.

The project's relevance and sustainability depends heavily on it being accepted throughout society. This includes the sharing of the financial burden. Therefore, the distributional impacts of different financing mechanisms should always be considered.

References

Amdal, Erik, Baardsen, Gunnar, Johansen, Kåre and Welde, Morten (2007). Operating costs in Norwegian toll companies: A panel data analysis, *Transportation* 34 (6): 681-695.

Asian Development Bank (1997). *Guidelines for the economic analysis of projects*, Economics and Development Resource Center, Manila.

Berg, Peder, et al. (1999), *Management of Public Investments*' (in Norwegian), Norwegian Ministry of Finance, Oslo.

Buchanan, James (1975). The Samaritan's dilemma. In Phelps, Edmund (ed.), *Altruism, morality and economic theory*, New York: Russel Sage foundation.

Econ Pöyry (2008). *Nye modeller for finansiering av jernbaneinfrastruktur*, report 2008-079, Econ Pöyry, Oslo.

Engwall, M. (2002). The futile dream of the perfect goal. In: Sahil-Andersson, K. and Soderholm, A. (eds). *Beyond Project Management: New Perspectives on the Temporary Permanent Dilemma*. Libe Ekonomi, Copenhagen Business School Press: Malmo, Sweden, pp. 261-277.

Goodwin, Phil (2004). *The economic costs of road traffic congestion*, University College London, London.

Grønn, Erik (2003). *Forelesninger i offentlig økonomi*, Cappelen Akademiske Forlag, Oslo.

Finansdepartementet (2005). *Veileder i samfunnsøkonomiske analyser*, Finansdepartementet, Oslo.

Flyvbjerg, Bent, Bruzelius, Nils and Rothengatter, Werner (2003), *Megaprojects and Risk: An Anatomy of Ambition*, Cambridge University Press, UK.

Flyvbjerg, Bent (2007). Policy and planning for large-infrastructure projects: Problems, causes, cures. *Environment and Planning B: Planning and Design*, Vol. 34 (4), 578-597.

Heijden, Kees van der. (1996), *Scenarios. The Art of Strategic Conversation*, JohnWiley & Son, UK, ISBN 0-471-96639-8

Helland, Lars and Sorensen, Rune (2009). Geographical redistribution with disproportional representation: A politico-economic model of Norwegian road projects, *Public Choice*, 139 (1), 5-19.

IMEC, Miller, R. and Lessard, D. (1999), *The Strategic Management of Large Engineering Projects*, MIT Press, US.

Levinson, David (2010). Equity Effects of Road Pricing: A Review, *Transport Reviews*, 30 (1), 33-57.

Linehan, C. and Kavanagh, D. (2004). From project ontologies to communities of virtue. Paper presented at the 2nd International Workshop, "Making Projects Critical".13-14th December 2004: University of Western England, UK.

Løwendahl, B. and Wenstøp, P. (2002), *Grunnbok i strategi*, NKS forlaget, Oslo.

Miller, R. and Hobbs, B. (2005). Governance regimes for large complex projects. *Project Management Journal* 36 (2): 42-50.

Minken, H., et al. (2009), *Konseptvalgsutredninger og samfunnsøkonomiske analyser* (Concept Appraisals and Economic Analysis), The Institute of Transport Economics (TØI), Oslo.

Mintzberg, Henry, Ahlstrand, Bruce and Lampel, Joseph (2005), *Strategy Bites Back: It is Far More, and Less, than You Ever Imagined…,* Prentice Hall, UK.

Morris, P.W.G., Hough, G.H. (1991), *The Anatomy of Major Projects. A Study of the Reality of Project Management*, JohnWiley & Sons, Chichester, UK, ISBN 0-471-91551-3

Naess, Petter (2004), *Bedre behovsanalyser. Erfaringer og anbefalinger om behovsanalyser I store offentlige investeringsprosjekt*. Concept rapport nr. 5, NTNU, Trondheim.

NOU 1997:27 *Nytte-kostnadsanalyser. Prinsipper for lønnsomhetsvurderinger i offentlig sektor,* Finansdepartementet, Oslo.

OECD (2008). *Infrastructure to 2030*. Policy Brief, OECD.

Olsson, N. (2006). *Management of flexibility in projects*. International Journal of Project Management 24: 66-74.

Ostrom, Elinor, Gibson, Clark, Shivakumar, Sujai and Andersson, Krister (2001). *Aid, Incentives and Sustainability: An Institutional Analysis of Development Cooperation*, SIDA, Stockholm.

Parry, Ian and Bento, Antonio (2001), Revenue Recycling and the Welfare Effects of Road Pricing. *Scandinavian Journal of Economics, 103 (4),* 645-671.

Pinto J.K., Slevin D.P., 1988, *Project Success - Definition and Measurement Techniques*, Project Management Journal, Vol. XIX No. 1, February 1988

Pinto, J. K. (2006), 'Organizational Governance and Project Success: Lessons from Boston's Big Dig', Presentation at the International Symposium on Project Governance, Norwegian University of Science and Technology, Trondheim.

Rosen, Harvey S. (1995). *Public Finance.* Fourth edition, McGraw-Hill, United States.

Samset K., (2008), Projects, Their Quality at Entry and Challenges in the Front-end Phase, in Williams, Samset and Sunnevaag (eds.), *Making Essential Choices with Scant Information*, Palgrave Macmillan

Samset Knut, (2010), *Early Proiect Appraisal Making the Initial Choices*, Palgrave Macmillan, UK.

Samset, K. (1998), *Project Management in a High-uncertain Situation: Uncertainty, Risk and Project Management in International Development Projects*, Norwegian University of Science and Technology, Trondheim.

Scheibehenne B. and von Helversen B., (2008), Useful Heuristics, in Williams, Samset and Sunnevaag (eds.), *Making Essential Choices with Scant Information*, Palgrave Macmillan

Simon, Herbert (1979), *Models of Thought*, Yale University Press, US.

Vista Analyse (2010). *Raskere og smartere samferdselsutbygging: Innspill til modeller for finansiering og gjennomføring av samferdselsprosjekter*, Næringslivets Hovedorganisasjon, Oslo.

Welde, Morten (2011). Accuracy of demand and operating cost forecasting for toll road projects. *Transport Policy*, 18 (5), 765-771.

Wideman, R.M., (2005), *Project Success – New Metrics and Measurements*, http://www.maxwideman.com/guests/metrics/metrics.htm

Williams, Terry, (2008), Decisions Made on Scant Information, in Williams, Samset and Sunnevaag (eds.), *Making Essential Choices with Scant Information*, Palgrave Macmillan

World Bank (1996), *Evaluation Results* 1994, The International Bank for Reconstruction and Development, Washington DC.

Paper IV

Chapter 18

Quality Assurance in Megaproject Management: The Norwegian Way

Gro Holst Volden and Knut Samset

Abstract

Governance regimes for major public investment projects comprise the processes and systems the financing party must implement to ensure a successful investment. Such regimes typically include a regulatory framework to ensure adequate quality at entry, compliance with agreed objectives, and sound management and resolution of issues that may arise during a project.

The challenges in securing quality at entry are considerable: i.e., identifying the right type of conceptual solution that is economically viable and relevant with respect to needs and often conflicting priorities in society beforehand; avoiding underestimating costs, overestimating utility and making unrealistic and inconsistent assumptions; and securing essential planning data and adequate contract regimes, among others.

This paper presents the Norwegian governance regime for public megaprojects and the lessons learned after 15 years. The regime involves external quality assurance of key decision documents at two gateways at the front end of projects: i) before the choice of conceptual solution and ii) before Parliamentary approval and funds appropriation. The results are promising. Most important, the regime has given the government greater control over the total cost of its investment project portfolio. Furthermore, the regime ensures that decisions regarding the choice of conceptual solution are made based on a broad assessment of overall needs and goals, as well as alternative ways of achieving these goals.

Keywords: project governance, public investment projects, quality assurance, cost estimation

Introduction

Public investment projects do not always meet the expectations of different stakeholders. Cost overruns are apparently the most common failure reported in the media. In studying more than 4000 large government funded projects, Morris and Hough (1991) found that cost overruns were typically between 40% and 200%. Flyvbjerg et al. (2003) analyzed 258 infrastructure projects in 20 countries over 70-year period and concluded that nine of ten projects had cost overruns. Further, Pinto (2006) claims that a culture has developed (in the US) whereby decision makers no longer see any reason to give credence to figures presented in the early phase of projects and instead acknowledge already at that stage that cost overruns will occur.

Another serious type of problem associated with projects is that they may not be able to produce the anticipated effect, rendering public resources wasted.

The three levels of project success, all of which are important, can be defined as follows (Samset, 2008):

1. Operational success: The project is delivered as promised with both time and cost efficiency.
2. Tactical success: The project produces the maximum utility/benefit for users at the lowest possible cost.
3. Strategic success: The project contributes to the desired societal development (as intended within its long-term objective), at the lowest possible cost and in a financially sustainable manner.

These three levels are in accordance with the three levels of achievements identified in the project management literature, i.e., (1) the outputs, (2) the outcome (first-order effects for users), and (3) the long-term effects for society.

In practice, the focus is primarily on operational success. However, as major public investments typically have a broader impact on society, an assessment of a project's tactical and strategic performance should be a vital aspect of the overall assessment of its success.

Public megaprojects are devised from needs that are politically expressed through dialogue between various stakeholders. The process of devising such projects typically involves government at various administrative levels, local government, political institutions, the public, the media, and contractors and consultants in the private sector. Such processes are often complex, disclosed and unpredictable, as described and analyzed in an in-depth study by Miller & Lessard (2000) of 60 major projects focused on the reconciliation of uncertainty and feasibility in the front-end phase. They can also be affected by deception and irresponsibility if stakeholders pursue hidden agendas rather than strive for openness and social responsibility, as discussed by Miller & Hobbs (2005) and Flyvbjerg et al. (2003).

In the field of project management, research has focused on the project itself and the improvement of the involved processes and procedures rather than on the governance framework that could or should provide direction and help improve the outcome of these processes. Project governance has become an issue of importance in the project management community only recently; see, for example, Müller (2009). Governance regimes for major investment projects comprise the processes and systems the financing party must implement to ensure a successful investment. Such a regime typically includes a regulatory framework to ensure adequate quality at entry, compliance with agreed objectives, and management and resolution of issues that may arise during a project, among others. It may also include an external quality review of key governance documents. However, the government, as represented by the responsible ministry, would have neither the necessary competence nor the need to interfere in the design and management of a project at the operational level. It would have a tactical and strategic perspective, and it should have a restricted role in facilitating structured, responsible and efficient preparation and implementation.

Flyvbjerg et al. (2003) discusses the ambitions, risks and effects related to megaprojects based on a large sample of projects. The authors conclude that the problems with such projects mainly concern negligence regarding risks and a lack of accountability among project promoters, whose primarily aim is to develop projects for private economic or political gain rather than for public benefit. To resolve the megaproject paradox, they suggest that: (1) risk and accountability should be much more centrally placed in decision making regarding megaprojects; (2) regulations should be in place to ensure that risk analysis and risk management are carried out; (3) the role of government should be shifted to reduce its involvement in promoting projects, maintain an arm's length distance, and restrict its involvement in forming and auditing the public interest objectives of megaprojects. Also that: (4) four basic instruments should be employed to ensure accountability in decision making by (a) ensuring transparency, (b) specifying performance requirements, (c) devising explicit rules to regulate project construction and operations, and finally (d) involving risk capital from private investors, where the

assumption is that private investors' willingness to invest would initially provide an effective indication of the viability of a project.

Front-end governance of megaprojects in Norway

In 1997, the Norwegian government initiated a study to review the systems for planning, implementing and monitoring large public investment projects because of a series of negative experiences with cost overruns, delays and low project success in general. The study reviewed eleven project cases in the transport, defense and construction sectors and focused on (1) whether the documentation that provided the basis for decisions was adequate when the project was approved and (2) whether project implementation was satisfactory. The study (Berg et al., 1999) found that of the eleven projects, only three were completed within the original budget; cost overruns for the other eight were as high as 84%. Moreover, the underlying documentation was deficient in a number of projects. The study ultimately concluded that failures in the initial phase of projects prior to the decision to proceed were generally the main cause of the low success rate for projects.

Further, a challenge with public investment projects in Norway has been that planning processes are often sectorial and locally based. The front-end phase has typically been a bottom-up process where ideas are generated locally by those who benefit from the project, and there may be strong incentives to overestimate utility and underestimate costs. Such incentives, referred to as perverse incentives in Samset et al. (2014), create a classic principal-agent problem. Broader economic analysis is typically conducted at a later stage when the conceptual solution has already been selected.

In 2000, the Norwegian Ministry of Finance introduced a mandatory quality-at-entry regime to address the challenges described above. Initially, the aim of the regime was to reduce the problems with cost overruns, i.e., to ensure operational success. In 2005, the regime was expanded to include quality assurance of the choice of conceptual solution in order to ensure that the right projects are started and that unviable projects are rejected, i.e., to improve tactical and strategic success. The regime was designed to improve analysis and decision making in the front-end phase, particularly the interplay between analysis and decision-making. It was based on the notion that the necessary rules for decision-making were already in place; however, there were no binding rules to ensure quality and consistency of analysis and decisions.[i]

To accommodate the needs and practices of different ministries and agencies, the scheme was devised to avoid being strict and overly comprehensive. Thus, it establishes requirements for the type of documentation that must exist, but it does not require that public agencies use specific tools, formats, and so forth and does not interfere with project implementation as such. Rather, the aim of the scheme is (1) to ensure political control with fundamental go/no-go decisions, (2) to ensure an adequate basis for decisions, where the focus is on essential matters rather than on the details, and (3) to establish a distinct set of milestones and decision gates that would apply to investment projects in all sectors. To fulfil these aims, the scheme was devised (1) to anchor the most essential decisions in the Cabinet itself, (2) to introduce a system for quality assurance that is independent of government and sufficiently competent to overrule analysts, and (3) to ensure that the governance regime is compatible with the procedures and practices of the affected ministries and agencies.

Under the Norwegian Quality-at-Entry regime, highly skilled external consultants, pre-qualified by the Ministry of Finance, are now assigned to perform quality assurance of the decision documents in all public investment projects with a total budget exceeding NOK 750 million (approximately EUR 80 million). During the first four years, such quality assurance was performed in some 50 projects, where cost estimates and decision documents were scrutinized prior to funds appropriation by Parliament. The involved consultants had expertise within project management and cost engineering. Based on this initial experience, the regime was then expanded, and it now involves two separate quality assurance processes to ensure the basis for decision making regarding (1) the choice of conceptual solution (QA1) and (2) the budget, management structure, and contract strategy for the chosen project

alternative (QA2). For QA1, broader competence, including competence in social science and economics is required.

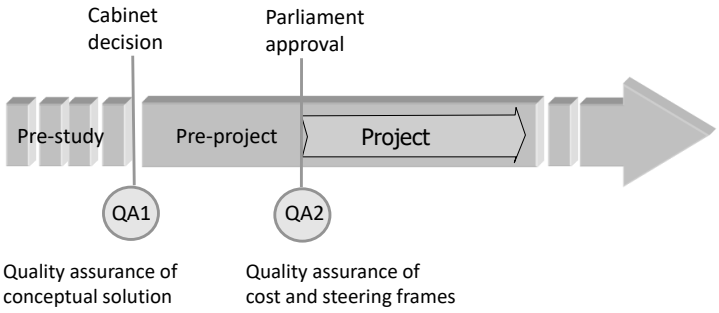


Figure 1. The Norwegian Quality-at-Entry regime for major public investment projects

The review process is fairly similar in both QA1 and QA2. The reviewer first receives documentation from the sectorial ministry and its subordinate agency and then examines the documentation to check whether it provides a sufficient basis for decision-making. If the documentation is insufficient, additional information may be requested. The reviewers also conducts independent analyses and calculations (an uncertainty analysis and, in QA1, a cost-benefit analysis). Finally, the reviewer writes a report and presents it to the sectorial ministry and the Ministry of Finance. The report is generally then made public.

QAl is a qualifying step for QA2, and QA2 is a qualifying step for submission to the budget process. Having QAl performed on a project does not guarantee that QA2 will be performed, and having QA2 performed on a project does not guarantee that the project will be prioritized by Parliament. The QAl process begins with a decision in the sectorial ministry responsible for the project and the participation of the Ministry of Finance as a quality body. When the external reviewer submits his report, the case is evaluated by the Ministry of Finance and presented to the Cabinet, which then decide whether to proceed with the project. The sectoral ministry may also decide to stop the process after QAl. If a decision is made to continue with a pre-project, the resulting document will be subjected to external review (QA2). After the QA2 report is delivered, the Cabinet still has two options – either stop the project or allow it to enter into the budget process, but without any guarantee that it will be prioritized (Christensen, 2011).

The data

After 15 years of operation (2015), nearly 200 projects have been subjected to QA2 reviews, and about 80 of these are completed and in the operational phase. The QA1 scheme has been in operation for ten years, and about 70 projects have so far been through a Conceptual Appraisal (CA) followed be an external QA1 review; however, none of these projects have been finalized thus far. To date, only eleven of the projects have undergone both QA1 and QA2.

With few exceptions, the projects subjected to QA1 and QA2 represent major public investments with an expected investment cost above the threshold value of approximately EUR 80 million. For most of the projects, the cost estimates range from EUR 70-300 million; however, some of them have a much higher cost estimate. For example, the acquisition of new fighter aircrafts is estimated to cost about EUR 7 billion.

About half of the projects fall under the purview of the Ministry of Transport (mainly road and rail); the other half are mostly projects under the Ministry of Defense, and construction and ICT projects in different parts of government. The agency for the construction of public buildings (Statsbygg) under the Ministry of Local Government and Modernization is also involved, as is the Ministry of Finance in its role as the manager of the QA scheme.

Since its inception, the involved parties have gained considerable experience with the QA scheme. A trailing research program funded by the Ministry of Finance continuously collects data on the projects and the QA process and thus contributes to the learning process. However, information on the effects of the scheme is only now becoming available, as it takes time to plan and implement large investment projects. Indeed, the process to conduct CA and QA1 reviews before the Cabinet's decision typically takes 2-4 years. Most of the projects then enter the pre-project phase, which ends with a QA2 review and Parliament's approval, and if approved they proceed to an engineering and construction phase, which typically takes 2-5 years, in exceptional cases up to 10 years. Only after this point can the final cost, time, and quality of delivery – in other words, indicators of the project's operational success – be registered. In order to observe the long-term effects for users and society, one has to wait even longer, until the project has been in operation for some years.

The present study is based on information from the QA1 and QA2 reports, CA reports and other project documents that provided the basis for the external quality reviews, parliamentary propositions including proposed cost and steering frames, as well as final reports and evaluations from ministries and agencies and other documents regarding project delivery and effects.

Data were collected between 2010 and 2015. Agencies and ministries responsible for the projects have provided support in data collection and quality control of figures. Approved cost frames and steering frames were collected from parliamentary propositions and resolutions. For the completed projects, the final costs were based on project accounts produced by the responsible agencies and reported to their respective ministries. These are public figures that are presented to the Parliament and signed off by the Auditor General each year. There may be some weaknesses regarding the quality of these cost data, as explained further below. However, it should be noted that the researchers have been through a quite thorough process of verification including reviews of accounts and interviews with responsible officers, to clarify issues, get details about price indexes used, etc. Aass (2013) and Welde (2014) provide a more detailed account of the data and the researchers' analyses to ensure validity of information. It should also be noted that the projects we are dealing with here are the country's very large ones, which receive special attention from both agencies, ministries and politicians, with strict requirements for analysis up front and external quality assurance. The projects are therefore well documented and monitored as compared with other projects outside the QA scheme[ii].

We reserve the right of inaccuracies related to the following errors:

- Final cost for completed projects could be somewhat incomplete. So far there is no evidence that agencies were trying to "hide" project costs, but the researchers' possibilities to verify if all real project costs were included in the accounts were of course limited.
- In some cases, cost figures may only be available long after the project is finalized due to ongoing judicial conflicts with suppliers regarding compensation for changes and additions.
- Agencies' reporting of the final costs may be incomplete. The agencies have varying practices for registering the final cost. This study includes all projects for which we had access to the final cost, which might create a bias towards agencies that provide researchers with timely and accurate data, such as the Norwegian Public Roads Administration. The proportion of road projects in the study is 58%, whereas that in the overall population of quality assured road projects is 48%. We regard this discrepancy as acceptable, and it should be noted that it is only a matter of delay, not of missing data as such. One should however also be aware that the relatively limited number and the predominance of road projects limit the accuracy and

strength of any statistical analysis in the study. The composition of projects in several different sectors also implies that conclusions are not representative of projects in individual sectors.

- Calculation of price adjustment and use of indexes may cause some inaccuracy. The agencies' methods and practices regarding price adjustment vary somewhat. Road projects are price regulated with the official index for road construction. In other projects, the agencies' own price adjustments are applied. The largest agencies have established their own cost indexes and procedures for adjusting the cost and steering frames on the basis of real cost development in their sectors. In the period covered in this study, all such indexes have increased above the Consumer Price Index (CPI).
- Correspondence between project scope and cost accounting may be somewhat divergent

In relative terms, however any discrepancies between the reported and the actual final cost are presumably small and constitute at most no more than a few percent of the total cost. For example, as the implementation phase in most projects is limited to a few years, the choice of price index is not expected to have a significant effect on the level of cost compliance or on differences between sectors.

These data on final costs include all 67 projects that had been completed at the time of study. With a few exceptions outlined below, these projects constitute the whole population of QA2 projects where the final cost had been established and reported. To our knowledge, 14 additional projects that underwent QA2 are completed, but their final costs have not yet been established. These projects are not discarded but only temporarily suspended, they will be included when cost figures are available. Rather than relying on forecasts for the final costs we prefer to wait until reliable data are available. Our study is updated annually based on a steadily increasing number of projects.

Only four projects have been excluded from the study. Three are PPT projects on account that their final cost will not be available due to a non-disclosure agreement between the concessionaries. One road project is not eligible since it underwent considerable changes in scope throughout its implementation so that meaningful comparisons between budget and cost border on the pointless.

To determine the effect of the QA1 scheme is another issue altogether. We would have to wait until investment projects were 3-5 years into the operational phase before their effects could be evaluated in a tactical and strategic perspective. The current situation (year 2015) is that none of the projects subjected to QA1 has yet been completed; thus, it will take time before the effect of this more fundamental and comprehensive intervention can be evaluated. So far, a total of 11 of the first QA2 projects have been evaluated ex post in terms of their effects on users and society, nine by external evaluators and two by our own researchers. This has provided us with valuable information, however, since they have not been through QA1, they will serve only as a control group for subsequent evaluations of the impact of QA1 projects.

Quality assurance of the choice of conceptual solution (QA1)

The current procedure during the QA1 phase is for the responsible Ministry to prepare a Conceptual Appraisal (CA) report or pre-feasibility study of the investment case. This report should include the following documents:

1. *Needs analysis*. In this document, all stakeholders and affected parties are identified, and the relevance of the anticipated investment in relation to their needs and priorities is assessed.
2. *Overall strategy*. Based on the prior analysis, consistent, realistic and verifiable immediate and long-term objectives are specified in this document.

3. *Overall requirements*. This document identifies all the requirements, such as functional, aesthetic, physical, operational and economic requirements, that need to be fulfilled.
4. *Possibilities study*. With the "opportunity space" delimited by the needs, objectives and requirements, this document provides the limits to what is possible and identifies realistic alternative conceptual solutions.
5. *Alternatives analysis*. At least two alternatives and the so-called zero alternative (no project) are analyzed to specify their operational objectives, essential uncertainties, cost estimates, and so forth, and the alternatives are subjected to a full cost-benefit analysis that is reported in this document.
6. *Guidelines for the pre-project phase*. This document includes a suggested implementation strategy for the preferred alternatives.

The CA report is now being scrutinized by external reviewers (QA1). Also, they perform a complete cost-benefit analysis of the alternatives based on guidelines from the Ministry of Finance. The reviewers present their findings in a report containing their assessment and advice regarding the following:

- Uncertainties likely to affect the project
- The anticipated economic benefits and costs of the concepts analyzed
- The ranking of the alternatives
- The management strategy

The purpose of QA1 is to assist the Ministry in ensuring that the decision regarding the choice of conceptual solution has been subjected to a fair and rational political process. Ultimately, of course, the concept is selected through a political process in which external reviewers play no role. The reviewers' role is limited to controlling the professional quality of the underlying documents that provide the basis for the decision. As a fundamental requirement, at least two viable alternative concepts in addition to the zero alternative should be reviewed.

The Ministry now analyses the documents and presents the case to the political level. As the project owner, the state must determine how to best solve the underlying problem that triggered the project and the associated societal needs. The Cabinet now makes the decision regarding the choice of conceptual solution and decides whether to proceed with the pre-project phase.

Lessons related to QA1

Ten years after the first QA1 report was produced, it is still too early to evaluate the effects of the scheme. Our knowledge is limited by the type of projects that have undergone QA1, the quality of CA and QA1 reports, and the resulting decisions. Indirectly, one can also infer some of the spinoff effects in government, industry and academia after the introduction of the scheme.

As of 2015, approximately 70 projects have undergone CA and QA1. There is little doubt that the quality of the CA reports has improved steadily over time and that there is a convergence towards a common best practice. The same trend can be observed with the QA1 reports – quality assurers have gained years of experience and shown a positive learning curve (Samset and Volden, 2013). Some reviews in the literature have already examined the performance of the CA/QA1 process in the transport sector and stakeholders' experience with the scheme; see, for example, Rasmussen et al. (2010), Statens vegvesen (2012), and Bjertnæs (2012). These studies suggest that the CA/QA1 process may consume time and resources, but overall, agencies seem to benefit from the scheme. In particular, the scheme provides a more systematic approach to the early identification of project ideas than the prior system. Rather than going straight to selecting road sections and determining a technical solution, planners are forced to take a broader perspective and to discuss societal aspects, which

allows ideas to mature and stimulates creativity in the agencies. The process in the scheme also increases the likelihood that the most effective option will be included in the analysis.

The QA1 scheme allows the ministries and government to have a more direct influence in the early stages of the process in comparison to local stakeholders, who have traditionally had a significant influence, especially in road projects. However, there is still room for improvement. One in-depth study of 17 projects (Samset et al., 2013) specifically examines how the opportunity space is defined and utilized in CA reports. A recurrent problem is that the conceptual solution has already been selected before the CA process, either because of path dependency in the agencies or political constraints and limitations. Another study (Statens vegvesen, 2012) suggests that quality assurers seem to give disproportionate attention to economic considerations and that they should balance economic impacts with the achievement of various political objectives. Finally, some ministries and agencies have drawn attention to the futility of undergoing the full CA/QA1 process in cases where, in their opinion, there are simply no alternatives apart from one feasible conceptual solution.

In studying the CA and QA1 recommendations and the resulting decisions for the first 70 QA1 projects, trailing researchers have found that quality assurers agree with the sectoral ministry on the ranking of concepts in one-third of the cases. In the remaining two-thirds of the cases, the quality assurer and the sectoral ministry disagree on the ranking of concepts. The QA1 reports more often recommend the zero alternative or a more economically feasible concept. In the QA1 report, the quality assurer often criticizes the sectoral ministry for its failure to explore the entire opportunity space, particularly with respect to less expensive concepts, during the CA process. Most of the 70 projects have now been through political treatment by the Cabinet, and almost 80% of them have entered into the pre-project phase with one (or sometimes more than one) concept. In only 6% of the cases, the Cabinet has rejected the project altogether, normally in accordance with the QA1 recommendations. Not surprisingly, we observe that when the quality assurer approves the CA recommendation in the QA1 report, the Cabinet normally follows the recommendation. The QA1 process thus increases the confidence that the proposed concept is the most efficient and effective alternative. However, when the recommendations diverge, the outcome is less predictable. In such cases, the Cabinet follows the recommendation by the sectoral ministry more often than the QA1 recommendation, but in some cases, project proposals are withdrawn, sent back to the sectoral ministry for new CA appraisal, or the Cabinet chooses a completely different concept (Grindvoll, 2015). The results indicate that the Cabinet is now more informed about the consequences of projects, mostly the economic consequences, and that they take QA1 recommendations seriously; however, the choice of a project concept clearly remains a political decision.

Notably, there is no tradition in Norway of prioritizing public investment projects according to their anticipated economic viability, and this is particularly the case with road projects (Welde et al., 2013). This lack of prioritization according to viability applies to both politicians and government agencies. The QA1 scheme can ensure only that decision makers are well informed about both alternatives and their economic implications. Over time however, it may become more difficult to select conceptual solutions that are obviously ineffective and that are clearly inferior to other alternatives. Evidence also suggests that an independent QA1 report showing that an investment is poorly justified may be influential and thus essential for the government to make a sound decision in a controversial case. One example is the proposal by sports enthusiasts and local communities to hold the Winter Olympics in Norway in 2018 and later in 2022; the proposal was for the state to guarantee the investment cost. In both cases, the Cabinet rejected the proposal, though only after the QA1 report indicated that the benefits were overestimated and that the costs were underestimated.

One noticeable impact of the CA/QA1 scheme is that the ministries' opportunity space has been broadened during the appraisal process owing to the advice from quality assurers, as they play a role as not only controllers but also advisers. In several cases, the CA documents were rejected by the external reviewers, resulting in a second round in the appraisal process, with new conceptual alternatives. In addition, there is reason to believe – although it is difficult to prove – that many of the most poorly conceived investment proposals are now screened out before they even reach the

CA/QA1 stage. Such proposals can be rejected early because of the improved processes and procedures in the involved ministries and agencies, which likely constitutes the most important beneficial effect of the QA1 scheme.

Quality assurance of the chosen project and its budget (QA2)

QA2 is performed at the end of the pre-project phase, and the aim is to provide the responsible ministry with an independent review of its strategic management document before it is submitted to Parliament for approval and funds appropriation. This review constitutes not only a final control measure to ensure that the budget is realistic and reasonable but also a forward-looking exercise to identify managerial challenges ahead. The analysis should help substantiate the final decision regarding the project's funding, and it should be useful during implementation as a reference for control.

As inputs to the quality assurance review, the agency is obliged to provide the following:

(1) The overall management document (steering document),
(2) A complete base estimate for costs and, if relevant, income/revenue, and
(3) An assessment of at least two alternative contract strategies.

The quality assurer reviews and verifies these documents and conducts a separate analysis of success factors/pitfalls and the overall uncertainty scenario. The cost uncertainty analysis should be based on the base estimate, and it should stipulate the expected additions in order to establish the expected costs and associated uncertainties. The quality assurer gives recommendations regarding (1) the proposed cost frame, including necessary contingency reserves and the agency's steering frame, and (2) the appropriate strategy for managing the project in order to keep within the cost frame, including the management and authorization of contingency reserves.

The proposed cost frame is established by using standard procedures for stochastic (probability-based) cost estimation, based on either mathematical-analytical methods or simulation tools. The result is a cumulative probability distribution of the investment cost, as illustrated below. The proposed cost frame is normally P85 with deductions for possible simplifications and reductions (the so-called reduction list) that can be implemented during the project if there were a danger that the cost frame would be exceeded. The agency's steering frame is lower, normally at the P50 level, in order to avoid incentives to use contingency reserves; see figure 2.

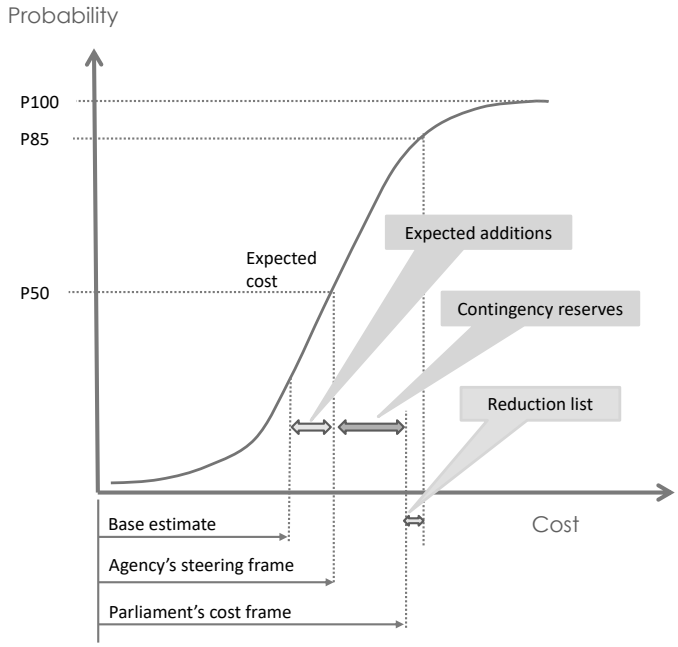


Figure 2 Stochastic cost estimation. Definition of key terms.

Parliament and the responsible ministry are naturally not required to follow these recommendations. The final overall cost frame for the project is decided by Parliament, and the ministry will then determine the steering frame for the executing agency. In professional terms, the QA2 review heavily relies on project management expertise, i.e., how to ensure that the project outputs are delivered on time, with the agreed quality and within the cost frames. Contract strategy is an important part of the exercise, as are elements of economics, including incentive theory, transaction cost theory, and organization theory more generally. Quality assurers are expected to have expertise in all these areas.

Lessons related to QA2

Although QA2 recommendations regarding cost and steering frames are only advisory, we find that in about 70% of the projects, the approved cost frame is identical to the QA2 recommendation, and in the remaining 30% of the projects, there are only minor deviations. Further, the final steering frame is identical to the QA2 recommendation in 54% of the projects, and with two exceptions, the deviations are within +/- 10% (Samset and Volden, 2013). These results indicate that the QA2 process and the stochastic cost estimation techniques used during this process are trusted as a basis for determining the budgets of major public investment projects.

The first analysis of cost compliance was presented in 2013 based on all QA2 projects completed by the end of 2012, with an established final investment cost (Samset and Volden, 2013 and 2014). In all, there were 40 such projects at that time, and the results were promising[iii]. As many as 32 of the 40 projects, i.e., 80%, were completed within or below the cost frame. Moreover, some of the projects had significant savings, totaling about EUR 500 million (mostly road projects). The total net savings for the projects taken as a whole were more than EUR 300 million or about 7% of the total investment. This is an exceptionally good result in comparison with what one could expect based on past experience and findings from a number of studies in other countries.

A total of 78 projects submitted to QA2 review have now been completed (2015), and the final cost has been established for 67 of these (84%). The results from the update of the previous analysis where these projects are included (Welde, 2015), are presented below.

Final cost relative to the approved cost frame

Figure 3 shows the difference between the final cost and the cost frame approved by Parliament, where the latter largely corresponds to the P85 estimate. The data show that 53 of the 67 projects, i.e., 79%, were completed within or below the cost frame. The total net savings for the projects taken as a whole was almost EUR 600 million, or about 7% of the total investment.

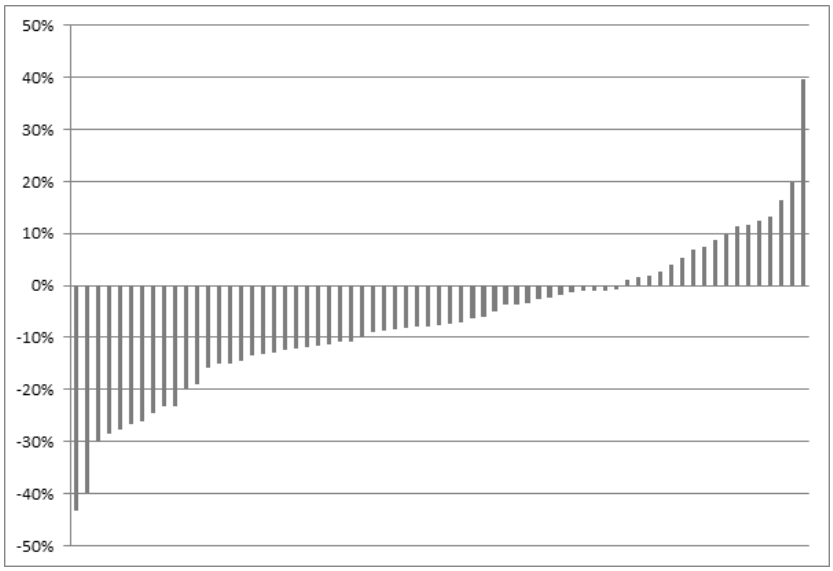


Figure 3. Deviation between the final cost and the cost frame approved by Parliament (N=67)

Cost deviation by sector

Figure 4 shows the extent to which projects in different sectors complied with their approved cost frames. The number of projects is still too small to draw any firm conclusions, but the defense sector notably has had no projects with cost overruns. The railroad sector also performed above average, with only 14% of the projects with cost overruns.

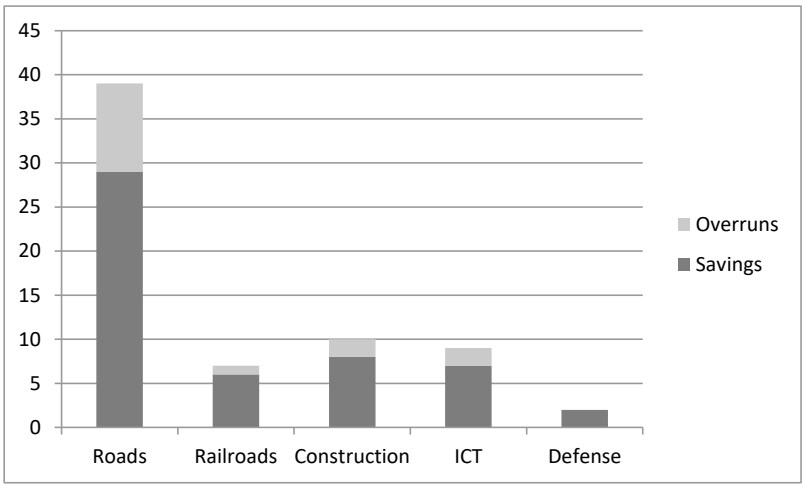


Figure 4. Number of projects with cost overruns and savings by sector (N=67)

Cost deviation by date

Another factor that may influence the problem of cost compliance is the date on which the project was commissioned. Indeed, the outcome of a project may be positively affected by learning effects, or positively or negative affected by economic cycles, in the sense that price changes may be greater than the variation captured by the uncertainty analysis.

Figure 5 shows the difference between the final cost and the approved cost frame for the projects, which are now sorted by time of inception, commissioned from 2000 to 2012. We find a vague tendency for cost overruns to have occurred in the middle part of the period, i.e., 2004-2008. On average, the projects in this period had final costs that were relatively equal to their cost frames (0% deviation), whereas the mean deviation for projects commissioned before and after this period was 13%, i.e., considerable cost savings. The difference between the two means is statistically significant at the 0.05 level. This vague tendency may be due to strong cost increases in the construction industry that occurred in this period. Alternatively, the subsequent Global Financial Crisis of 2007-2008 may have had unforeseen consequences. Beyond these explanations, it is difficult to determine any cause of the vague tendency observed.

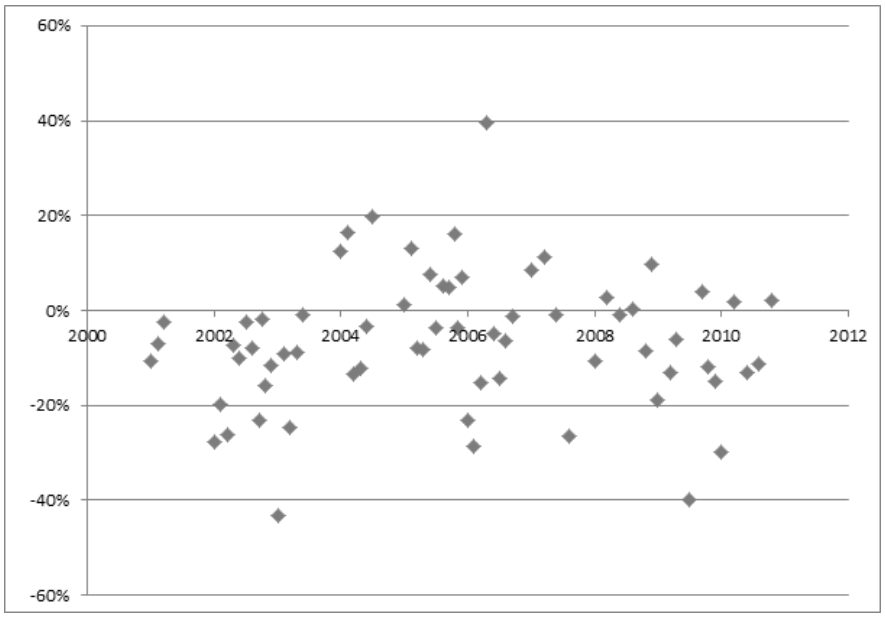


Figure 5. Deviations between the final cost and the approved cost frame at the time of commissioning for the project (N=67)

Final cost in relation to the agency's steering frame

As mentioned above, the steering frame for the executing agency largely coincides with the estimated median (P50). Given the uncertainty associated with project implementation, one must not only expect deviations but also accept them. With steering frames at P50, we should expect equal numbers of overruns and underruns, and with a sufficiently large portfolio, the average for the whole portfolio should be close to the median.

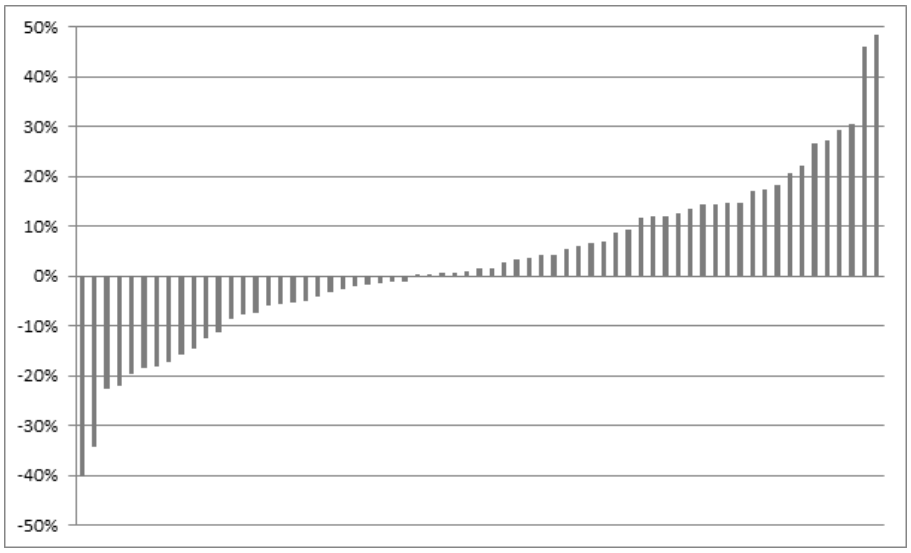


Figure 6. Deviation between the final cost and the agreed steering frame for the project (N=65)

The differences between the final cost and the steering frames are illustrated in figure 6, and the results are as expected. The differences are almost symmetrically distributed around the median, indicating that cost control at the portfolio level is good. The distribution is slightly skewed to the right, however, with 48% of the projects below and 52% above the steering frame. There is an average positive deviation of 2.8%. Ideally, the deviation should be zero.

As mentioned, researchers are now in the process of performing more extensive ex-post evaluations of projects that have undergone QA2 and are now in their operational phase where benefits for users and society may be measured. The results for the first 11 projects evaluated indicate that, overall, their rate of operational success is high, i.e., not only in financial terms. Only two of eleven projects exceeded their cost frames, whereas five exceeded their steering frames. Furthermore, only two experienced delays, and four had (insignificant) shortcomings related to quality and functionality. The projects were essentially well organized and executed. Most risk factors that do indeed materialize were identified in the QA2 reports. However, notably, in one of the projects, expensive adjustments and upgrading were necessary in the first few years after it began. This finding shows the importance of focusing on the life cycle cost, not exclusively the investment cost.

The results from the first projects that have undergone QA2 show that the majority have been completed within the cost frame. Further, the deviation between the final cost and the steering frame is almost symmetrically distributed around the median. These results indicate that at the portfolio level, the Norwegian state is now effectively controlling costs in major investment projects. Moreover, as the deviations are both positive and negative to almost the same extent, there is no incentive to spend contingency reserves when unnecessary. This result is likely due to the practice of establishing a lower steering frame for the executing agency (typically at the P50 level).

The data reveal a tendency for overruns to have occurred in the middle part of the period (i.e., projects started in 2004-2008). Monitoring how the results develop over time and in different sectors is thus important. The first set of completed projects may be somewhat biased in the sense that these projects were implemented in a time-efficient manner. Road projects are somewhat prevalent among the first completed projects, and many of them were rather uncomplicated. By contrast, a number of defense and ICT projects that started many years ago remain unfinished, and we are still waiting for their results.

Discussion

Considering the above results, QA2 seems to have helped improve cost management and ensure cost control in major public projects in Norway. The situation in the 1990s was that major cost overruns were the norm rather than the exception, in both Norway and other countries. International research (Flybjerg et al, 2003) has shown that the situation has neither improved nor worsened over the past 70 years. However, caution should be made when comparing the results to the situation in the 1990s and in other countries. What QA2 ensures, is that projects are now more mature when they are approved by Parliament, that the cost estimate is realistic, and that it includes necessary contingency reserves (corresponding to P85 level minus a reduction list). An important part of the picture is that the projects are normally approved (and the cost frame set) at a later stage in the planning process today than in the 1990s. Previously, the cost frame was often based on earlier and less realistic estimates and did not always include provisions for uncertainty.

Generally speaking, the encouraging results concerning cost compliance could be interpreted in two ways, either that investments projects that have undergone QA2 are now implemented at a lower cost than before, or that the cost frame is established later based on more mature and therefore more realistic estimates. While projects in the 1990s experienced large overruns in relation to their budget, we see today that projects may have overruns compared to "very early estimates» (Welde et al., 2014). What QA2 cannot prevent is an increase in the project content, scope and cost during the process that

precedes QA2. Before the cost frame is set, projects have often undergone a long front-end phase where user groups and stakeholders have great expectations. When the project is presented to Parliament, it may be too late to reject the proposal even in cases where the project scope has grown beyond what is considered an efficient solution to the societal needs. It is at the QA1 stage where the cost estimate is compared with expected benefits to determine whether the project is worth implementing. However, if the cost estimate increases between QA1 and QA2, the assessment of the project's tactical and strategic success from the QA1 stage will no longer be valid. Only now can we observe the first projects that have been subjected to both QA1 and QA2, and the researchers examining the effects of the QA scheme will expand their focus accordingly. If an increase in scope between QA1 and QA2 is a prevalent problem, a solution could be to change the QA scheme in the future so that the cost estimate at QA1 is binding for the subsequent pre-project phase (the so-called "Design to cost" principle).

The issue of selection bias always needs to be investigated. A relevant question is whether the magnitudes of overruns have changed not only between the large projects but overall across all projects. In such a situation, it may be difficult to conclude that the implementation of quality assurance for large projects alone was the major cause of the differences in the magnitudes of cost overruns among large projects. Odeck et al. (2015) tested for this issue of selection bias in a similar study. They found that while cost overruns in small road projects had remained constant, large road projects under the QA2 scheme had experienced significant reduction in the magnitudes of overruns during the same period.[iv]

The Norwegian Quality-at-Entry regime is essentially a top-down regulatory scheme that was introduced to enforce a qualitative change in government practice and to improve quality at entry for major investments. It does not interfere with current procedures in the involved ministries and agencies but merely aims to improve existing documents that provide an essential basis for political decision making. Naturally, the regime was initially somewhat controversial, as it entailed the involvement of external experts in matters handled by the ministries and agencies, as well as an additional cost and time delay. In reality, however, it has essentially received constructive responses from the ministries and agencies, which have adapted their practices to meet the new quality requirements and, in some cases, also adopted the scheme as a self-regulatory procedure.

A recent study (Kvalheim et al, 2015) examines whether the Norwegian government is better informed when making decisions about major public investment projects today compared with the period before the QA scheme was introduced. Based on a sample of projects from before and after the QA scheme was introduced, as well as interviews with 24 public planners and leaders, the study shows that

1. the quality of the decision documents has improved considerably;
2. the essential factors for the choice of conceptual solution are covered in the CA/QA1 analyses;
3. the premises for making a decision are well documented;
4. recommendations are clear and transparent; and
5. the projects are sufficiently mature and ready for Parliamentary approval after QA2.

A significant feature of the Norwegian QA scheme is the spin-off effects that it seems to have had on both the government and private sector. In the period after the scheme was introduced, we find a clear trend of improved practices in the areas of cost estimation and budgeting, risk assessment and strategic planning. Moreover, there is growing awareness in the government regarding the need to improve the quality of decision documents, broaden the scope of analyses to include alternative concepts, and avoid making overly detailed analyses at an early stage. Such awareness has also proliferated into the consulting and construction industries, which have clearly responded positively to the new procedures and requirements in these areas in their role as suppliers to the public. We can also see that front-end management has become an issue within the community of professionals in project management, and training courses are now being offered by a number of institutions and

consultants. Improved practices have also been adopted and institutionalized by various government agencies. Sectors not subjected to the QA scheme, notably health authorities (Myrbostad et al., 2010), electric utilities (St. Meld. Nr 14 (2011-2012)), and the Oslo municipal authority (Oslo Kommune 2011), have voluntarily introduced variants of the scheme. Other countries have also shown interest in the scheme. For example, in Sweden, a variant of the CA report, directly inspired by the Norwegian scheme, was introduced in 2013 as a new step in early planning. The Province of Quebec, Canada, has also introduced a similar scheme.

As mentioned, one fundamental aspect of the governance regime is that at least three alternative conceptual solutions should be analyzed, at an early stage when options are still open. The alternatives should be analyzed to the same degree of specification to ensure that the assessment of the alternatives is fair. This requirement has triggered a debate regarding what defines a conceptual solution. For instance, should conceptual solutions constitute different technical solutions to the same problem (e.g., a bridge versus a tunnel in an infrastructure project for crossing a fjord), or should conceptual solutions be determined based on differences in the combined effects of different projects in the broadest sense? Whatever the answer, as the regime has put this issue on the public agenda, it has had a notable effect on analysts, politicians and the public. This issue, and the emphasis on economic analysis, might prove significant in the aim to identify relevant alternative concepts and select the most viable project alternative.

Conclusion

In this paper, we have discussed the necessity of governance regimes in securing the interests of the financing party in public investment projects, i.e., in improving the overall decision making and success of public projects, from both an operational and a tactical and strategic perspective. We have presented and discussed the Norwegian Quality-at-Entry regime, which requires two external quality assurance interventions before project implementation. In this top-down regulatory regime, major decisions are anchored at the political rather than the administrative level of government. The approach of limiting interference in existing practices and procedures inherent in the scheme seems to be effective, and it may facilitate the development of self-regulatory schemes, which may ultimately make central top-down interventions unnecessary. Openness and transparency seem to be essential in improving the governance of public investments.

Lessons so far indicate that projects subjected to budgetary quality assurance (QA2) are now largely completed within their cost frames. Hence, at the portfolio level, the state is able to more effectively control the cost of major investment projects. Perhaps the most intriguing aspect of this study, as compared to other studies of cost compliance in projects, is that our figures are consistent, based on the same principle and methodology for all projects, and therefore comparable. The good results concerning cost compliance should however not be confused with evidence that Norway implements its major projects more cost-efficiently than other countries. What QA2 cannot prevent is an increase in the project content, scope and cost during the process that *precedes* QA2. It should be a topic for future research to investigate whether this is a prevalent problem.

It is too early to determine whether quality assurance of the choice of conceptual solution (QA1) has increased the tactical and strategic success of projects, but it is clear that the systematic appraisal of the choice of conceptual solution is considered useful.

The Norwegian QA regime is a novel and simple approach to improving project governance. With a few exceptions, most notably the lack of requirement for risk capital from those who initiate and benefit from the project, the scheme is in line with Flyvbjerg et al.'s (2003) recommendations. It ensures that the projects are mature and well-defined when presented to Parliament, it requires the use of stochastic cost-estimation methodology, and it contains incentives for the agency not to use the whole cost frame. So far the scheme seems to work well in a Norwegian context. However, this does

not imply that it will be preferable in any context. Other countries with different institutional contexts, specific types of projects, and major, long-lasting projects may require other approaches.

Project governance has become an issue in the project management community only recently. To move forward in this field, numerous questions need to be answered. What procedures are applied in different countries and agencies, and what are their effects? What would it take to develop more effective governance regimes at the international, government or corporate level in order to ensure maximum utility and return on investment for society and investors? What is the optimal mix of regulations, economic means and information in improved governance regimes for major investment projects? A challenge for the project management community will be to shift the focus beyond the delivery of the project itself to the broader issues of the project's utility and effects. A recent study compared governance schemes for major public investment projects in Norway and five other countries (Samset et al., 2015). There are many similarities between the schemes, but also differences regarding for example the number of project phases, decision points and interventions, and who performs quality assurance. All schemes are of a fairly recent date, and it is too early to explore their effects and compare their success, but this would be an interesting topic for future studies. An increased understanding and sensitivity in this area could be of mutual benefit to both the financing and the implementing parties.

References

BEMELMANS-VIDEC M., RIST R., and VEDUNG E. (eds.) (1998). *Carrots, Sticks, and Sermons: Policy Instruments and Their Evaluation*. Transaction Publishers, New Brunswick, New Jersey. 1998

BERG, P., ØSTBY, L.-E., LILLEBY, S., STYRVOLD, S., HOLAND, K., KORSNES, U., RØNNING, K. and JOHANSEN, F. (1999). *Styring av statlige investeringer. Sluttrapport fra styringsgruppen*, Ministry of Finance.

BJERTNÆS, A. M. (2012). *Konseptvalgutredninger – intensjon og forbedringer*, Master thesis, NTNU

CHRISTENSEN, T. (2011). "The Norwegian front-end governance regime of major public projects: A theoretically based analysis and evaluation", *International Journal of Managing Projects in Business*, Vol. 4 No. 2, 2011 p. 221

FINANSDEPARTEMENTET (Ministry of Finance) (2013). *Reglement for økonomistyring i staten Bestemmelser om økonomistyring i staten*. Oslo: Finansdepartementet

FLYVBJERG, B., BRUZELIUS, N. and ROTHENGATTER, W. (2003). *Megaprojects and Risk: An Anatomy of ambition*, Cambridge University Press, Cambridge

GRINDVOLL, I.L. (2015). *Hva har skjedd med KS1-prosjektene? – Status per mars 2015*, Concept working paper

KVALHEIM, E.V., CHRISTENSEN, T., SAMSET. K. and VOLDEN, G.H. (2015). *Har regjeringen fått et bedre beslutningsunderlag? Om effekten av å innføre konseptvalgutredning (KVU) og ekstern kvalitetssikring (KS1 og KS2) for store statlige investeringsprosjekter*, Concept working paper

MILLER, R., and HOBBS, B. (2005). *Governance regimes for Large Complex Projects*. University of Montreal. Paper for the EURAM 2005 conference.

MILLER, R., and LESSARD, D.R. (2000). *The strategic Management of Large Engineering Projects: Shaping institutions, Risk and Governance.* Cambridge Mass.: MIT Press.

MORRIS, P.W.G. and HOUGH, G.H. (1991). *The Anatomy of Major Projects. A Study of the Reality of Project Management*. Chichester, UK; John Wiley & sons.

MÜLLER, R. (2009). *Project Governance*. UK, Gower Publishing

MYRBOSTAD, A., ROHDE, T., MARTINUSSEN, P. and LAUVSNES, M. (2010). "Regime for planlegging og beslutning i sykehusprosjekter", Concept report No. 25

ODECK, J., WELDE, M., VOLDEN. G.H. (2015). The Impact of External Quality Assurance of Costs Estimates on Cost Overruns: Empirical Evidence from the Norwegian Road Sector. *European Journal of Transport and Infrastructure Research*, 15 (3), pp. 286-303.

OSLO KOMMUNE (2011). *Konseptvalgutredning (KVU) i Oslo kommune. Krav og veiledning*

PINTO J.K., SLEVIN, D.P. (2006). *Organizational Governance and Project Success: Lessons from Boston's Big Dig*, Concept Symposium 2006, NTNU, Trondheim

RASMUSSEN, I., HELDAL, N., HOMLEID, T., IBENHOLT, K., SKJELVIK, J. M. and VENNEMO, H. (2010). "På vei til kvalitet? Evaluering av KS1 i transportsektoren", *Vista Analyse report* 2010/10

SAMSET, K. (2008). *Prosjekt i tidligfasen: Valg av konsept*, Tapir Akademisk forlag

SAMSET, K. and VOLDEN, G.H. (2013). «Investing for Impact. Lessons with the Norvegian State Project Model and the first investment projects that have been subjected to external quality assurance", *Concept report* No. 36, Ex Ante Academic Publisher, Norway

SAMSET, K., ANDERSEN, B. and AUSTENG, K. (2013). «Mulighetsrommet. En studie om konseptutredninger og konseptvalg", *Concept report* No. 34, Ex Ante Academic Publisher, Norway

SAMSET, K., VOLDEN, G.H., WELDE, M. and BULL-BERG, H. (2014). "Mot sin hensikt. Perverse incentiver – om offentlige investeringsprosjekter som ikke forplikter", *Concept report* No. 40, Ex Ante Academic Publisher, Norway

SAMSET, K. and VOLDEN, G.H. (2014): "Front end Governance of Major Public Projects - Lessons with a Norwegian Quality Assurance Scheme", International Journal of Architecture, Engineering and Construction, 2014; Volume 3(2)

SAMSET, K., G.H. VOLDEN, N. OLSSON and E.V. KVALHEIM (2015): Styringsregimer for store offentlige prosjekter. En sammenliknende studie av prinsipper og praksis i seks land. (Governance schemes for major public investment projects: A comparative study of principles and practices in six countries). *Concept report* No. 46, Ex Ante Academic Publisher, Norway

STATENS VEGVESEN (The Norwegian Public Roads Administration) (2012). *Evaluering av KVU/KS1. Færre og bedre KVU-er*

ST.MELD. Nr.14 (2011-2012). *Vi bygger Norge – om utbygging av strømnettet*

WELDE, M., ELIASSON, J., ODECK, J. and BÖRJESSON, M. (2013). "Planprosesser, beregningsverktøy og bruk av nytte-kostnadsanalyser i vegsektor. En sammenligning av praksis av Norge og Sverige", *Concept report* No. 33, Ex Ante Academic Publisher, Norway

WELDE, M. (2014). *Oppdaterte sluttkostnader – prosjekter som har vært underlagt KS2 per mai 2014.* Concept working paper

WELDE, M., SAMSET, K., ANDERSEN, B., AUSTENG, K. (2014). Lav prising – store valg. En studie av underestimering av kostnader i prosjekters tidligfase. (Low estimates – high stakes. A study of underestimation of costs in projects' earliest phase). *Concept report* No. 39, Ex Ante Academic Publisher, Norway

WELDE, M. (2015). *Oppdaterte sluttkostnader. Prosjekter som har vært underlagt KS2 per september 2015*, Concept working paper

WORLD BANK (2000). Reforming Public Institutions and Strengthening Governance, A World Bank Strategy, Washington DC, November 2000

AASS, T. (2013). *Kostnadskontroll i prosjekter som har vært underlagt ekstern kvalitetssikring KS2*. Concept working paper

Notes

[i] A parallel here would be private financial institutions, where investment projects are assessed almost exclusively based on a review of the investors' credibility and collateral available, with little regard to the substantive issues related to or characteristics of the investment project as such.

[ii] All agencies are obligated to keep accounts according to the Regulations for State Financial Management (Finansdepartementet, 2013). The accounts provide a basis for control over the allocation of appropriations and the basis for analysis of activities, including final costs in individual projects. Final costs cannot be tampered with without compromising the accounts. The Auditor General audits agency accounts annually.

The accrued expenses in the projects are collected annually through the agencies accounting systems and followed up as part of the agencies' internal control systems. During the construction phase the agencies reports accrued expenses to the responsible ministry and provides forecasts for final costs. Upon completion, total annual costs are adjusted to a common price level using the appropriate indexes and used internally as part of the agencies internal evaluation systems.

The data are owned by the responsible agencies, but are publicly available for research and other purposes.

[iv] We do not have access to data on final costs on smaller projects from other agencies than the Norwegian Public Roads Administration, so we cannot conclude with certainty that quality assurance alone is the main reason for changes in the magnitudes of overruns, but the study by Odeck et al. (2015) concluded that quality assurance has led to a reduction in cost overruns among road projects which make up a majority of the projects in our study.

Paper V

PAPERS

Governance of Major Public Investment Projects: Principles and Practices in Six Countries

Gro Holst Volden, *Norwegian University of Science and Technology, Trondheim, Norway*
Knut Samset, *Norwegian University of Science and Technology, Trondheim, Norway*

ABSTRACT

This article compares the Norwegian scheme for quality assurance of major public projects with similar project governance schemes in five other OECD countries.[1] All schemes have been introduced since the turn of the millennium and seem to be fairly consistent with recommendations from the project management literature. There are also a number of differences between the six schemes, for example, with regard to parties and roles, comprehensiveness, flexibility, organization, and whether portfolio management is covered. It is too early to make conclusions about their relative effects, but the evidence thus far indicates that there is much to learn across countries.

KEYWORDS: project governance; governance framework; public projects; front-end

Project Management Journal, Vol. 48, No. 3, 90–108

Published online at www.pmi.org/PMJ

INTRODUCTION

Public investment projects amount to large sums, both in relative terms and absolute figures. The McKinsey Global Institute (2013) estimates global infrastructure spending to be at 4% of the total global gross domestic product, mainly delivered as large-scale projects. However, public investment projects face a number of challenges and have varying reputations. There is broad literature on what Hall (1981) termed "great planning disasters," which are projects with cost overruns, time delays, and either no benefits or very limited benefits, and that are sometimes so controversial and infeasible that they end up being closed down or severely altered. The problem of cost overrun is particularly well documented (Morris & Hough, 1987; Flyvbjerg, Skamris Holm, & Buhl, 2003a; van Wee, 2007). For example, Flyvbjerg et al. (2003a) analyzed 258 infrastructure projects in 20 countries over a period of 70 years, and concluded that the cost overruns were significant and the situation had not improved during the period. The more serious, but equally common, problem is when projects do not meet the expectations of users and society. For example, Pinto (2006, p. 7) quotes from an *Infoworld* article describing, "a U.S. Army study of IT projects [that] found that 47% were delivered to the customer but not used; 29% were paid for but not delivered; 19% were abandoned or reworked; 3% were used with minor changes; and only 2% were used as delivered." Similarly, Flyvbjerg, Bruzelius, and Rothengatter (2003b) showed that benefit shortfalls are a consistent problem in the transport sector.

These problems are not limited to the public sector—see, for example, Merrow (2011), who documents similar challenges in the private sector. The public sector, however, has some additional challenges, including multiple objectives, difficulties in measuring success, and having to deal with a wide array of external stakeholders in the democratic decision-making processes (Klakegg & Volden, 2016). Public projects are the outcome of a political tug-of-war between stakeholders in society, whose needs and priorities will concur or conflict to varying degrees. The outcomes of such processes are not always predictable. This is clearly shown in Miller and Lessard's study of 60 international projects (Miller & Lessard, 2000). Some authors emphasize dishonesty and "strategic explanations" as the causes of project failure, including deliberate misrepresentation in project appraisal by promoters (Flyvbjerg et al., 2003b), which is referred to as "perverse incentives" by

[1]This article is based on a research project funded by the Concept research program and retrieved from www.ntnu.no/concept. Preliminary results were presented in Samset, Volden, Olsson, and Kvalheim (2016).

Volden and Samset (2015). However, the public sector, too, has some internal challenges, such as a weak capacity for designing a strategic vision, lack of skills, and lack of coordination among levels and actors, as noted by the Organisation for Economic Co-operation and Development (OECD) (2015b).

The very largest of such projects are the most crucial: they are "too big to fail," they are very expensive, and they have high levels of inherent uncertainty and risk (Le Quesne & Parr, 2016). Special measures are therefore required to ensure successful implementation and outcome. In order to deal with these challenges, some governments have established designated governance schemes for the very largest projects. Norway was a pioneer in this endeavor and introduced an overarching framework for the governance of major public projects in the year 2000. See, for example, Volden and Samset (2017) for a presentation of the Norwegian framework and its effects, some of which are very encouraging; other countries have introduced similar frameworks in recent years. In this article we provide a description and a comparative analysis of how such project governance schemes are currently being organized and handled at the central government level in six countries: Norway, the United Kingdom, Denmark, the Netherlands, Canada (Quebec Province), and Sweden. Common to all schemes is that they are intended for project governance by a central government and applied to projects that involve particularly high costs, risk, and complexity, or are highly innovative. For example, in Norway, there are 20 to 30 such projects annually.

Our contribution to the literature is the compilation of a set of innovative project governance schemes, in which we highlight their differences and similarities and present the preliminary evidence of their impact. The results should not only be of academic interest, but should also provide information for other countries considering the introduction of similar mechanisms for improving the success of major public projects, including the OECD's ongoing work to establish a common framework for governance and delivery of infrastructure (OECD, 2015b), which seems to have focused more on delivery models and less on the strategic project perspective. All schemes are relatively recent, however; therefore, it is too early to determine with certainty their impact and degree of success, and this should be a topic of future studies.

This article starts with definitions of key concepts and principles related to project governance and presents key findings from the literature, while highlighting the importance of the front-end phase and role of central government. Each country's governance scheme and its underlying stage-gate models are described, as well as the involved parties and their roles, the use of independent quality assurance in the process, and a number of other elements. Similarities and differences between the schemes are explored to discuss the significance of principles and practices of the different approaches to project governance.

Governance of Public Sector Investment Projects

Governance

In general terms, *governance* relates to "all of processes of governing, whether undertaken by a government, market or network, whether over a family, tribe, formal or informal organization or territory and whether through the laws, norms, power or language" (Bevir, 2013, p. 1). The term governance means "to steer." In political science, it refers to what happens at the government level in a society. It concerns the role of government in facilitating the attainment of societal objectives. The government generally has three types of policy instruments at its disposal: the stick, the carrot, and the sermon, corresponding to regulation, economic means, and information (Bemelmans-Videc, Rist, & Vedung, 1998). The instruments may be either affirmative or negative. The model has its parallel in the regime of the World Bank (World Bank, 2000), in which the regulation element is described in terms of rules and restrictions, the economic element in terms of competition pressure, and the information element in the forms of transparency and assistance.

Governance is often used as a normative concept, whereby the quality of governance is compared to a standard of "good governance." For example, the United Nations Development Programme (UNDP) (2006) defines good governance as "among other things *participatory, transparent*, and *accountable*. It is also *effective* and *equitable*. And it promotes *the rule of law*" (our italics). Similarly, the Council of Europe (2014) suggests 12 principles for good governance, including *sustainability* (long-term orientation) and *competence* and *capacity*. Regardless, the social and economic consequences of poor governance policies and systems may be considerable.

A related term is *corporate governance*, which refers to the mechanisms, processes, and relations by which corporations are controlled and directed. Müller (2009) distinguishes between the traditional "shareholder perspective," which limits corporate governance to a question of how to incentivize management to deliver good financial results, and the "stakeholder perspective," which is broader and takes a wide range of other stakeholders into account. According to the OECD (2015a), good corporate governance involves a set of relationships between the organization's management, its board, its shareholders, and other stakeholders. Moreover, good corporate governance requires a structure defining how the organization's goals should be determined, how such goals should be realized, and how this should be followed up (OECD, 2015a).

Project Governance: Principles and Components

The term *project governance* has only recently become an important issue in the project management community and literature. It refers to the processes,

systems, and regulations that the financing party must have in place to ensure that projects are successful (i.e., that relevant and sustainable project alternatives are chosen and delivered efficiently) (Volden & Samset, 2017). The Project Management Institute (PMI) (2013) defines project governance in a similar way, as "an oversight function that is aligned with the organization's governance model and that encompasses the project life-cycle [by providing] a comprehensive, consistent method of controlling the project and ensuring its success by defining and documenting and communicating reliable, repeatable project practices." A key project governance issue is that the interests of the implementing agent will not necessarily be aligned with those of the financing party or project owner. Project governance seeks to ensure that an implementing agent, in this case represented by the project manager, will act in conformity with the interests of the owners (Tirole, 2001). Project governance is thus a system of appropriate checks and balances that enables transparency, accountability, and defined roles, while at the same time supporting project managers in delivering their objectives. This corresponds well with what Morris and Geraldi (2011) define as the institutional level of managing projects, which focuses on shaping the context and conditions to support and foster projects, although Morris and Geraldi focus more on the support function than the governance function. As noted by Crawford et al. (2008), there is a possible conflict of interest facing a project sponsor (owner), between the "governance perspective" and the "support perspective." On one hand, the sponsor should have an external focus, representing the enterprise and the client's interest, and on the other hand, he or she must have an internal focus, providing project management with support to fulfill their role efficiently. Crawford et al. find that the sponsor role is played out quite differently in different organizations. In our study, the focus is primarily on the governance perspective.

Various definitions and typologies of project governance are suggested in the literature. Williams, Klakegg, Magnussen, and Glasspool (2010) distinguish between governance *of* projects, which aims at efficient delivery, and governance *through* projects, which aims at choosing the right concepts and ensuring that effects are realized and are sustainable. Müller, Shao, and Pemsel (2015) distinguish between *project governance* and *governance of projects*, where the former refers to the governance of a single project, and the latter to the governance of groups of projects, such as a program or portfolio.[2] In a similar manner, Too and Weaver (2014) note that publications discussing project governance can be classified into two main groups. The first group focuses on governance of single projects, typically involving several actors and stakeholders, when a contract will specify the specific governance arrangements for that project. The second group of publications examines governance models linking different project-related levels (project, program, and portfolio) within an organization, and thus sees project governance as a subset of corporate governance. In our study, the focus is on governance schemes applying to all major investment projects at the national level. Accordingly, our perspective is the *governance of projects* in Müller's terminology, but we take the central government perspective rather than a given organization's perspective. A natural implication is that we emphasize governance *through* projects somewhat more than the governance *of* projects (cf. Williams et al., 2010).

[2]Project governance includes, among many other things, the governance part of the project management methodology, the role of the steering group, and the sovereignty and authority with which the project manager can manage his or her project. By contrast, governance of projects includes issues such as the level of institutionalization of project management, for example by using similar reporting systems, methodologies, or project selection techniques across the group of projects (Müller et al., 2015).

Flyvbjerg et al. (2003b) discusses which criteria should underpin megaproject governance regimes. Based on a large set of empirical data, they found that the main problem with major public projects is that the stakeholders have a self-interest in their implementation (whether financial or political); they underestimate the risk and they are not held accountable to central government, which adopts a more overarching perspective of maximizing public benefits. The authors talk about the "megaproject paradox," and propose the following alleviating measures:

1. Risk and accountability must be accorded much more of a key role in decision-making processes.
2. Risk analysis and risk management requirements must be imposed.
3. The authorities should remain at "arm's length" and not become involved in promoting the project, but limit their role to formulating overarching objectives and ensuring that such objectives are attended to by the project.
4. In order to bring about responsible decisions, one should:
   - ensure transparency;
   - specify performance requirements;
   - impose clear requirements for the construction and operation of the project; and
   - involve capital from private investors since their willingness to invest will be a project viability test.

Haanes, Holte, and Larsen (2006) reviewed different models for decision making in major public projects based on best practice in Norway and other countries and suggest the following minimum requirements:

- Clearly defined project phases
- Clearly defined decision points
- Quality assured basis for the decisions
- Simplicity
- Some degree of standardization and common terminology

Likewise, Narayanan and DeFillippi (2012) state that project governance schemes should incorporate five elements: stage-gate approval processes, formal roles and responsibilities, stakeholder representation, quality assurance, and contracts and sign-offs.

More recently, several standards and guides have been developed to address project governance models, in particular as part of corporate governance. For example, the Association for Project Management (APM, 2011) has established 13 principles for the governance of project management, and has defined four main components of schemes that adhere to them:

1. Portfolio management—ensuring that each project is aligned with key business objectives
2. Project sponsorship—providing a link between the permanent and the temporary organization, typically by defining a project sponsor or project board as the "governance agent," with decision making, directing, and representational accountabilities
3. Project management capability—ensuring that the teams responsible for projects are capable of achieving the objectives that are defined at project approval points, which is a question of skills, available tools and processes, and a clear mandate (among others)
4. Disclosure and reporting—ensuring that project reports provide timely, relevant, and reliable information that supports the organization's decision-making processes, without fostering a culture of micro-management

Such principles and guides may be more or less detailed, and more or less mandatory. Some project governance models are behavior oriented, requiring that certain detailed rules are followed (e.g., common project management methodology), whereas others are outcome oriented and give more autonomy to the project manager. These two "paradigms" may also be denoted as bottom-up and top-down (Müller, 2009). The top-down model is more often found in organizations with a high level of trust and a high level of project management skills.

Some organizations have established project management offices (PMOs). A PMO is an internal group or department that defines and maintains standards for project management; provides training, monitoring, and reporting on active projects and portfolios; and, in some cases, takes a more strategic role, acting as the owner of the project portfolio. PMOs may take many forms, as demonstrated by Hobbs and Aubry (2008), but they often have a central role in a project governance model (Morris & Geraldi, 2011; Müller et al., 2014).

In this article, we focus more on the structural than the non-structural elements of project governance. However, it should be noted that project governance is not only about laws and regulations, as it is not possible to determine every action. Based on Foucault's work, Müller et al. (2014) introduced the term *governmentality* in the project management literature. Governmentality is a combination of "governance" and "mentality," and addresses the human side of governing—the attitude that governors have toward those they govern, and whether governance is enforced through strict rules or through soft "cultural" values that members of an organization share and respect. Similarly, Klakegg and Meistad (2014) divide governance into structure-based and relationships-based governance. The former incorporates the elements mentioned above, such as stage-gate approval processes, roles and responsibilities, and quality assurance; whereas relationships-based governance typically includes non-hierarchical elements such as leadership, motivation and incentives, resource allocation, trust and ethics, alliances and involvement of stakeholders, informal relations, and communication.

According to Miller and Hobbs (2005), large complex projects will require a governance system that is not static and hierarchical, such as is commonly used for smaller projects. There needs to be scope for changes along the way, because both the planning and implementation phases of large public projects last for a long time. Governance will therefore assume different forms in the various phases of the project cycle. This highlighting of *flexibility* is supported by Müller et al. (2014), who seek to identify "organizational enablers" for good governance and governmentality. The most prevalent finding of their study is the importance of flexibility—the lower levels of governance require flexibility in the choice of methods and processes, whereas the higher levels of governance require flexibility in people's mindsets and attitudes toward work. Furthermore, there are needs for vision and values provided by top management and management's development of a culture that fosters flexibility and self-responsible employees.

The Importance of the Front-End

A project's life cycle consists of several phases (Figure 1). The *front-end phase* is the stage when the project only exists conceptually, before being operationalized. This encompasses all activities from when the idea is conceived until a final implementation decision is made. A distinction is commonly made between the *conceptual phase*, the *pre-study*, and the *pre-project*, as shown in Figure 1. In the conceptual phase, the conceptual solution and the overall project strategy are decided, and thus the key premises underpinning the project, as well as its characteristics and objectives. In the pre-study and pre-project phases, the decisions are more concrete with regard to contractual strategy, mode of delivery, and subsequently the detailed project design with regard to budget, activities, scope, schedule, and quality. This is followed by the *implementation phase*, which encompasses anything that happens after a final funding decision has been made, and includes *detailed engineering* and actual *construction*. Finally, the *operational phase* consists of commis-

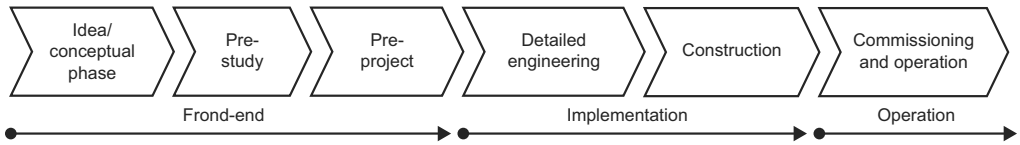


Figure 1: Project phases.

sioning and, subsequently, operation and maintenance.

A governance framework for the full life cycle of the project should be prepared at the outset, given that certain phases are more critical and in need of governance arrangements than others (HM Treasury, 2007). A number of authors have highlighted the importance of paying more attention to the front-end of projects to ensure project success (Shenhar, 2004; Williams & Samset, 2010; Morris, 2013, Samset & Volden, 2016). Morris (2013) highlights the importance of taking a holistic and "big picture" perspective on the project, and notes that in the early years, the project management community had an extremely narrow focus, reflecting only on the project itself and ignoring the critical front-end phase in which the most essential and overarching issues are decided. Many of the factors that later create problems in the construction phase, leading to projects delivering too late and over budget, arise early in the project definition stage (Morris, 2009). Williams and Samset (2010) note that the choice of concept has the largest impact on strategic project success and is thus highly critical. Other fundamental issues in the front-end are: to ensure realistic cost estimates (and counteract tactical budgeting); to ensure a rational planning process and a transparent democratic process; and to achieve predictability over time, since the front-end phase often extends over more than one parliamentary cycle.

A study of more than 1,000 projects, conducted by the World Bank, may provide solid evidence for the importance of the front-end phase (World Bank, 1996). A thorough review of the scope and quality of prior checks, prior assessment, and project design before the implementation of projects was linked to whether these turned out to be successful or not when examined in retrospect. The World Bank concluded that no less than 80% of the thoroughly prepared projects were successful, whereas as much as 65% of those initiated without proper preparation turned out to be unsuccessful. A corresponding study of 23 Norwegian projects delivered similar findings (Whist & Christensen, 2011).

The "Top Layer" and the Role of Central Government

This study is concerned with project governance from the perspective of central government, regarding investment projects that are funded by the state and implemented by line ministries and state agencies. In Norway, municipalities and counties are responsible for their own investments and may have their own governance schemes, which are not discussed here. We discuss how the governance of projects is currently organized and practiced at the overarching level. A governance framework is hierarchical, in the same way as a management system, where the top level is accountable for the whole system but delegates the responsibility and authority for defined actions to subordinate levels (Too & Weaver, 2014). Thus, central government, ultimately on behalf of the whole population, should set the conditions for projects (as well as other public sector activities) to deliver value to society; it should also impose *overarching requirements* with regard to, for example, structures, processes, and outcomes, but should not intervene in detailed project implementation (Samset, Berg, & Klakegg, 2006). Responsibility for implementing projects and programs is delegated to the different line ministries and agencies, which define the specific governance arrangements necessary to ensure tactical and operational project success.

Taking "the central government perspective" does not imply that we believe that central government can always be regarded as one unit and that all government decisions are made rationally. In practice, public project decisions are made through political processes in which agreements about goals and fundamental assumptions cannot be taken for granted (O'Leary, 2012), and in which there are many examples of irresponsible behavior, even from the top level (Miller & Hobbs, 2005). It is important to note that project governance structures and processes, which focus more on improving administrative processes than on political processes, do not ensure good *decisions*; they simply provide the framework within which good decisions can be made. This is probably the best one can do within a democratic political system.

An important part of governance schemes should be to ensure that decisions are lifted up to the appropriate level. Accordingly, the government itself should be involved in the management process on a strategic level, such as approving very large and critical projects. This is in line with the reform processes often referred to as Post-

New Public Management (Christensen, 2009), which is based on the premise that such an approach will enhance effectiveness and efficiency, without losing political impact.

The Study, Selected Countries, and Methodology

The starting point for this study was the Norwegian project governance framework, which the authors have followed for a number of years. The framework was an attempt to resolve or mitigate some common challenges observed in public projects in the 1990s, and the preliminary results are encouraging (Volden & Samset, 2017). However, it is only one of many possible ways to set up a project governance scheme, and our intention has been to review replicable systems in other selected countries, relate them to the Norwegian system and each other, and to discuss the following questions: Are they apt to ensure project success as intended? What are the differences and similarities between the schemes? What can Norway learn from the other countries and vice versa?

The other countries included in this study—the United Kingdom, Denmark, the Netherlands, Canada (Quebec Province), and Sweden—were selected primarily because they too are at the forefront in developing a public sector investment project governance system, with schemes introduced after the turn of the millennium. Quebec is merely one of several provinces of the Canadian Confederation, but has extensive independence in the area of infrastructure investments, and is included with the other studied countries due to its early initiative and advanced project governance scheme. Another determining factor was that all of the studied countries were OECD countries with a high level of economic development. There are, however, significant differences between the countries, not the least in their demographic and natural conditions, which implies that they differ also in their economic prospects for developing infrastructure. Norway (and to some extent also Sweden and Quebec) has a small population, long geographic distances, and areas that are virtually uninhabited, but nonetheless has a broad political consensus that the scattered and remote settlements should be maintained by building roads and public infrastructure. It is therefore obvious that the criteria for project selection may include societal objectives other than "value for money" in economic terms.

Two existing studies have compared the Norwegian governance framework with the British one (Williams et al., 2010), and with the British and Dutch frameworks (Klakegg, Williams, & Shiferaw, 2015), respectively. These studies constitute an important background for our study. Williams et al. (2010) conducted a case study and concluded that in all the four projects examined, the governance framework was useful in its own way, but also that there was some potential for improvement, such as more assessment of the project during the early stages (which has since been introduced in the United Kingdom). Klakegg et al. (2015) conclude that consistent project governance provides rewards, but they note that effort must be made to preserve the effect, otherwise it might "wear off." Another relevant study was conducted by Trafikanalys (2012), which has presented and discussed the systems regarding planning and assurance of transport projects in the Nordic countries, focusing mostly on cost figures. Other than the above-mentioned studies, we are not aware of any studies focusing on project governance models on a country level. Our study comprises more up-to-date descriptions of the governance frameworks in the same three countries as those studied by Klakegg et al. (2015), along with three additional countries. It is still no more than a case study, but it allows for comparisons that are somewhat more systematic and for evaluations of the development, content, context, and preliminary effects of the governance frameworks.

Some countries have more than one scheme, for example, depending on the sector. In these cases, we restrict the study to the governance models that concern the largest sectors measured by investment volume. Other schemes are cross-sectoral, such as the United Kingdom, Quebecian, and Norwegian schemes, and apply to all types of infrastructure investments. A common feature of all schemes, however, is that they are used for large investment projects that entail high costs or are highly complex.

This study is principally based on document reviews, backed up with interviews with key informants at the ministry level in the relevant countries and/or persons with special knowledge of the various schemes, in order to obtain documentation and verify the descriptions of the schemes. The documentation provided by the governments has varied. In some countries, the authorities have provided thorough descriptions of their schemes, and in the United Kingdom they have even made evaluations publicly available, whereas in other countries limited descriptions have been provided; therefore, we have had to supplement them with other sources, such as research reports and interviews. The information concerning the scheme in the Netherlands was primarily obtained from a doctoral dissertation that focused specifically on that scheme (Shiferaw, 2013).

In order to compare governance in the various countries, we have examined the development and content of the schemes, including which objectives countries have defined for them, which internal and external parties have been involved, their duties and responsibilities, how decisions have been made at the political level, and how the schemes have been structured at the project level. The reference point has been a scheme adhering to the recommendations from the literature, including both the overall principles of good governance and the more specific recommendations concerning project governance schemes. These principles and recommendations include stage-gate

approval processes with clearly defined phases, decision points and quality assurance, highlighting the front-end, lifting decisions to a high political level, being simple and flexible, promoting a portfolio perspective, and transparent processes and decisions.

The Governance Schemes

Norway

The background to the Norwegian governance scheme was a series of negative experiences with cost overruns, delays, and limited viability of some public investments in the 1990s, resulting in a government-initiated study to review the systems for the planning, implementation, and monitoring of large public investment projects. The authors of the study (Berg et al., 1999) concluded that the underlying documentation was deficient in a number of projects and that failures in the front-end phases were generally the main cause of problems during implementation. The authors proposed the introduction of an external quality assurance (QA) scheme in the decision phase for the largest public projects.

The QA scheme, introduced in the year 2000, is often referred to as the State Project Model, and is mandatory for investment projects with an anticipated budget exceeding 750 million Norwegian Kroner (approximately US$90 million).[3] It involves some 20 to 30 projects per year, mostly in the building, transport, construction, and information and communications technology (ICT) sectors. Initially, the purpose was to improve project efficiency, with a special focus on cost and delivery, but it was expanded in 2005 to enhance the effectiveness of the investments (i.e., more successful projects in terms of higher benefits for each Norwegian crown spent through, for example, improved cost control and conceptual solutions).

The Ministry of Finance is responsible for the administration of the QA scheme, which in principle, involves a very simple model with only two decision gates. No specific changes to the procedures of the various government agencies are required with respect to, for example, the implementation of the model, project organization, and use of steering groups, PMOs, or project sponsors, thus enabling them to implement their projects as before. Current requirements, however, are somewhat stricter with regard to the planning documents, intended to assure quality and the comprehensiveness of analyses. It is also a requirement that at least two conceptual solutions should be analyzed in addition to the zero option. This is intended to counteract the tendency for path dependency, which has largely characterized established practice. In contrast to previous practice, the documents prepared by the agencies (in some cases by the line ministries) have to be quality assured by external advisors before being submitted for appraisal at the political level. The quality assurers are pre-approved private consultants who have framework agreements with the Ministry of Finance. They have a limited mandate that requires them to examine the *quality* of the documents and not to address the political issues relating to the choice of project. They are also required to perform a separate independent, probability-based, cost estimation and a business case.

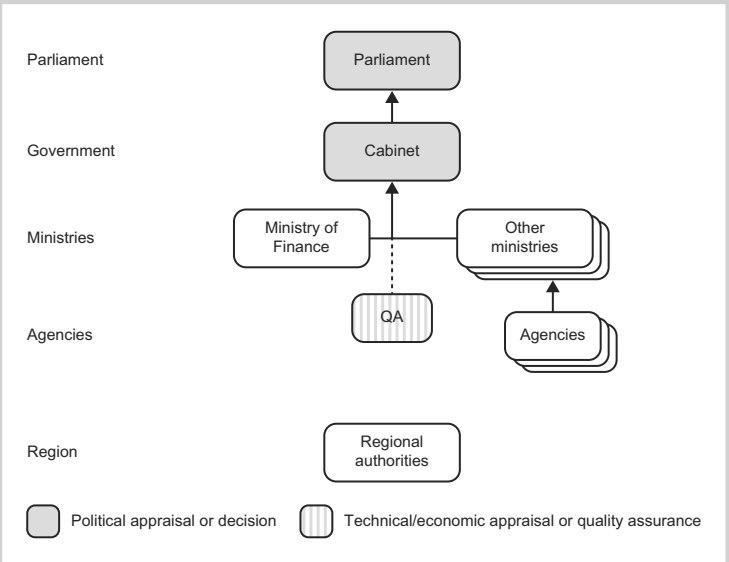


Figure 2: Investment project governance in Norway.

Figure 2 shows the roles and principles in investment project governance in Norway. Individual ministries are responsible for new investment initiatives, the vast majority of which are initiated and planned by a subordinate agency. These planning documents are then subjected to external quality assurance on behalf of the relevant line ministry and the Ministry of Finance. The line ministry will summarize the findings and recommendations in a memorandum, which will be submitted to the Cabinet for political appraisal before the matter is presented to Parliament for its approval and final decision.

The State Project Model involves two stages, as shown in Figure 3. The first stage concerns the choice of concept. The agency's pre-study (comprising an

[3]NOK 750 million

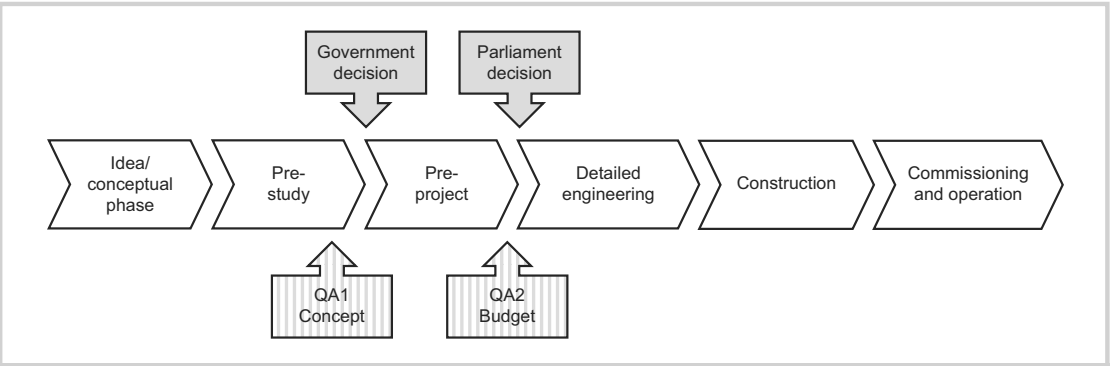


Figure 3: The Norwegian State Project Model.

assessment of needs, alignment with government strategy, the opportunity case, and the business case) is subjected to an external quality assurance of the choice of concept (QA1). At this stage, it will be decided at the central government level whether to reject the project or to move on to the pre-project phase, and in such cases which concept to choose.

At the next stage, when the pre-project has been finalized, the agency has to present an overall project management document, which provides information on, for example, objectives, budgets and target cost, implementation strategy, and contract strategy. This document is then subjected to external quality assurance of the cost estimate and management documentation (QA2). Budgets are based on formal uncertainty analyses and stochastic cost estimation. The recommended budget will commonly be close to the P85 level, and the recommended target cost for the responsible agency is normally lower and close to the P50 level.[4]

[4]With stochastic (probability-based) cost estimation based on either mathematical analytical methods or simulation, the result will be a cumulative probability distribution of the investment cost. P85 implies that the cost will be at or below this level with 85% probability. Similarly, there is a 50% chance that a budget at P50 will be adhered to. The budgeted cost should include a residual reserve and therefore be higher than the expected cost. At the same time, the target cost for the agency should be more ambitious, to give incentives for efficiency and cost control. In Norway, the difference between the budgeted cost and the target cost is kept as a residual reserve, normally on a ministry level.

The line ministry and the Ministry of Finance will summarize the quality-assured documents and the recommendations based on them, in a memorandum to the government. Special prominence is then given to the proposed budget; thereafter, the government will submit the matter to Parliament, which will make the final decision and stipulate both the budgeted cost that commits the responsible ministry, and the target cost that commits the agency. Alternatively, Parliament may reject the project at this level.

The Other Case Countries—Establishment and Scope

In common with Norway, the background to the governance schemes in the other five case countries was negative experiences from past projects, especially with regard to cost overruns and delays:

- Like Norway, the United Kingdom was a pioneer. In the year 2000, a separate unit—the Office of Government Commerce (OGC)—was created at HM Treasury, to manage a scheme applying to the largest and riskiest public projects. Initially, it focused on budgets and project management documentation, drawing on experts from the private sector, and a number of follow-up points throughout the project life cycle. The OGC developed a standardized gateway process and public project methodology that came to be widely disseminated. Subsequently, the scheme was strengthened, with focus increasingly being placed on the front-end (a so-called Starting Gate review was introduced), on portfolio management, and on the education of public project leaders. In 2011, a new unit, the Major Projects Authority (MPA), was established, with a stronger mandate, given directly by the Prime Minister, and this unit reports jointly to HM Treasury and the Cabinet Office. In 2016, the MPA merged with Infrastructure UK to form the Infrastructure and Projects Authority (IPA). Preliminary evaluations suggest positive effects of the scheme on project management and cost savings. Main sources for a description of the scheme and experiences are HM Treasury and Cabinet Office (2011) and National Audit Office (2012).
- Denmark, inspired by the United Kingdom and Norway, launched a scheme applying to transport projects, in the wake of a study of cost overruns in 12 transport projects (Ministry of Transport and Building, 2010a, 2010b, 2010c, 2015). A financial management model was established in 2003 to streamline decision-making processes for the various sectors. In 2007, the financial management model

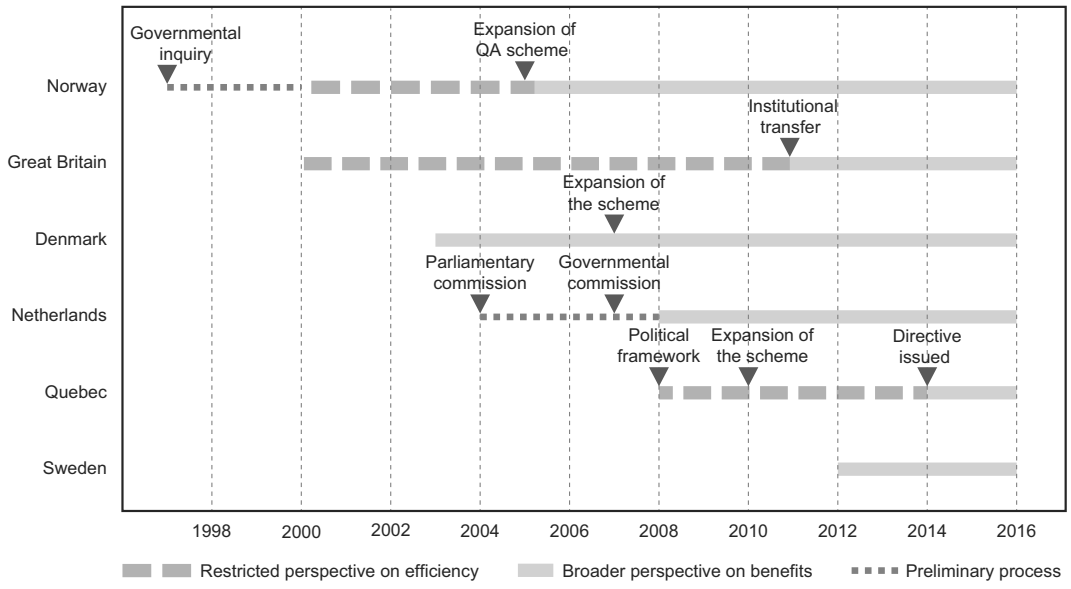


Figure 4: Introduction of investment project governance schemes in six countries.

was expanded, both by requiring projects in excess of US$36 million[5] to be subjected to external quality assurance, and by adding an experience-based correction factor to the cost estimate. Decisions are lifted to the parliamentary level.

- In the Netherlands,[6] each ministry is responsible for its own major projects. In 2008, the Ministry of Infrastructure and the Environment, which has by far the largest portfolio, introduced an integrated investment program, MIRT, which includes a mandatory stage-gate process. The predominant issue to be addressed was how to avoid cost-overruns and speed up the implementation of major infrastructure projects, but also more generally to ensure a robust foundation for major projects, with broad participation from affected parties, commitment at the political level, and the assessment of several alternative conceptual solutions.
- In Quebec, in 2008, the Treasury Board established a political framework for the governance of large public sector investment projects. The scheme was revised and strengthened in 2010, and again in 2014, when it was given in the form of a *directive*, with increasing focus on the front-end (Secretariat du Conseil du trésor, 2014). The organization of the scheme has been developed and strengthened over time. In 2014, a unit reporting to the Treasury Board, the Société Québécoise des Infrastructures (SQI) was established as the project manager for all major infrastructure projects, in association with the sponsoring line ministry.
- Sweden was the last of the case countries to introduce a governance scheme, which happened in 2012, and only for transport projects (see Trafikverket, 2014). Traditionally, the transport agency has had a rather independent role, but decisions regarding major projects are now lifted up to a government level, as in the other two Scandinavian countries. The decision base for the choice of concept includes assessments of needs and alternative options, and more formalized uncertainty analyses of cost estimates have entered into use in recent years.

The background to and development of the various schemes in the six cases are summarized in Figure 4. It is interesting to observe that several countries have expanded and strengthened their schemes over time, and reorganized the management of them. Generally, the purpose of most of the schemes initially related to the efficiency aspect in the implementation of the projects. Later, a somewhat broader perspective on the front-end and the choice of concept was adopted in Norway, the United Kingdom, the Netherlands and, finally, in Quebec. In Sweden, the requirement for conceptual appraisals as well as environmental assessments

[5]DKK 250 million

[6]Main source: Shiferaw (2013).

have long been a key focus. Denmark still has a somewhat narrower focus, but includes economic and transport analyses.

Overview of the Schemes

A detailed comparison of the various governance schemes in the six case countries is presented in Table 1. The findings concerning specific elements of the schemes are discussed separately in the following subsections.

Projects Covered

The United Kingdom, Quebec, and Norway have governance schemes that in principle apply to all sectors where the state is responsible for infrastructure projects (funding, procuring and, in many cases, implementing and operating). The other countries have schemes that only apply to one or some sectors, and certain sectors are exempted, as in Norway.

In all countries, the schemes focus on projects with central government funding that are large, complex, or otherwise involve risk on the part of the central government. Only three countries have introduced a general threshold value defining which projects should be encompassed: Norway, Denmark, and Quebec. A threshold value is a simple criterion for deciding whether a project is subjected to the regime, but its application may seem rigid and not always optimal. This criticism has been leveled against the Norwegian scheme. In the United Kingdom, the Infrastructure and Projects Authority makes an overall assessment as to whether a project should be encompassed by the scheme, and it has chosen to include a considerable number of modernization projects that are "small" in terms of investment cost, but highly complex and innovative, and thus risky.

Parties and Roles

Figure 5 provides an overview of how project governance is organized in the six case countries. The gray boxes imply influence over decisions taken at the various gates in the stage-gate models, with a special focus on the front-end (choice of concept and final approval of the project), and the pattern-filled boxes indicates the quality assurance function.

We find that the government plays a key role as a decision maker in all countries, primarily with regard to the final choice of project alternative. In the Scandinavian countries, the final approval is elevated all the way to the parliamentary level. Presumably, this has to do with these countries being relatively small, but also because they normally have minority governments, and thus need support at the parliamentary level. It may also be noted that many central government–funded investment projects in Scandinavia, especially within transport, are highly politicized, and not viewed exclusively as measures for national economic growth (Boge, 2006).

By contrast, in the United Kingdom and Quebec the Treasury has an important role in advising the government, based more on economic and technical considerations than on political concerns. Klakegg et al. (2015) generally hold that the UK scheme is somewhat more "business-like" than the Norwegian one; it is largely based on best practice in the private sector, and attaches major weight to financial and profitability issues. In the Netherlands, the role of advising the government is performed by an inter-ministerial commission (ICRE) with representatives from the various ministries, and with the Ministry of Finance having a very strong position. It should also be noted (although not shown in Figure 5) that the Dutch scheme involves broad participation of stakeholders in the front-end of public projects in a more systematic manner than those in the other countries—the purpose being to pull discussions toward the front-end and avoid tugs-of-war in later stages.

Most project appraisals are conducted at the agency level in all countries, with their sponsoring line ministries being involved to varying degrees. Quebec stands out in that the new designated government agency, SQI, is responsible for all infrastructure projects across sectors. The quality assurance function is performed by parties independent of those who conduct the appraisals, and these reviewers have a key role in most countries, feeding their advice directly into the decision-making process.

Quality Assurance Reviews

Independent quality assurance reviews are performed in all the countries. In Norway, the use of external experts has been controversial. The criticism is partly that it prevents the development of adequate expertise within the public administration; partly that the consultants do not possess enough sector competence; and partly that when a group of consultants is pre-qualified for such work, it may achieve something akin to a monopoly position. The same kind of criticism is heard in Denmark. In Sweden, where much of the quality assurance takes place on an ad hoc basis and internally within government agencies, the criticism is rather that it becomes difficult to ensure that the quality assurance is sufficiently independent and professional. In the Netherlands, Quebec, and the United Kingdom, designated public bodies are established to perform the quality assurance function. In Quebec, quality assurance is performed both internally in the SQI, and then again by SCT at the Treasury Board before the project is presented to political decision makers.

An important principle of all schemes is that the external quality assurance arrangement only has an indirect impact on the decision-making process. The decisions are to be made at the political level, and the recommendations of the quality assurer have advisory status only.

The Stage-Gate Models

All six countries use stage-gate models in their governance schemes, defining the number of project phases, decision points, and the types of analyses

Governance of Major Public Investment Projects

Criteria/Country	Norway	Denmark	Sweden	The Netherlands	United Kingdom	Quebec
Who initiates the QA process?	Ministry of Finance	Ministry of Transport	Agency	A designated government agency	A designated government agency	A designated government agency
Who decides the choice of concept?	Government	Parliament	Agency or Government	A designated government agency	Treasury[1]	Council of Ministers
Who determines the budget?	Parliament	Parliament	Agency or Government	Government	Treasury	Government
Sectors included[2]	All, with some exceptions[3]	Transport sector	Transport sector	Infrastructure projects	All sectors[4]	Infrastructure projects
Threshold value (million)	NOK 750	DKK 250	No	No	Large projects[5]	CAD 50
Who appraises the project?	Agency or ministry[6]	Agency	Agency and regional authority	Responsible government agency	Agency or ministry	A designated government agency
Who performs quality assurance?	External consultants	External consultants	A designated government agency, and internally	A designated government agency	Independent quality assurers[7]	A designated gov. agency
Requires co-funding from promoters	No	No	No, but may happen	For all in excess of EUR 60 billion	Desired, but no requirement[8]	To be considered, not required
Budgeted cost	P85 (normally)	Basic calculation + 20%[9]	In the portfolio		Estimate plus supplement	Estimate plus supplements[10]
Target cost	P50 (normally)	Basic calculation + 10%	Budget[11]		Estimate plus supplement	Budget
Decision points	2	2	2	3	5	5
QA or advisory interventions	2	2	Ongoing	1	6	Ongoing
Transparency	Yes	Limited	Limited	Limited	Some	Limited
Portfolio management as part of the scheme	No	No	No	Yes	Yes	Yes

Notes:

[1]Concerns approval of business case; the line ministry may have determined the choice of concept much earlier

[2]Some countries may have different schemes in some sectors

[3]All, except for health, oil/gas, and state enterprises

[4]Central government infrastructure investments and ICT/restructuring projects

[5]No threshold value; relevant factors are size, complexity, requirement for a separate statute, and the degree of innovation

[6]External resources are drawn on in some cases, from the private or public sector, including QA resources

[7]Both private and public sector technical experts

[8]This varies between sectors

[9]The 20% supplement is managed at the portfolio level and is transferable from one year to the next

[10]The government should be informed if it is anticipated that the budget will be overrun

[11]Recently based on stochastic cost estimation (P50).

Table 1: A comparison of the governance schemes in six countries.

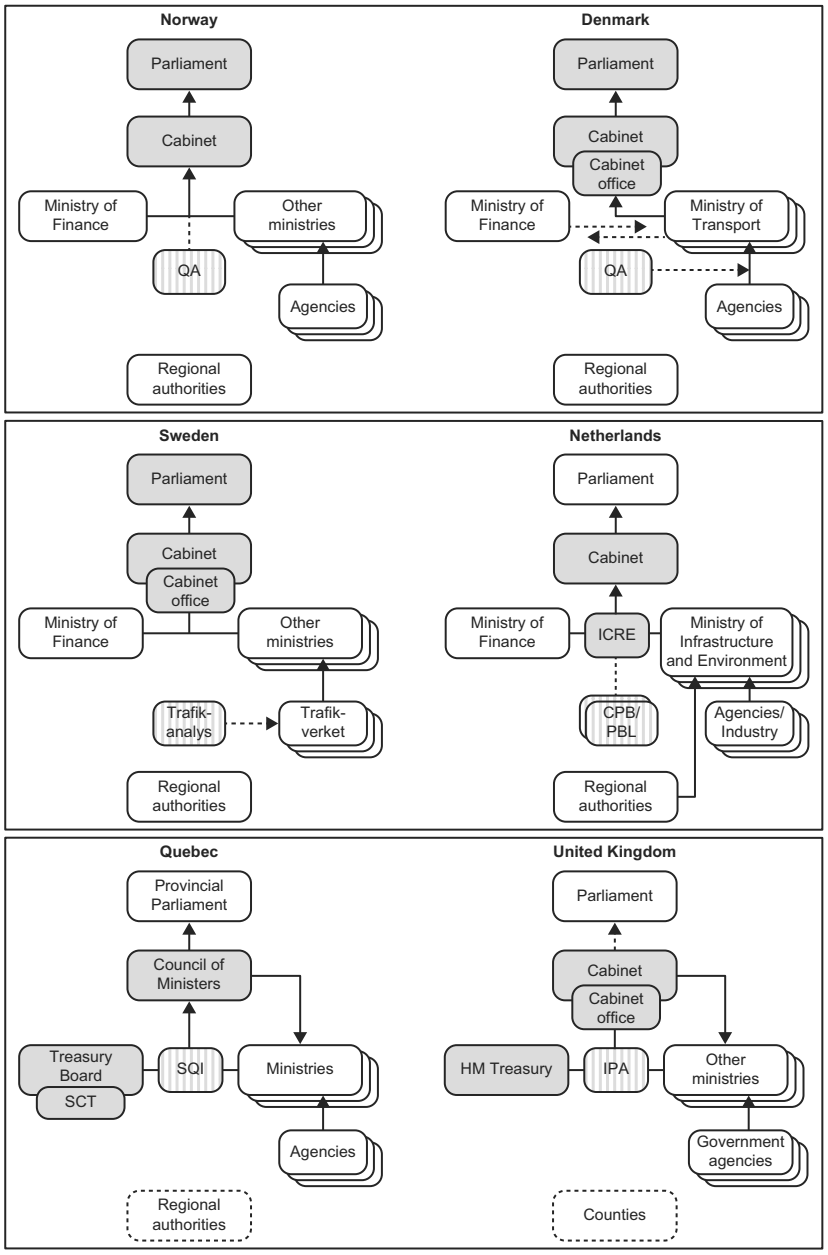


Figure 5: Responsibility for appraisal/quality assurance and decisions under the various governance schemes in six countries.[7]

[7]ICRE, inter-ministerial commission for improvement of the structure of the economy in the Netherlands; CPB, Netherlands Bureau for Economic Policy Analysis; PBL, Netherlands Environmental Assessment Agency; SCT, Secretariat du Conseil du trésor; SQI, Société Québécoise des Infrastructures; IPA, Infrastructure and Projects Authority.

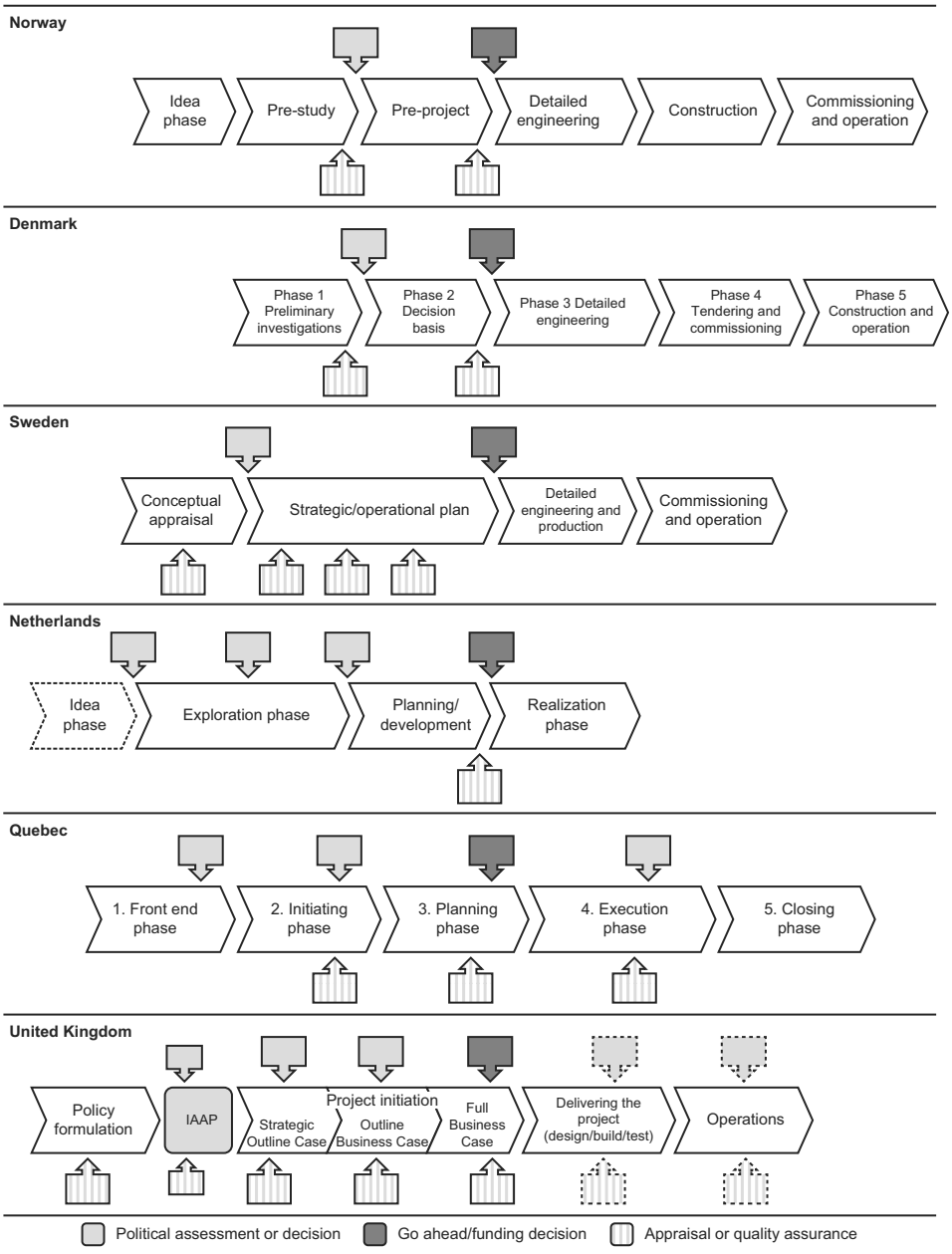


Figure 6: Summarized versions of the six stage-gate models.[8]

[8]IAAP is an integrated assurance and approval plan.

and independent reviews required at the various stages. The number and names of the phases are more or less the same in all countries, but we find larger variation in the number and locations of reviews and decision points, as shown in Figure 6. The Scandinavian countries are distinguished by formal decision points and quality assurance being limited to the front-end phase, whereas the other three countries have follow-up points during project implementation and closure, and in the United Kingdom, for some projects all the way into the operational phase.

As far as the number of decision points is concerned, Denmark and Norway have the simplest schemes, with only two decision points. With regard to quality assurance, the Netherlands have the simplest arrangement, with only one review. Quebec and the United Kingdom have the largest number of decision gates. The UK scheme is the most comprehensive, involving the most detailed control measures and requiring the preparation of a separate plan for the subsequent follow-up and quality assurance of each project. However, the UK model is also flexible in the sense that the number of intervention points and their scopes are decided on a project-by-project basis and may be changed throughout the project.

It should also be noted that the scope of a review varies. The reviews in the Norwegian scheme are rather time consuming, inasmuch as the quality assurer is required to perform his or her own independent analyses, and not only oversee the work that has been done. By contrast, in the UK scheme, the number of checkpoints is large, but each quality assurance exercise is slightly simpler.

In Norway, the first decision point concerns the choice of concept, after the pre-study phase. In recent years, some of the other countries have introduced a formal decision gate at an even earlier stage. In the United Kingdom, for example, the Starting Gate review process was introduced in 2011, clarifying the strategic premises underpinning the choice of alternative concepts, but not involving technical analyses of specified alternatives at this stage. The first stage of the business case is not a detailed appraisal of alternatives, but rather a rough analysis, with the purpose of reducing the opportunity space from a long list to three or four alternatives. Similarly, the Dutch scheme is strongly focused on early assessment of solutions to a problem and broad involvement of stakeholders. This is an interesting observation, as it is generally appreciated that premises laid down at this stage may have a decisive impact on the actual choice of concept. In Norway, early experiences indicate that at the QA1 stage many premises are already laid down and some stakeholders have high expectations related to a specific solution. In such cases, we may see that the pre-study includes alternatives that are variants of the same concept rather than truly different solutions.

Cost Estimation

As far as cost control is concerned, a key element of the Norwegian governance scheme has been the introduction of a *budgeted cost* and a distinct, lower *target cost* for the agency. The difference between the two figures is the contingency reserve, which is normally controlled by the line ministry. The figures are based on probability-based cost estimation (using the "successive principle") and are reviewed by external consultants who will normally recommend a budgeted cost at or close to P85, and a target cost at P50. Parliament's decision normally follows the recommended figures.

Norway and, recently, Sweden too are apparently alone in using probability-based estimation in each project. Denmark has an advanced system and methods for cost estimation, including an extensive cost database, but a basic cost estimate is applied, to which is added a general supplement of 10% for the agency and 20% for the ministry. The 20% supplement is thus available at the portfolio level, and is transferable from one year to the next. Hence, the latter provides the ministry with somewhat more freedom of action than under the Norwegian scheme. In the United Kingdom, there does not seem to be a distinction made between target cost and budgeted cost, but an uncertainty level is chosen for each case (e.g., P50 if central government is willing to assume a high risk of cost overruns or if the project forms part of a large portfolio) and optimism bias correction factors are used, based on rules of thumb tailored to the chosen uncertainty level. The other countries apply a budget that has to be adhered to, but may add a notional supplement that is not to be exceeded; however, if this does happen, the government must be informed.

We have not been able to address specifically the experiences of individual countries with the various budget estimation principles in this study, but this would be an interesting issue for potential follow-up. Lessons from the Norwegian model thus far indicate that projects under the scheme are now largely completed within their cost frames (Volden & Samset, 2017). The deviation between the final cost and the target cost is almost symmetrically distributed around the median. Hence, at the portfolio level, the government is able to control the cost of major investment projects more effectively. Whether this can be explained by the use of stochastic estimation, thorough external quality assurance, or the practice of establishing a lower target cost for the agency, or a combination of these, remains to be proved.

Co-Funding Requirement

In all six countries, the governance schemes are applicable to projects with central government funding; however, they are often initiated locally and benefit specific groups or regions, thus giving rise to perverse incentives (Volden & Samset, 2015). The conditions attached to such funding differ between the countries. The Scandinavian countries

stand out in that generally they do not require co-funding from those who will benefit from the projects. The exception is the road sector in Norway, where a significant element of user charges has been introduced in recent years. The Netherlands is distinguished by *requiring* co-funding from local authorities who come forward with a project proposal. The rationale is that this signals commitment and a willingness to pay, which increases the likelihood that the project idea is feasible. There is also a requirement that all investment initiatives in excess of EUR 60 million (approximately US$64 million) have private co-funding. The rationale is that this will result in more weight being attached to long-term revenue flows (in the form of user fees) as well as efficient project implementation. In the United Kingdom and Quebec, the central government has signaled a desire for co-funding from local authorities and the private sector in certain areas, although there is no requirement.

Transparency

Transparency is a key criterion for defining good governance. As noted by Klakegg and Volden (2016), the public sector depends on transparency as a means of strengthening accountability, where the private sector has competition. In major public projects, it is a question of ensuring that the decision-making processes and administrative processes are well documented. There is also a prerequisite for another governance principle—participation—in order to give stakeholders and the general public an opportunity to express their views in the process.

The Norwegian scheme attaches great importance to transparency. The Ministry of Finance currently funds a research program to follow the scheme closely and collect information about the projects. All QA reports are published on the program's website.[9] This has undoubtedly made all the actors involved, including the reviewers, put a lot of effort into their work and has resulted in high-quality plans and estimates. Furthermore, as the projects are finalized and enter their operational phases, cost figures and other project results are made available to the public.

None of the other countries seems to practice the same level of transparency as Norway, although several of them have expressed a concern for this matter. In the United Kingdom, the IPA publishes valuable information about major projects in its annual reports, although most of the data are on the group level and published with a considerable time lag. Preliminary evaluations of the UK scheme recommend that more data be published earlier and at the project level. In other countries, there is hardly any publicly available information about the projects.

The Portfolio Perspective

The Norwegian governance scheme focuses on requirements applicable to individual projects, and does not impose explicit portfolio evaluation requirements. The same is essentially the case for the schemes in all other Scandinavian countries. Nonetheless, it must be expected that the high level of transparency will make it easier for the line ministries to make decisions from a portfolio perspective. It must also be expected that overall project risk and the need for a contingency reserve will be influenced by whether or not the project forms part of a larger portfolio.

In the Netherlands, the MIRT program was introduced along with the requirements applying to the individual projects. The intention was to ensure coherence and synergy and to facilitate portfolio management within the Ministry of Infrastructure and the Environment. The UK and Quebecian schemes are also intended to include a portfolio perspective, inasmuch as a central government unit is responsible for compiling data on all infrastructure projects in the portfolio, thus making it possible to analyze and manage them collectively. These units are also responsible for training and facilitating learning across sectors; thus, both the IPA (United Kingdom) and, to some extent, the SQI (Quebec) have similarities with an organization's strategic PMO, although in this case working on the central government level. It should also be noted that in the United Kingdom, quality assurance is to be performed not only on individual projects but also at the program and portfolio levels at regular intervals. However, there is much to suggest that this potential has not yet been realized.

Assessments and Conclusions

A number of international studies have highlighted the problems of managing public investment projects with respect to operational, tactical, and strategic aspects. Special measures are therefore required to ensure successful implementation and outcomes. Norway was a pioneer and, in the year 2000, introduced an overarching governance framework for major public projects. The framework and its effects, some of which are very encouraging, have already been presented in earlier literature. In recent years, a number of countries have introduced similar frameworks in which independent quality assurance is duly coordinated with the decision points. Six schemes are presented and compared in this study.

We found that the six governance schemes have many characteristics in common. They were all established for project governance by central government, and they apply to large projects that involve particularly high costs, risk, and complexity, or that are highly innovative. They all apply a stage-gate model at the project level, with clearly defined roles and responsibilities, including independent quality assurance reviews of project documentation at specified decision points. They also have measures to avoid optimism bias in the cost estimates, and they place key decisions, as well as responsibility, for managing the scheme at a high level in the system.

[9]Retrieved from http://www.ntnu.no/concept/ks-rapporter

Overall, the schemes seem to be fairly consistent with the recommendations from the literature; some exceptions are that only the Netherlands requires co-funding from beneficiaries to obtain state funding, and only Norway highlights transparency at the project level. Furthermore, there is potential for improvement in several countries when it comes to integrating the portfolio perspective. The various schemes' emphases vary somewhat, but we have observed a general development over time toward more focus on the front-end and the choice of concept. This is in line with general development within the project management community (Morris, 2013). All six countries now require needs assessments and the evaluation of alternative conceptual solutions, which demonstrates that their importance is duly acknowledged.

At the same time, we know that the final project choice is not only the result of systematic investigation of alternatives by professionals and experts. In many cases, the politicians' priorities carry more weight, and this needs to be tolerated within a democratic political system. It is nevertheless essential in a project governance scheme to bring in technical and economic expertise at an early stage in order to identify and, if possible, eliminate the worst alternatives or conceptual solutions. Within a political reality, there is no guarantee that the best alternative will be chosen, but we can possibly avoid the worst ones. To quote Herbert Simon (1976), in many cases the realistic scenario would be not to aim for "maximizing," but to put the bar at "satisficing."

There are a number of significant differences between the six schemes, such as in the use of internal or external experts, in the demarcation between the political and technical spheres, and in the comprehensiveness of the schemes, the organization of the schemes, and the extent to which projects are assessed individually or as part of a public project portfolio. Some of the differences can probably be explained by historical and cultural differences, such as the Scandinavian countries' involvement of Parliament in the approval of individual projects. However, both Norway's and Denmark's use of private consultants as opposed to the United Kingdom's and Quebec's use of a government unit, is not what might be expected (cf., the Nordic "strong state" tradition versus the Anglo/American market orientation). All in all, we are faced with two main types of project governance schemes: the schemes in the Scandinavian countries and the other schemes. The former are relatively simple in terms of the number of intervention points, although these may be comprehensive in terms of which analyses are to be performed. The schemes do not intervene significantly in existing processes and practices, but impose new qualitative requirements with regard to appraisal and documentation. The schemes in the three other countries are more ambitious and extensive, with more follow-up points, also during the implementation phase. With regard to Williams et al.'s (2010) distinction between governance of and governance through projects, it might be claimed that whereas the Scandinavian schemes are only about governance *through* projects, the other schemes are also about governance *of* projects. The United Kingdom and Quebec have altered the organization of their schemes several times, and now have centrally placed units with a clear mandate to managing the quality assurance function, as well as responsibility for the support and development of expertise, and compiling and publishing data on the portfolio level. In Quebec, a central organization is even mandated to serve as project manager for all major infrastructure projects.

The Norwegian scheme currently aims to achieve control over costs and progress, and also to ensure that investments deliver economic benefits. The scheme is intended to have a disciplining effect, both on the agents responsible for the projects and on their sponsoring ministries. The impact on cost control seems quite satisfactory (Volden & Samset, 2017). However, we need more knowledge about the effects of the various measures, such as the use of probability-based estimation, the role of private sector reviewers, the use of a lower target cost for the agency, and the focus on increased transparency. There are also objections relating to, for example, time and resource use, how the use of private consultants prevents the development of central government expertise, and the scheme being rigid and inflexible. It has also been argued that QA1 takes place too late and that the analysis of alternatives may turn into more of a ritual exercise than a forceful tool used to identify the best conceptual solution. In this regard, it would be useful to learn more about the experiences obtained with interventions at an earlier stage in other countries, such as Starting Gate reviews in the United Kingdom.

The schemes described in this article were all introduced in recent years and have not been in operation sufficiently long for any conclusions to be drawn as to their effects. The ultimate question is whether some schemes are more effective than others in improving project delivery as well as outcome, and to what extent an effective scheme can be applied also in other countries. This will be a topic for future research. The fact that there are several different governance schemes in operation is positive, and they might inspire alternative ways of organizing and implementing such schemes in the future. It should be noted that we have focused only on the top layer of project governance introduced by central government, assuming that the governance arrangements at the level below (e.g., line ministry, department, and agency) are in place to ensure tactical and operational project success. A topic for future research could be to address the question of whether the rather simple schemes in the Scandinavian countries are matched by the necessary requirements, guidelines, and training on the lower levels. Furthermore, we have only looked at the structural elements of a project governance scheme. Future research should

also seek to determine how these work together with the relationship-based elements on different levels of government.

A further hypothesis, which is perhaps too difficult to test, relates to the trickle-down effects, if any. This concerns whether improvements in project governance and governmentality on other levels and for smaller projects can be attributed to the overarching schemes discussed in this article. To date, the indications from the Norwegian scheme are that the spinoffs may be considerable, not only in the public sector but also among the external quality assurers, project management consultants, contractors, and suppliers, and in the research community.

References

Association for Project Management (APM). (2011). *Directing change: A guide to governance of project management.* Princes Risborough, UK: Association for Project Management.

Bemelmans-Videc, M. L., Rist, R. C., & Vedung, E. (Eds.). (1998). *Carrots, sticks, and sermons: Policy instruments and their evaluation*. New Brunswick, NJ: Transaction.

Berg, P., Andersen, K., Østby, L. E., Lilleby, S., Styrvold, S., Holand, K., Korsnes, U, Rønning, K., Johansen, F., & Kvarsvik, T. (1999). *Styring av statlige investeringer. Sluttrapport fra styringsgruppen*. Oslo, Norway: Finansdepartementet.

Bevir, M. (2013). *A theory of governance*. Berkeley, CA: Gaia Books, University of California.

Boge, K. (2006). *Votes count but the number of seats decides: A comparative historical case study of 20th century Danish, Swedish and Norwegian road policy*. Dr. Oecon dissertation. Series of Dissertations 4/2006. Oslo, Norway: Norwegian Business School.

Christensen, T. (2009). *The Norwegian front-end governance regime of major public projects: A theoretically based analysis*. Concept rapport nr. 23. Trondheim, Norway: Concept-programmet.

Council of Europe (2014). *The 12 principles for good governance at local level, with tools for implementation.* Retrieved from www.coe.int/t/dgap/localdemocracy/Strategy_Innovation/12principles_en.asp

Crawford, L., Cooke-Davies, T., Hobbs, B., Labuschagne, L., Remington, K., & Ping, C. (2008). Governance and support in the sponsoring of projects and programs. *Project Management Journal, 39*, 43–55.

Flyvbjerg, B., Skamris Holm, M. K., & Buhl, S. L. (2003a). How common and how large are cost overruns in transport infrastructure projects? *Transport Review, 23*(1), 71–88.

Flyvberg, B., Bruzelius, N., & Rothengatter, W. (2003b). *Megaprojects and risk: An anatomy of ambition.* Cambridge, UK: Cambridge University Press.

Hall, P. (1981). *Great planning disasters.* Harmondsworth, UK: Penguin Books.

HM Treasury and Cabinet Office (2011). *Major project approval and assurance guidance.* London, UK: HM Treasury.

HM Treasury (2007). *Project governance: A guidance note for public sector projects.* London, UK: HM Treasury.

Hobbs, B., & Aubry, M. (2008). An empirically grounded search for a typology of project management offices. *Project Management Journal, 39*(1), 569–582.

Haanes, S., Holte, E., & Larsen, S.V. (2006). *Beslutningsunderlag og beslutninger i store, statlige investeringsprosjekter*. Concept rapport nr. 3. Trondheim, Norway: Concept-programmet.

Klakegg, O. J., & Meistad, T. (2014). *Individual relations – the core line in project based organizations*. IRSPM Conference, 2014, Ottawa, Canada.

Klakegg, O. J., & Volden, G. H. (2016). Governance in public projects: The Norwegian case. In R. Müller (Ed.), *Governance and governmentality for projects: Enablers, practices and consequences*, pp. 129–156. New York, NY: Routledge.

Klakegg, O. J., Williams, T., & Shiferaw, A. T. (2015). Taming the "trolls:" Major public projects in the making. *International Journal of Project Management, 34*, 282–296.

Le Quesne, T., & Parr, T. (2016). *Major capital programmes: A discussion document based on insights from recent experience.* London, England: Infrastructure and Projects Authority.

McKinsey Global Institute. (2013). *Infrastructure productivity: How to save $1 trillion a year*. McKinsey Global Institute.

Merrow, E. W. (2011). *Industrial megaprojects: Concepts, strategies and practices for success.* Hoboken, NJ: John Wiley & Sons.

Miller, R., & Lessard, D. R. (2000). *The strategic management of large engineering projects.* Cambridge, MA: MIT Press.

Miller, R., & Hobbs, B. (2005). Governance regimes for large complex projects. *Project Management Journal, 36*(3), 42–50.

Ministry of Transport and Building (2010a). *Ny anlægsbudgettering—Hovednotat. (New project budgeting—Main report).*

Ministry of Transport and Building (2010b). *Terms of reference for ekstern kvalitetssikring af beslutningsgrundlag på niveau 1. (Terms of reference for the external quality assurance of decision base on level 1).*

Ministry of Transport and Building (2010c). *Terms of reference for ekstern kvalitetssikring af beslutningsgrundlag på niveau 2 (Terms of reference for the external quality assurance of decision base on level 2).*

Ministry of Transport and Building (2015). *Evaluering af ny anlægsbudgettering (Evaluation of new project budgeting).*

Morris, P. W. G., & Hough, G. H. (1987). *The anatomy of major projects.* Chichester, UK: John Wiley & Sons.

Morris, P. (2009). Implementing strategy through project management: The importance of managing the project

front-end. In T. Williams, K. Samset, & K. J. Sunnevåg (Eds.), *Making essential choices with scant information*, pp. 36–97. New York, NY: Palgrave Macmillan.

Morris, P. W. (2013). *Reconstructing project management*. West Sussex, UK: John Wiley & Sons.

Morris, P. W. G., & Geraldi, J. (2011). Managing the institutional context for projects. *Project Management Journal, 42*(6), 20–32.

Müller, R. (2009). *Project governance: Fundamentals of project management*. New York, NY: Gower.

Müller, R., Pemsel, S., & Shao, J. (2014). Organizational enablers for governance and governmentality of projects: A literature review. *International Journal of Project Management, 32*, 1309–1320.

Müller, R., Shao, J., & Pemsel, S. (2015). *Governance and governmentality of projects. Summary Report*, April 2015.

Narayanan, V. K., & DeFillippi, R. (2012). The influence of strategic context on project management systems: A senior management perspective. In T. Williams, & K. Samset (Eds.), *Project governance: Getting investments right*, pp. 3–45. New York, NY: Palgrave Macmillan.

National Audit Office (2012). *HM Treasury and Cabinet Office. Assurance for major projects*

O'Leary, T. (2012). Decision-making in organizations. In T. Williams, & K. Samset (Eds.), *Project governance: Getting investments right*, pp. 175–220. New York, NY: Palgrave Macmillan.

Organisation for Economic Co-operation and Development (OECD). (2015a). *Principles of corporate governance*. Paris, France: OECD.

Organisation for Economic Co-operation and Development (OECD). (2015b). *Towards a framework for the governance of infrastructure*. Paris, France: OECD.

Pinto, J. (2006). *Organizational governance and project success: Lessons from Boston's big dig.* Concept International Symposium 2006, Trondheim, Norway.

Project Management Institute (PMI). (2013). *A guide to the project management body of knowledge (PMBOK® guide)* – Fifth edition. Newtown Square, PA: Author.

Samset, K., Berg, P., & Klakegg, O. J. (2006). *Front-end governance of major public projects*. Paper presented at the EURAM Conference 2006, Oslo, Norway.

Samset, K., & Volden, G. H. (2016). Front-end definition of projects: Ten paradoxes and some reflections regarding project management and project governance. *International Journal of Project Management, 34*(2), 297–313.

Samset, K., Volden, G.H., Olsson, N., & Kvalheim, E.V. (2016). Governance schemes for major public investment projects. A comparative study of principles and practices in six countries, Concept report no. 47, Trondheim, Norway.

Secretariat du Conseil du trésor. (2014). *Directive sur la gestion des projets majeurs d'infrastructure publique*. Quebec, Canada: Secretariat du Conseil du trésor.

Shenhar, A. J. (2004). Strategic Project Leadership® toward a strategic approach to project management. *R&D Management, 34*, 569–578.

Shiferaw, A. (2013). *Choice of project concept and decision-making: An international perspective*. Thesis for the degree of Philosophiae Doctor. Trondheim, Norway: Norwegian University of Science and Technology

Simon, H. (1976). *Administrative behavior* (3rd ed.). New York, NY: The Free Press.

Tirole, J. (2001). Corporate governance. *Econometrica, 69*(1), 1–35.

Too, E., & Weaver, P. (2014). The management of project management: A conceptual framework for project governance. *International Journal of Project Management, 32*(8), 1382–1394.

Trafikanalys. (2012). *Kvalitetssäkring och kostnadskontroll i de nordiske länderna*. Rapport 2012:6. Stockholm, Sweden: Trafikanalys.

Trafikverket. (2014) *Planläggning av vägar och järnvägar*. Version 1.0. Borlänge, Sweden: Trafikverket.

United Nations Development Program (UNDP). (2006). *UNDP and governance: Experiences and lessons learned*. Lessons-Learned Series No. 1. New York, NY: United Nations Development Program.

van Wee, B. (2007). Large infrastructure projects: A review of the quality of demand forecasts and cost estimations. *Environment and Planning B: Planning and Design, 34*, 611–625.

Volden, G. H., & Samset, K. (2015). *Perverse incentives in the front-end: Public funding and counterproductive projects*. IRNOP conference, London, England, 22–24 June 2015.

Volden, G. H., & Samset, K. (2017). Quality assurance in megaproject management: The Norwegian way. In B. Flyvbjerg (Ed.), *The Oxford handbook of megaproject management*. Oxford, UK: Oxford University Press.

Whist, E., & Christensen, T. (2011). *Politisk styring, lokal rasjonalitet og komplekse koalisjoner: Tidligfaseprosessen i store offentlige investeringsprosjekter*. Concept rapport nr. 26. Trondheim, Norway: Concept-programmet.

Williams, T., Klakegg, O. J., Magnussen, O. M., & Glasspool, H. (2010). An investigation of governance frameworks for public projects in Norway and the UK. *International Journal of Project Management, 28*, 40–59.

Williams, T., & Samset, K. (2010). Issues in front-end decision making on projects. *Project Management Journal, 41*, 38–49.

World Bank (1996). *Evaluation results: 1994*. Washington, DC: World Bank.

World Bank (2000). *Reforming public institutions and strengthening governance*. Washington DC: World Bank.

Gro Holst Volden holds the position of Research Director of the Concept Research Program on Front-end Management of Major Investment Projects, at the Norwegian University of Science and Technology in Trondheim, Norway. Her own research is within project governance, public decision processes, and

assessment and evaluation of major public investments. She in an economist from the Norwegian School of Economics, and has a prior career as a senior advisor in the consulting industry and in the government bureaucracy, aiming to increase the value for money of investments, regulations, and other public measures. She can be contacted at gro.holst.volden@ntnu.no

Knut Samset is Professor of Project Management at the Norwegian University of Science and Technology in Trondheim, Norway. He is Director of the Concept Research Program on Front-end Management of Major Investment Projects, and founder and senior partner of Scanteam, an international consultancy based in Oslo, Norway. His academic background is in engineering and social science, and he holds a PhD in risk management. He can be contacted at knut.samset@ntnu.no

Paper VI

Received 14 April 2017
Revised 19 November 2017
Accepted 28 November 2017

The hierarchy of public project governance frameworks

An empirical study of principles and practices in Norwegian ministries and agencies

Gro Holst Volden
Norwegian University of Science and Technology, Trondheim, Norway, and
Bjorn Andersen
Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, Trondheim, Norway

Abstract

Purpose – The purpose of this paper is to study public project governance frameworks in various ministries and agencies in Norway, following the introduction of such a framework on the topmost level (i.e. the cabinet) which applies to the very largest projects.

Design/methodology/approach – The study is methodologically designed as a qualitative assessment of project governance frameworks that apply to state-funded investment projects in selected sectors, based on data gathered through document reviews and interviews.

Findings – The study finds that all of the agencies have introduced their own project governance frameworks, which are basically consistent with the recommendations from the project management literature and with the cabinet's overall requirements in Norway. By contrast, only one ministry has taken a formalized role as a project owner. Governance tasks thus seem to be extensively delegated to the subordinate agencies. This even includes strategic tasks such as project selection and portfolio management, and implies there is a risk that public project governance has a narrow and internal focus.

Originality/value – The paper is a first step toward a better understanding of public project governance as a hierarchical system and the relationship between project owners on three levels, the cabinet, the sectoral ministry, and the government agency.

Keywords Project sponsor, Project governance, Governance framework, Project owner, Public projects

Paper type Research paper

Introduction

Major projects are increasingly used for delivering public goods and services, such as transport infrastructure, defense acquisitions, public buildings, and major ICT projects. They amount to large sums of money and their scale tends to increase (Flyvbjerg, 2014). The McKinsey Global Institute (2013) estimated global infrastructure spending at approximately 4 percent of total global gross domestic product, mainly delivered as large-scale projects.

However, public investment projects face a number of challenges and have varying reputations. There is a wide literature on what Hall (1981) termed "great planning disasters," projects with cost overruns, time delays, and either none or very limited benefits. In one of the most extensive studies to date, Morris and Hough (1987) examined more than 3,500 major public projects across different sectors and countries. They found that cost overruns were typically in the range 40-200 percent and that very few projects were cheaper than budgeted. A number of more recent studies have confirmed that cost

International Journal of Managing Projects in Business
Vol. 11 No. 1, 2018
pp. 174-197
Emerald Publishing Limited
1753-8378
DOI 10.1108/IJMPB-04-2017-0040

overruns are common in infrastructure projects in the UK (MacDonald, 2002), Canada (Berechman and Wu, 2006), Australia (Love *et al.*, 2012), Norway (Odeck, 2004), Sweden (Lundberg *et al.*, 2011), the Netherlands (Cantarelli *et al.*, 2012), Slovenia (Makovšek *et al.*, 2012), and China (Ansar *et al.*, 2016). Flyvbjerg, Skamris Holm and Buhl (2003) studied 258 transport projects from 20 countries over a period of 70 years and found that the problem of cost overruns was consistent and applied to nine out of ten projects. Some of the worst examples of overruns are found in sporting events. The authors of a review of all Olympic Games in the period 1960-2016 concluded that the average cost overrun was 156 percent, and that all games had overruns (Flyvbjerg *et al.*, 2016). By contrast, in ICT projects, Flyvbjerg and Budzier (2011) found that the average overrun was moderate (27 percent), but there was a "fat tail," implying that one out of six projects were "black swans," with an average overrun of 200 percent.

Cost overrun is a widespread phenomenon. However, the more serious, but equally common problem is when projects do not meet the expectations of users and society. In extreme cases, the whole investment could be wasted. Samset (2003) argued that in order to be truly successful, projects must perform well tactically and strategically, not only operationally. Flyvbjerg, Bruzelius and Rothengatter (2003) and van Wee (2007) documented that in transport projects the demand estimates were of equally poor quality as the cost estimates, and that benefits were often overestimated. Pinto (2006) quoted from an *InfoWorld* article describing how 29 percent of ICT projects were paid for but not delivered to the customer, 47 percent were delivered but not used, and 19 percent were abandoned or reworked; only a small share of projects was used as delivered or with minor changes. Flyvbjerg and Budzier (2011) referred to a German study that found that 34 percent of companies undertook projects that were not aligned with corporate strategy, and 67 percent of companies failed to terminate unsuccessful projects.

These problems are not limited to the public sector, as highlighted by, for example, Merrow (2011), who documented similar challenges in the private sector. However, the public sector faces some special challenges, including multiple objectives, difficulties in measuring success, and having to deal with a wide array of external stakeholders in the democratic decision-making processes (Klakegg and Volden, 2017). Public projects are the outcome of a political tug-of-war between stakeholders in society, whose needs and priorities will concur or conflict to varying degrees. The outcomes of such processes are unpredictable, as is well illustrated in a study of 60 international projects (Miller and Lessard, 2000). We could also add that the public sector often has internal challenges, such as a weak capacity for designing a strategic vision, lack of skills, and lack of coordination among levels and actors, as noted by the OECD (2015).

Cantarelli *et al.* (2010) offered four explanations for project failures, each of which may be relevant to varying degrees in specific projects, but generally they may reinforce each other: technical, psychological, economic, and political. The technical explanation is related to the poor project design, incomplete estimation, and lack of tools, methods, and experience. The psychological explanation is based on the concept of planning fallacy and optimism bias, and involves people's cognitive bias and cautious attitude toward risk. The economic explanation has to do with lack of time and resources invested in the planning phase. The political explanation is closely related to the stakeholders' incentives and may cause deliberate manipulation of estimates in order to increase the chance that a specific project will be selected.

Special measures are therefore required to ensure that the right projects are selected and efficiently implemented. In line with the four explanations for project failure, a wide range of measures may be relevant, such as training and improved estimation techniques to avoid technical problems, independent reviews to prevent optimism bias, enough time and resources for early planning, and economic incentives for "true speech" and accountability. A number of organizations, particularly in the private sector, have introduced project governance frameworks in recent years, and the literature on project governance is growing.

Some governments have even established an overarching governance scheme that applies to the largest investment projects across ministries and sectors, to deal with common challenges and to ensure project success. Norway was an early mover in this respect and since the year 2000 the country has required external quality assurance (QA) of decision documents, and that key decisions are elevated to the highest political and administrative level (i.e. the cabinet and parliament). The scheme and preliminary effects, some of which are very encouraging, are presented and discussed in works by Samset *et al.* (2006), Williams *et al.* (2010), and Volden and Samset (2017a).

However, the Norwegian and similar frameworks introduced at the topmost level of the hierarchy apply only to the largest, most complex, and/or highly innovative public projects, and they focus only on the most critical decision gates, while leaving it to the various ministries and agencies to define the more specific governance arrangements necessary to ensure the success of all projects. In this paper, we explore the scope of project governance frameworks on these subordinate levels, including their main components, their comprehensiveness, and their differences and similarities. We are particularly interested in the hierarchy of the project governance frameworks, with the cabinet at top, the sectorial ministries on the next level, and below them the agencies in which the projects are actually implemented.

We start by reviewing the relevant literature concerning public project governance schemes, and conclude that the hierarchy of such frameworks is an underresearched topic. Thereafter, we present our research questions, methodology, and data, before we present and discuss the findings from seven government agencies and seven ministries in Norway. The final chapter offers some conclusions and recommendations, including the need to strengthen the strategic focus of the public project owner's role.

Extant research

Governance in relation to projects

In general terms, governance relates to "all processes of governing, whether undertaken by a government, market or network, whether over a family, tribe, formal or informal organization or territory and whether through the laws, norms, power or language" (Bevir, 2013, p. 1). Governance can thus be studied at different levels and in different fields, such as public governance, corporate governance, global governance, and project governance. A key governance issue is that the interests of the implementing agent will not necessarily be aligned with those of the financing party (Tirole, 2001). Principal-agent theory has been useful to understand this constellation (Eisenhardt, 1989).

The term project governance has only recently become an important issue in the project management community and literature. It refers to the processes, systems, and regulations that the financing party (the project owner) must have in place to ensure that relevant and sustainable project alternatives are chosen and delivered efficiently (Volden and Samset, 2017a). Project governance is thus a system of appropriate checks and balances that enables transparency, accountability, and defined roles in the project, while at the same time supports project managers in delivering their objectives. This corresponds well with what Morris and Geraldi (2011) defined as the institutional level of managing projects, which focuses on shaping the context and conditions to support and foster projects.

When a project is being undertaken by an organization (which is normally the case), a related term is corporate governance, which refers to the mechanisms, processes, and relations by which the corporation is controlled and directed. A much-cited textbook by Müller (2009) defines project governance as a subset of corporate governance, wherein the project governance framework is established to allow projects to achieve organizational objectives and foster implementation that is in the best interests of all stakeholders and the corporation itself. The Project Management Institute (PMI) (2013) defined project governance in a similar way, as "an oversight function that is aligned with the

organization's governance model and that encompasses the project life cycle [by providing] a comprehensive, consistent method of controlling the project and ensuring its success by defining and documenting and communicating reliable, repeatable project practices."

Project governance is a relatively new topic of the project management community, and the literature on the topic is fragmented (Ahola *et al.*, 2014). Different conceptual models have been suggested to categorize the various streams of the literature. Williams *et al.* (2010) distinguished between governance of projects, which aims at efficient delivery, and governance through projects, which aims at choosing the right concepts and ensuring that effects are realized and are sustainable. This corresponds well with Samset's (2003) levels of project success (i.e. operational success defined by efficiency and cost compliance, and tactical and strategic success in terms of impact on users and society).

Müller (2017) made a distinction between project governance on the one hand and governance of projects on the other hand. Project governance refers to the governance of a single project, and includes such topics as the project manager's sovereignty and authority, involvement of various stakeholders, and the use of project boards. By contrast, governance of projects refers to governance of groups of projects within an organization, and includes questions such as the institutionalization of project management methodologies, reporting systems, project selection techniques and program and portfolio management. A similar categorization was made by Too and Weaver (2014) and Ahola *et al.* (2014).

Most of the project governance literature has its origins in the private sector, but the findings and recommendations are also relevant to the public sector. Some studies focus on the governance of state-funded projects at a country level, in relation to political processes and policy forming. Their perspective is on overarching institutional arrangements established by the central governments to ensure that projects succeed across different public organizations (Williams *et al.*, 2010; Klakegg *et al.*, 2015; Volden and Samset, 2017b).

Project governance frameworks

The project governance scheme should be established by the topmost level of the organization, to set the context and the regulatory frameworks within which the projects are implemented. In the following, we will briefly summarize some key findings and recommendations from the literature concerning project governance frameworks and their components. The authors of various studies have emphasized different aspects, depending on their level of analysis, but also on which explanations for project failure they support (cf. Cantarelli *et al.*, 2010). Some authors, such as Flyvbjerg, Bruzelius and Rothengatter (2003), who believe project failure is caused by strategic misrepresentation and irresponsible behavior, highlight economic incentives, accountability, and transparency, whereas others emphasize better tools, techniques, training, and support. In most cases, a combination of measures is recommended. For example, Siemiatycki (2015) discusses remedies to prevent cost overruns and recommends performance monitoring and pay-for-performance contracts as well as better training of project leaders and forecasting techniques that are more precise.

Haanæs *et al.* (2006) reviewed different models for decision making in major public projects based on best practice in Norway and other countries, and suggest the following minimum requirements:

- clearly defined project phases;
- clearly defined decision points between the phases;
- quality-assured basis for the decisions;
- simplicity; and
- a certain standardization and common terminology.

Likewise, Narayanan and DeFillippi (2012) suggest that project governance schemes incorporate five elements:

(1) stage-gate approval processes;

(2) formal roles and responsibilities;

(3) stakeholder representation;

(4) QA; and

(5) contracts and sign-offs.

Certain project phases are more critical and in need of governance arrangements than others. A number of authors have highlighted the crucial role of the front-end phase (Shenhar, 2004; Williams and Samset, 2010; Morris, 2013; Samset and Volden, 2015). This is the stage from when the idea is conceived until a final implementation decision is made, and during which it is still possible to make changes or to terminate the project, at an affordable cost. Many of the factors that later create problems in the construction phase, leading to cost overruns and other problems, are typically present early in the project definition stage (Morris, 2009). Williams and Samset (2010) note that the choice of concept has the largest impact on strategic project success and is thus highly critical. Similarly, Müller (2009) emphasizes that the selection and prioritization of projects is a key issue in a project governance scheme, and that it is closely related to the organization's portfolio management.

A number of standards and guidelines have recently been developed to address project governance models further, in particular as part of corporate governance. For example, the Association for Project Management (APM) (2011) has established 13 principles for the governance of project management and defined four main components of such schemes: portfolio management, project sponsorship, project management capability, and disclosure and reporting.

Such principles and guidelines may be more or less detailed and more or less mandatory. Some project governance models are behavior oriented (i.e. require that certain detailed rules are followed, such as common project management methodology), whereas others are outcome oriented and thus give more autonomy to the project manager. These two "paradigms" may also be denoted as bottom-up and top-down (Müller, 2009). Each organization should establish a project model according to its needs, but some standard models exist. For public organizations, the most commonly used scheme internationally is PRINCE2® (Projects IN Controlled Environments, see www.axelos.com). The scheme was developed in the UK, originally for ICT projects, but has since been developed into a more general framework, integrating the governance of projects, programs, and portfolios and with an associated certification scheme.

Some organizations have established project management offices (PMOs) that often have a central role in a project governance scheme (Hobbs and Aubry, 2008; Morris and Geraldi, 2011; Müller *et al.*, 2014). Other institutions commonly used in the governance of individual projects are the project sponsor, the project board (or steering committee), and various advisory groups and quality committees (Müller *et al.*, 2017).

Independent QA is an important element of a project model too, as it may ensure more realistic estimations of cost, risk, and benefits, as well as a transparent planning process. All four explanations for project failure identified by Cantarelli *et al.* (2010) suggest that independent reviews should mitigate the problem, as they ensure that sufficient competence, experience, and resources are brought in, they provide an outside view, and they provide disincentives to manipulate estimates. Volden and Samset (2017b) studied project models at country level and found that in five of the six schemes there were truly

independent reviews of decision documents at key decision gates – in two countries by external consultants from the private sector and in three cases by a designated government agency.

Flexibility is crucial. Although it may be useful to have a common project governance scheme for all projects in the organization, the scheme should not be static, as the need for governance may vary across projects and project phases (Miller and Hobbs, 2005). Müller *et al.* (2014) identified "organizational enablers" for good governance and governmentality, and their most prevalent finding was the importance of flexibility.

It should be mentioned that although we here focus on governance frameworks, with its formal roles and regulations, there is also a human side of governance. Müller (2017) introduced the term governmentality, which is a combination of "governance" and "mentality," and addresses such aspects as top management's attitudes and ambitions regarding project work, support, and confidence in the project manager, and more generally the cultural values that members of an organization share and respect. The two types of governance, structure-based and relationships-based, will interact and strengthen each other (Klakegg and Meistad, 2014).

The ambiguous project owner role

Project governance and project ownership are closely related, as it is the owner who should be responsible for introducing a project governance scheme to ensure that projects are successful. However, there is much confusion in the literature about this role. In large public projects, the government (i.e. the cabinet) may be seen as the owner, ultimately on behalf of all citizens. Similarly, in private projects, the board of directors is the project owner on behalf of all shareholders (Klakegg and Shannon, 2013). In the next step, the role of project owner may be delegated from the true owners to individuals or groups, so-called "governance agents," according to clear instructions defined by the project governance scheme. In the project management literature, the role of the project sponsor is often highlighted (Helm and Remington, 2005; Kloppenborg *et al.*, 2009). The APM (2011) has defined a long list of responsibilities for sponsors, reflecting the multifaceted nature of the role. The responsibilities include, for example, owning the business case, keeping the project aligned with the organization's strategy and portfolio direction, focusing on realization of benefits, clarifying the decision-making framework, providing resources, supporting the project manager, and engaging other stakeholders.

Project owners as well as sponsors may face a conflict of interest regarding the "governance perspective" and the "support perspective," also referred to as the external and the internal perspective on the project (Ahola *et al.*, 2014; Crawford *et al.*, 2008). On the one hand, project governance should have an external focus, representing the organization and the client's interest, and on the other hand, it should have an internal focus, providing project managers with support so that they fulfill their role efficiently. Olsson and Berg-Johansen (2016) studied seven Norwegian projects and found that the "support perspective" was present in all projects, whereas the strategic and external perspective, focusing on the business case and benefits realization, was more or less absent.

We have already defined the project owner as the financing party. It should be added that the owner is also the ultimate commissioner of the investment, and the one who will control the asset in the operational phase. PRINCE2® distinguishes between three project owner roles, that should all be represented in the project board: the executive, who takes care of the business perspective and provides the funding; the user who establishes whether the project is meeting the needs of the people who will be directly working with the outputs; and the supplier or "do-ability perspective," which provides confidence that the project's outputs will be achieved with available resources. Similarly, Klakegg and Olsson (2010)

distinguish between three strategic owner functions (financing, commissioning, and judicially administering) and five tactical owner functions (controlling, broker/facilitator, planning, executing, and operating).

Governance hierarchies – an under-researched topic

It follows from the above that the literature clearly recommends that a project governance framework should be established by the financing party (i.e. the topmost level of the organization) and be aligned with the organization's strategies and corporate governance model. Several authors have noted that the task is multifaceted, and that different perspectives must be balanced, such as, inter alia, the operational project perspective (governance of projects) and the tactical-strategic perspective (governance through projects), as well as the control and the support perspectives. Some have suggested that different governance functions be filled by different individuals. However, this discussion should also be related to the levels of the hierarchy and the allocation of project owner responsibilities optimally across the levels of an organization.

As noted by Too and Weaver (2014), a governance framework is hierarchical in the same way as a management system, where the top level is accountable for the whole system, but delegates responsibility and authority for defined actions to subordinate levels. One out of very few studies that have explored this topic is Zwikael and Smyrk (2012), who showed that there are principal-agent relationships at multiple levels of the organization, with the funder (the true owner) on top, who hires a project owner to be accountable for benefits realization, who in turn hires a project manager to be accountable for efficient output. Each level must evaluate the performance of the level below, and for this task they need the right success criteria, governance arrangements, and authority to make decisions. All these elements should be determined by the organization's project governance framework. In most cases, we would expect that the higher levels of the hierarchy should control the measures to ensure tactical-strategic success, whereas the lower levels focus on operational success. Similarly, that the higher levels place more weight on control and the lower levels on support. And that the higher levels govern with respect to outcome, and the lower levels with respect to behavior. But again, our knowledge is limited since there is a gap in the literature concerning these issues.

Furthermore, the above-mentioned authors discuss projects that take place within the framework of a single organization. We have not identified any studies that explain the allocation of project governance responsibilities across different organizations in the way it occurs in government-funded projects. The ultimate owner of a project funded by the national government is the cabinet, led by the prime minister, who is de facto responsible for all projects under the various ministries. However, this responsibility is normally delegated to the sectoral ministry, and in turn to the relevant subordinate agency where the project is implemented. Each level may want to introduce their own project governance scheme which is aligned with their strategies, project portfolio, competence level, etc. An interesting question is whether these project governance frameworks on various levels are consistent with each other. This is the topic that we seek to explore in this paper, and we find Norway to be a suitable case since a project governance scheme introduced at the topmost level (the cabinet) is already in place.

This study and the Norwegian public project context

The model which forms the basis for our research is shown in Figure 1. As illustrated, there are project owners on three levels of the government hierarchy: the cabinet, the responsible ministry, and the agency. Certainly, it is the parliament, and ultimately the people as voters and taxpayers, who are the real owners of public projects. However, here we limit our attention to the executing power, which, in a parliamentary system,

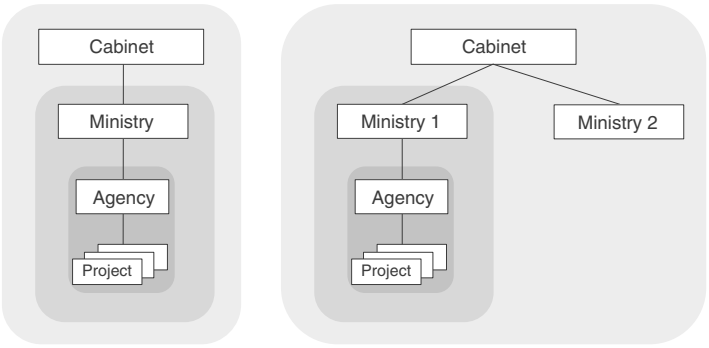


Figure 1.
The various levels of project ownership in public projects

emanates from the parliament. Furthermore, we focus only on the state and not on local and regional authorities.

Each gray-colored field defines an organization, with its own goals, strategies, types of projects and contextual factors, and associated project governance framework. All the three levels can be said to have ownership in, at least, the largest projects executed by the agencies. In this paper, we investigate to what extent the various levels actually exert project ownership in terms of introducing a stage-gate model or other governance arrangements, and engaging in key decisions concerning projects, programs, or portfolios.

There is a principal-agent relationship between each level. The ministry's goals and strategies will normally be broader than those of the agency. For example, whereas a public roads agency may wish to build as many highways as possible, with the most fancy and high-tech (and expensive) bridges, subsea tunnels, and so forth, the Ministry of Transport is responsible for all modes of transport, and should balance the need for mobility against life-cycle cost, the environment, and other concerns. A ministry cannot and should not be involved in all individual projects executed by the agencies, but it should require that the most critical decisions are elevated to ministry level, and it should ensure that the agencies have the necessary competence, capacity, procedures, and processes. At top of the hierarchy is the cabinet. The overall allocation of the government budget among the ministries should of course be determined on this level. Major public projects may have impacts that extend beyond a single ministry's field of responsibility. Certain project decisions may therefore be so important, or involve such a high level of conflict or risk that they should be elevated to the topmost level of the governmental system.

The picture may be more complex than the "pure" structure shown on the left-hand side of Figure 1, with only one ministry and one agency involved. In reality, many different variants exist, where various other ministries and/or agencies are among the most important stakeholders in the project (e.g. in the role of user or regulator). We would like to highlight one model in particular, as shown on the right-hand side of Figure 1, where two different ministries are shown as having ownership interest in the project: one ministry (Ministry 1) is formally responsible for the executing agency (providing the general mandate to operate its business), and another ministry (Ministry 2), which commissions the particular project. Ministry 2 will then be expected to take more of a user/customer perspective. As we will come back to, this is the case with building construction projects in Norway, and it implies a strong need to coordinate the exertion of ownership.

In Norway, an overarching project governance framework was established by the cabinet in the year 2000, and extended in 2005. This scheme represents the "top layer" of the project governance hierarchy. It applies only to the projects exceeding an expected cost of NOK750 million, and comprises only two decision points in the front-end of projects.

The decision documents to be produced before these two decision gates must have a certain content and they must be subject to an external quality assurance (QA):

(1) QA1 – QA of choice of concept before cabinet decision to start detailed planning. The basis for the QA1 exercise should be a needs analysis, the project goals, and overall requirements, a possibility study and a cost-benefit analysis of at least three alternative solutions. The reviewer should give recommendations regarding the ranking of alternatives and the decision strategy.

(2) QA2 – QA of the management base and cost estimates before the project is submitted to the parliament for approval and funding. The cost estimate must be based on stochastic estimation techniques, where two figures should be calculated, the P50 and the P85 (*Px* implies that there is *x* percent probability that the actual cost will be at or below this level, given the uncertainty in the project). The recommended budgeted cost should be set at or close to P85 whereas the target cost for the agency should be around P50. The difference between the two numbers is the uncertainty provision. On the portfolio level, it should be expected that the projects hit P50 on average so that no provision is spent.

The Ministry of Finance has entered into framework agreements with private consultants who perform the QAs. The final decision is of course a political one. The scheme is shown in Figure 2.

The QA scheme ensures that the project is well prepared and at a sufficient level of maturity when it is approved, and it ensures legitimacy for the final decisions. QA1 concerns the choice of concept and thus the tactical-strategic level of project success, whereas QA2 is intended at efficient project implementation (i.e. operational success). Volden and Samset (2017a) present and discuss the Norwegian scheme and preliminary effects. Generally, it adheres to the recommendations from the literature concerning the components of a project governance framework (stage-gate approval process, clearly defined responsibilities, QA, etc.). It is also a transparent system, as all the QA reports are made publicly available.

Other countries have introduced similar schemes in recent years. In the UK, the Cabinet Office and HM Treasury introduced a similar scheme in the year 2000, and other countries have followed in the subsequent years. The authors of several studies (Williams *et al.*, 2010; Klakegg *et al.*, 2015; Volden and Samset, 2017b) compare the Norwegian and the UK governance frameworks with each other and with those of other countries. Generally, they conclude that it is too early to determine the schemes' effect on project success, but that there is much to suggest that it has been positive. The schemes have much in common, and those that were first to be implemented have inspired the followers. There are a number of differences between the schemes: Volden and Samset (2017b) conclude that there are two main types of schemes: the schemes in the Scandinavian countries, which are simple in terms of the number of intervention points and which do not intervene significantly in the existing processes and practices in the agencies; and the schemes of the Anglo-American

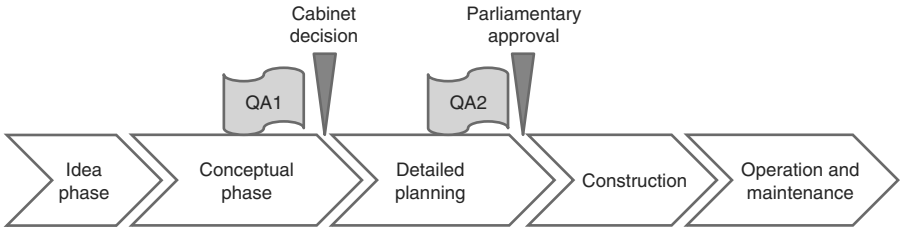


Figure 2.
The Norwegian quality assurance (QA) scheme – the "top layer" of project governance

countries, which are more extensive and behavior oriented, with follow-up points also in the implementation phase, and which exert more centralized control.

The Norwegian QA scheme applies only to the largest projects – approximately 20-30 projects each year. It applies only to the front-end phase, with no instructions concerning the implementation phase. It should also be noted that the QA scheme does not include a decision point between the idea phase and the conceptual phase. Furthermore, there are no instructions regarding project organization (such as the use of governance agents), nor about portfolio management. It is therefore implicitly assumed that the scheme is supplemented by more specific governance arrangements in ministries and agencies. By demanding high quality from the top of the pyramid, a trickle-down effect should be expected, in the form of higher standards at the lower levels. There should also be consistency between the three project governance schemes, in the sense that schemes on the lower levels include the requirements on the higher levels, and specify, refine, and adapt them to the specific sector or project type, to the extent necessary.

For the empirical part of this study, we raise the following research questions:

RQ1. Have Norwegian ministries and agencies introduced project governance frameworks, following the introduction of such a framework at the topmost level? If so, describe the content of the schemes, including any differences and similarities between sectors.

RQ2. Which level, agency or ministry, takes the most active role as project owner and initiator of governance arrangements?

RQ3. Are these frameworks consistent with recommendations from the project governance literature?

RQ4. Are the schemes in the hierarchy (cabinet, Ministry, agency) internally consistent?

RQ5. What are the theoretical and/or practical implications of the findings?

Data and methodology

In this study, we investigate project governance arrangements on the ministry and agency levels in seven government agencies and seven ministries in Norway. The study is methodologically designed as a qualitative, case-based assessment of project governance frameworks that apply to the state-funded investment projects in the selected sectors, based on data gathered through document reviews and interviews. We consider this to be a well-suited approach as the study has a descriptive and exploratory purpose rather than to draw universal conclusions. In line with Yin (2013), we believe that the concrete, context-dependent knowledge that can be obtained from case studies is highly valuable and that precisely because of the detailed understanding that is obtained, the results are often relevant to other contexts. We consider this an important step toward a better understanding of how public projects are governed and how project governance arrangements on lower levels underpin and reinforce the QA scheme on the topmost level. This knowledge should also be relevant to other countries with similar project governance frameworks on the cabinet level.

The included agencies and ministries are shown in Table I. The seven agencies were selected because they have had the most projects covered by the QA scheme. Samset and Volden (2013) documented that the projects that undergo external QA are divided as follows: 43 percent road projects, 9 percent rail projects, 14 percent defense projects (material projects and construction), 11 percent ICT projects, 18 percent civil building construction projects, and 5 percent others. All of the agencies are largely project-based organizations. Total investment volumes per year[1] in the sectors follow from Table I.

Type of project	Investment volume (mill. NOK, 2015)	Name of government agency	Responsible ministry	Client ministry and subordinate agency
Road projects	12,208	National Public Roads Administration	Ministry of Transport	
Railway projects	7,158	National Rail Administration	Ministry of Transport	
Defense material projects	10,211	Norwegian Defence Material Agency	Ministry of Defence	
Defense building construction projects	1,605	Norwegian Defence Estates Agency	Ministry of Defence	
ICT infrastructure projects in the labor and welfare sector	260	Norwegian Labor and Welfare Administration (NAV)	Ministry of Labour and Social Affairs	
ICT infrastructure projects in the Police	n/a	Police Directorate	Ministry of Justice	
Civil building construction projects	2,577	Statsbygg	Ministry of Local Government and Modernization	Examples Ministry of Culture/ Museum Ministry of Justice/ Prison Ministry of Education/ University

Table I.
Public organizations included in the study

Note that three of the ministries are included twice because they are responsible for more than one agency. Further, a special comment regarding the ownership of civil building construction projects is required (cf. the right-hand side of Figure 1). Statsbygg is a government agency specializing in providing facilities for various public institutions. The ministry responsible for providing the financial resources for Statsbygg is the Ministry of Local Government and Modernization. However, there is typically a second ministry involved, namely the one that owns the institution that will use the building (e.g. in the case of a prison, the Ministry of Justice). Hereafter, we refer to this ministry as the "client ministry."

Project governance arrangements can be more or less formalized. For example, a high level of trust and/or frequent communication with the subordinate level may reduce the need for formal requirements. However, as these projects take place within the state bureaucracy, we expect a certain degree of formalization. We therefore focus primarily on the structural aspects of project governance.

The empirical investigation covers the following main topics, which are extracted from literature and underpin the research questions. The list largely corresponds to our interview guide:

- Stage-gate models: have schemes or models defining project phases and decision points, been introduced for projects implemented by the agencies? If so, what characterizes them in terms of, for example, their comprehensiveness or the phases they cover?
- Ministry involvement: what level of involvement have the ministries chosen for projects implemented by their subordinate agencies?
- Roles in governance: are formalized governance agents appointed, such as project sponsors or project boards? Who fill these roles and what are their mandates?

- QA: to what extent is independent QA of decision documents an integrated part of the schemes?
- Threshold: which criteria (threshold level or other) are used to determine when decisions are to be elevated to a higher level and/or specific requirements must be adhered to? How flexible are such requirements in response to individual projects' needs and properties?
- Uncertainty provision: who controls the provision for uncertainty in the project budget? Do smaller projects not covered by the QA scheme also have project budgets that include a provision for uncertainty?
- Portfolio management: to what extent is project portfolio management an integrated part of the scheme, at each level?

This study builds on document reviews and semi-structured interviews with key interviewees from all of the included ministries and agencies (Table I), a total of 31 people. A challenge during data collection, and a finding in itself, has been the great variability in the availability of written descriptions of the project governance schemes. The Ministry of Defense stands out, with its' comprehensive descriptions of its project governance scheme Prinsix, including templates and guidelines, which are openly available on the Ministry's website. Other institutions have provided more or less detailed documentation at our request. Yet others have provided little information, mainly because such written documentation does not exist. This was particularly the case for the ministries' involvement in the projects and for governance arrangements in the earliest project phases. In these cases, data collection was based on oral sources. Generally, the document review was conducted before the interviews, and the interviewees were asked to verify our understanding of the model and explain some of its elements further when necessary. In few cases, it was an iterative process where we were given access to new documents after the interviews, in which case we were able to contact the interviewees again for subsequent follow-up questions.

All the ministries and agencies with projects that undergo the cabinet's QA scheme has appointed a contact person, who is available for questions about the quality assured projects and processes. We have access to this list and used it as a basis for contacting the ministries and agencies, and the contact persons helped us identify interviewees. We were looking for interviewees who were highly experienced in project work in general, and had high knowledge of the ministry's/agency's project governance arrangements, in addition to having special knowledge of their sectors. Some of the contact persons were themselves among the interviewees. The interviews were in most cases conducted with individuals, but in a few cases more than one person from the same agency or ministry was present simultaneously.

As indicated, the interviews were open and semi-structured, and based on the list of topics presented above. We used the stage-gate model as a starting point and asked the interviewee(s) to explain the life-cycle of a typical project in their sector, from the first idea arises until the operational and maintenance phase. We also asked questions like "what would happen […]" and "who would react […]" when something does not go according to plan. The interviewees were encouraged to talk freely, based on their own personal experiences and knowledge of the various topics. Each interview lasted one to two hours, with one or two researchers present. The researchers prepared a comprehensive interview report following each interview.

Data were collected between February and December 2016. Data from different sources were subsequently compiled and systematized by the researchers, topic by topic. Since the objects of study were few and the topics covered fairly broad, we did not use any formal

content analysis, coding or other quantitative or qualitative analysis software, neither for the document analysis not the interview analysis. We simply constructed a large table with the type of project (i.e. the sectors) along the *x*-axis and the selected topics along the *y*-axis, with comprehensive descriptions in the cells. We placed great emphasis on ensuring comparability across sectors, also in cases where varying terms were used.

A challenge when using interviews as a data collection method is that the information is inevitably affected by the interviewees' interpretations and subjective assessments. In our case, there was a potential risk that some of the interviewees might have had a self-interest in portraying their own efforts, competence, and project practice in a good light. We therefore emphasized triangulation of the information obtained from different sources (written documentation vs oral sources, and ministry vs agency). In a few cases, we discovered information that we perceived as inconsistent, and then had to go back to the interviewees and/or check a third source, to clarify the issues.

More generally, case studies are often considered "weak evidence" and biased toward verification. But as noted by Flyvbjerg (2006), the question of bias applies to all methods, including, for example, the choices of categories and variables in a quantitative study and the structure of a questionnaire. Experience indicates that case studies actually contain a greater bias toward falsification of preconceived notions. Our study, although not intended to test a set of hypotheses about cause and effect, rested implicitly on a set of assumptions, and we tried to be open to the fact that they might not hold. For example, even though we assumed that the topics drawn from the extant literature were the most important ones in describing these particular governance frameworks, we also searched actively for other aspects. Similarly, we tried in various ways to questions the assumption that any improvements in ministries and agencies' governance schemes after the year 2000 can be traced back to the introduction of the cabinet scheme.

Presentation and discussion of findings

In this section, we present and discuss the most important findings from our study. The presentation is basically structured according to the predefined list of topics from the data and methodology section, but more interesting and/or surprising findings are highlighted.

Stage-gate models

A key finding in the study was that all or most of the agencies seem to have invested heavily in their project competence and capacity in recent years. All of the agencies in the study have introduced formalized stage-gate project models and many of them update and improve their models regularly to ensure that they are consistent with best practice. Most of the project models were introduced during the last 15 years, some even more recent, and our interviewees believed that the QA scheme introduced in year 2000 on the topmost level had been a major trigger. "The OA scheme taught us which requirements we should ask in the planning of projects," said one agency interviewee. All of the models are well adapted to the QA scheme and ensure that the largest projects are well prepared for the two control points. The models also seem to have become increasingly comprehensive over time, with associated guidelines, templates, and procedures, and some interviewees were of the view that future improvements should be in the form of simplifications.

The two agencies with many ICT projects have both introduced a variant of PRINCE2® in recent years. The other agencies have introduced similar models, but without referring to any particular international model. Generally, there are many similarities between the models. They have between four and six project phases, with decision points between each phase. The names of the phases are fairly similar (but not identical) across agencies, and cover the conceptual phase, the detailed planning phase, and the construction phase,

as a minimum. By contrast, the idea phase is included only in one of the project models. The two included agencies with ICT projects stand out for highlighting benefits realization (i.e. the operations and maintenance phase) in their project models.

The models are, with one exception, introduced and managed on the agency level. Only one of the ministries, the Ministry of Defense, has established its own formalized stage-gate model, which applies to defense material projects and is consistent with what happens at the agency level. The interviewees from the ministries generally referred to the QA scheme when asked about project models. The ministries see it as their responsibility to ensure that the subordinate agencies prepare their largest projects according to the cabinet's requirements, but otherwise there are few regulations from the ministries' side. Table II shows the project models by the project type.

The earliest phase

As mentioned above, the project models generally do not cover the earliest phase, where the idea occurs and is developed into a conceptual solution. Some of the interviewees referred to this phase as "the political or strategic phase, which is not a part of the project." Only the Ministry of Defense has introduced a clear instruction as to how project ideas should be treated and who can approve an idea before transferring it to the conceptual phase. "We always approve the idea before any start-up activities on agency level can be initiated." The Ministry of Defense is clearly responsible for this phase, in close cooperation with the agency.

According to our interviewees, the project ideas occur in various ways. "It is not always clear where they come from," said an interviewee. In some cases, the idea is identified on the political level, or follows from a new policy or reform. An example provided by one of the respondents was the Pension reform in 2011, which implied a need for a renewal of the ICT systems in the Labor and Welfare Administration. Equally often, however, the idea "occurs" on the agency level, sometimes in close dialogue with internal or external user groups, and based on more or less systematic needs assessments. For example, the National Public Roads Administration has five regional departments, each of which is in close contact with municipalities and other stakeholders, and "picks up" user needs in various informal ways. In the case of civil building construction projects, the idea will often arise at the user agency level (e.g. a prison or a museum), which may start to explore alternatives, sometimes in

					Phases covered by the model				
Type of project	Stage-gate model	Owner of model	Number of phases	Number of decision points	Idea phase	Conceptual	Detailed planning	Construction	O&M
Road projects	Yes	Agency	6			X	X	X	X
Railway projects	Yes	Agency	6	6		X	X	X	
Defence material projects	Yes	Ministry/ agency	6	4	X	X	X	X	X
Buildings/locations for the defense	Yes	Agency	4	3		X	X	X	
ICT infrastructure projects in the health and welfare area	Yes	Agency	5	6		X	X	X	X
ICT infrastructure projects in the Police	Yes	Agency	5	5		X	X	X	X
Buildings/locations for civil government institutions	Yes	Agency	6	6		X	X	X	

Note: "X" indicates that a phase is covered by the stage-gate model

Table II.
Stage-gate models by the type of project

consultation with Statsbygg (Norwegian Directorate of Public Construction and Property), and eventually the ministry will become involved and the building will be formally commissioned.

The QA scheme is meant to ensure that the choice of concept is elevated to the cabinet level in the largest projects, but not even the QA scheme covers the earliest decision to develop a project idea. It is a well-known critique against the QA scheme that projects sometimes have developed too far when they reach the cabinet after QA1, by which time it is difficult to stop them.

Ministry involvement

The ministries are formally responsible for all projects implemented by subordinate agencies. For the largest projects, the ministries formally submit the decision documents to QA1 and thereafter to the cabinet for approval. However, for smaller projects and for all other project phases, the ministries' formal involvement is limited in most cases.

The Ministry of Defense stands out for being strongly involved in the projects in all phases, formally as well as informally, particularly in the front end. The Ministry defines itself as the project owner of all projects in subordinate agencies, and it designates a person in the Ministry to act as project sponsor, regardless of project size and complexity. Projects exceeding a certain threshold have to be elevated to the minister (i.e. the political level in the Ministry) for approval. An agency interviewee was not always satisfied with the extensive involvement. "The Ministry is quite detail-oriented, and always tells us which form to fill in," the person said. But also, "if the documents are of good quality they normally listen to us."

The other ministries do not use formal project models or assign themselves formal roles in relation to the projects. They may however govern projects in informal ways to the extent that they find it necessary, typically depending on scope, complexity, and political risk. The Ministry of Transport, as owner of large, project-based agencies within road and rail, sees no reason to get involved in individual projects "as long as everything goes according to plan and the project is not politically critical" as one interviewee puts it. Often, the subordinate agencies drive the process, even in the early phases. In civil building construction projects, the client ministry takes responsibility for the conceptual solution, but leaves the detailed planning to the implementing agency, Statsbygg. This is somewhat surprising, given that many strategic decisions and clarifications are needed in this phase as well. Some of our interviewees from the client ministries expressed the view that they, and their user agency, should have been more involved. But they find it difficult in the face of Statsbygg as the professional government construction agency, who "asks for our opinion only when its stage-gate model says so", as one interviewee puts it. The sponsoring ministry on the other hand, gets involved once the building is formally commissioned. Thereafter, the sponsoring ministry follows the projects, individually and as a portfolio, through the implementation phase, mainly with a cost control perspective.

The two ministries responsible for ICT projects have both appointed senior experts in the ministry to monitor the largest projects closely. "After all, the minister is the ultimate responsible, and he/she needs to know what is going on", said a ministry interviewee. This is not surprising, as these agencies have experienced serious problems with some ICT projects in the recent past. And also, to quote an interviewee, "because ICT projects change so rapidly that we may have to change the conceptual solution several times along the way." However, rather than taking a formal project sponsor role, the ministries rest on informal meetings, and the purpose of such activities is to obtain information early. The ministries may request information at a fairly detailed level if they find it necessary, and they sometimes do. If serious problems arise, they will be addressed in the regular meetings between the ministry and the agency's director general.

The project sponsor role

Our study confirms that the project sponsor role is commonly used in Norwegian public projects. All of the project models included in this study highlight this role. The responsibilities, tasks, and competencies required for the role are defined by the models. They normally state that the project sponsor has ultimate responsibility for the project, is the leader of the project board (if used), and the one who defines the project goals, appoints the project manager, initiates QA, and makes decisions beyond the projects managers' defined authority. However, our data also show that in many cases the sponsor is positioned at a low level in the hierarchy, and thus cannot be expected to take the strategic and high-order perspective on the project that he or she should. Again, the Ministry of Defense stands out in requiring that the project sponsor is located in the ministry, with an additional "local" sponsor in the agency. In all the other sectors, the project sponsor is located only at the agency level. Most of the project models require the role to be filled by a senior manager or even someone from top management. However, as noted by several interviewees, "it is a big challenge that top managers do not have the time and capacity to follow the projects sufficiently closely." Therefore, in practice, the role is habitually delegated further down in the organization, to a subordinate manager (typically the project manager's supervisor). Two agencies emphasized that the project sponsor should be someone who is also responsible for benefits realization (e.g. the department that will reap the benefits from an ICT project).

Another observation is that the project sponsor is often appointed rather late in the project life-cycle, after the project has been formally approved for funding, while having no role in the front-end phase, and sometimes not in the operational and maintenance phase either. In some agencies, the project sponsor role is transferred from one person to another as the project enters new project phases and the responsibility changes from one department to another.

Generally, our findings indicate that there is a risk that the governance tasks will be handled from an internal perspective, rather than actually representing user groups and the broader society. This is in line with Olsson and Berg-Johansen (2016) who distinguish between project owner type 1, with an external perspective, in line with recommendations in the literature, and project owner type 2, with an internal perspective, which is what they observed in practice.

Project boards

There is an extensive use of project boards in the agencies. Most of the project models recommend or require the use of boards for projects that are large, complex, or have interfaces with other agencies or key stakeholders, in which case these stakeholders should be represented.

However, an interesting finding is that many of these institutions bear more resemblance to advisory groups and project reference groups than to real steering groups. They often have a large number of board members, including user groups and various other internal and external stakeholders, who are there to obtain information and give advice, but do not necessarily, have a mandate to make decisions on behalf of their organizations. To quote one interviewee, "we tend to include the whole list of stakeholders, so that we do not need a supplementary consultation process." Another interviewee said "unfortunately, few decisions are made in these meetings. It is sometimes more like a tea party."

The ministries are normally not involved in such project boards at all. In civil building construction projects, an external advisory committee is sometimes used, on which the sponsoring ministry and the client ministry are both represented. Such committees do not make binding decisions; only recommendations. Traditionally they have been established after the project has been approved, to follow up during project implementation, but the

trend is now that they are established at an earlier stage. According to our interviewees, experiences are mainly positive, and the committees are found to strengthen the client perspective in the projects.

QA

As described above, the overarching project governance scheme on the top level, with its two mandatory QA reviews, applies to all the largest projects, independent of sector and agency – about 20-30 projects per year.

On the ministry level, there are few additional requirements for QA. The exception is, again, the Ministry of Defense, which uses "red teams" on certain high-risk projects. Furthermore, some ministries routinely consider whether the cabinet's QA scheme should be applied to projects just below the QA threshold. This has happened in several cases. As noted by one of the ministry interviewees, "the need for QA does not only depend on size, but a broader set of criteria." For example, ICT projects may be smaller than the other project types in terms of monetary values, but their complexity is often considerably higher.

On the agency level, all project models have requirements in place concerning independent QA. The scope and content of such reviews, and the extent to which they involve truly external experts, vary significantly. In most cases, the QA is performed internally in the agency by people who are independent of the particular project.

Uncertainty provisions and change management

Project budgets may include a provision for uncertainty. As noted, for the largest projects (covered by the cabinet's scheme), the cost estimate must be based on stochastic estimation techniques, and the budgeted cost is set at or close to P85 while the target cost for the agency is set at P50. Our data show that probability-based cost estimation has spread to smaller projects as well. All the studied projects models include requirements regarding cost estimation and uncertainty provisions, applying not only to QA projects, with budgeted costs and target costs expressed in terms of probabilities, i.e. Px.

There is assumed as an inherent incentive for the implementing party to increase scope and quality and/or to add some slack to the project implementation. Therefore, an effective way to avoid unnecessary spending may be to retain most or all of the provision at a higher level in the project hierarchy. However, our findings show that most of the agencies are given authority to spend the budgeted cost, even for the largest projects. Only three ministries delegate a lower target cost to their subordinate agencies, which must apply to their ministry to spend the provision. In the case of civil building construction projects, the external advisory committee will normally discuss the need to spend the uncertainty provision before it is approved by the sponsoring ministry.

On the agency level, the project models define how the provision is delegated further to lower levels in the project hierarchy. The target cost for the lowest level, the project manager, is generally between P35 and P45. However, there are considerable variations in how the decision hierarchy is defined and how the mandates are given, not least to what extent the project sponsor and/or board is given authority to make decisions concerning the provision.

Portfolio management

Volden and Samset (2017a) note that the Norwegian QA scheme focuses primarily on governance of individual projects and does not include explicit portfolio evaluation requirements. This is in contrast to, for example, the UK scheme, in which QA is required not only for individual projects but also at the program and portfolio levels at regular intervals.

Our study confirms that portfolio management is not considered an important issue in the ministries. An exception is, again, the Ministry of Defense, which takes an active role, not least in the project selection phase. The ministry regularly updates and manages its long-term investment plan for the sector and new projects are assessed against this project portfolio. The two ministries with many ICT projects have also been concerned with project portfolio management in recent years. However, instead of taking care of portfolio management themselves, the ministries require that high-quality portfolio management takes place in the agencies. In the transport sector, a national transport plan is updated by the ministry every four years, and includes all modes of transport (i.e. road, rail, sea, and air). However, the way this has been practiced so far is that the national transport plan is the sum of four independent portfolios, governed by the four agencies, rather than being managed as one holistic plan for the whole transport sector.

By contrast, the portfolio perspective is prominent in all the agencies. The agencies manage their respective investment plans, they compile data on progress and performance for the portfolio, and most of them have some flexibility to re-allocate funds between projects if necessary. However, the extent to which portfolio management is an explicit part of the project model varies. The agencies with many ICT projects stand out as rather advanced. No agencies have established PMOs with portfolio management responsibilities.

Key characteristics of the project governance arrangements in the various sectors are summarized in Table III.

Assessments and conclusions

Public investment projects in Norway and worldwide have traditionally been burdened with problems, to the extent that Flyvbjerg (2014) proposed an iron law: "over budget, over time, over and over again." There is no easy solution to these problems. In public sector projects, people rarely put their own money at stake, goals are often multifaceted, goal achievement is not easily measurable, and there are many stakeholders within and beyond central government who have opinions on the project. Over time, we have gained a better understanding of the challenges that arise in these political administrative processes, but we still have a long way to go before we fully know how to solve them. What we do know from the existing literature is that having a project governance framework in place will at least make the processes more predictable, and hopefully will result in a decision basis of higher quality, and more legitimate decisions. The Norwegian Government introduced a QA scheme in the year 2000. The scheme ensures that the basis for two key decisions in the front-end of the largest projects is quality assured by truly external experts and that decisions are elevated to the highest level in the political system. The scheme is very simple, applies only to the very largest projects and constitutes the "top layer" of a hierarchy of governance arrangements. Clearly, the success of public projects depends heavily on what happens in the ministries and agencies in which projects are actually implemented. The hierarchy of governance that surrounds public projects has not been thoroughly discussed in the literature. In this paper, we have explored this topic and provided some empirical findings in a Norwegian context.

The first research questioned posed earlier in this paper was to what extent Norwegian ministries and agencies have introduced project governance frameworks. The findings clearly show that the ministries and agencies in the studied sectors have invested heavily in their project competence and capacity since the turn of the millennium. All the agencies (and one ministry) have introduced a comprehensive project model with, for example, clearly defined phases, decision points, roles, responsibilities, and QA. There are more similarities than differences between the models, in terms of phases, decision points, roles, and requirements.

These findings indicate that Norwegian public projects are fairly well governed, first and foremost on the agency level. In light of the general "projectification" of society, it is difficult to

Table III.
Characteristics of project governance schemes by type of project

Type of project	Ministry role, overall	Agency role, overall	Project sponsor	Project board	Quality assurance additional to requirements in cabinet's QA scheme	Control of uncertainty provision in the budget (normally P85)	Target costs	Portfolio management
Road projects	Limited, informal	Active, owner of stage-gate model	Yes, located in the agency, normally delegated	Not often used. Advisory role. Ministry not involved	QA by agency	Agency	P50 for the regional department, normally P45 for project manager	To some extent, in the agency
Railway projects	Limited, informal	Active, owner of stage-gate model	Yes, located in the agency, normally delegated	Often used. Advisory role. Ministry not involved	QA by agency	Agency	P50 for the project sponsor, lower for the project manager	To some extent, in the agency
Defence material projects	Active, formal, owner of stage-gate model	Active	Yes, located in the ministry, and a local sponsor in the agency	Not often used. Advisory role. Ministry not involved	Cabinet requirements apply to smaller projects as well. QA by ministry and agency	Ministry	P50 for the agency, lower for project manager	Yes, in both the ministry and the agency
Buildings/ locations for the defense	Active, formal	Active, owner of stage-gate model	Yes, located in the ministry, and a local sponsor in the agency	Often used. Advisory role. Ministry not involved	Cabinet requirements apply to smaller projects as well. QA only ad hoc, by agency	Agency (ministry if > 100 mill. NOK)	P50 for the agency, lower for project manager	Yes, in both the ministry and the agency
ICT infrastructure projects in the health and welfare area	Active, informal	Active, owner of stage-gate model	Yes, located in the agency	Often used. Advisory role. Ministry not involved	QA by agency	Ministry	P50 for the agency, may be lower for project manager	Yes, in the agency, by requirement from ministry
ICT infrastructure projects in the Police	Active, informal	Active, owner of stage-gate model	Yes, located in the agency, normally delegated	Often used, supposed to have a steering function but in practice advisory. Ministry not involved	Cabinet requirements apply to smaller projects as well. QA by agency	Agency	P50 for the steering committee and P35 for the project manager	Yes, in the agency, by requirement from ministry
Buildings/ locations for civil government institutions	Limited, informal	Active, owner of stage-gate model	Yes, located in the agency, normally delegated	Only advisory committees are used, including sponsoring and client ministries	QA by agency	Ministry (agency may spend 20% of the uncertainty provision)	P50 for the agency, lower for project manager	To a limited extent. Client ministries have few projects each

determine whether the improvements on the lower levels are caused by the cabinet's QA scheme, but at least it seems to be an important triggering factor. Clearly, the threat of the project being critiqued by external quality assurers, and possibly rejected by the cabinet on their advice, provides an effective incentive for the agencies to work hard with the decision documents. Our study also confirms that the ministries relate strongly to the QA scheme and see themselves as responsible for ensuring that the requirements are met in the largest projects.

The second research question was which level, agency or ministry, takes the most active role as project owner and initiator of governance arrangements. Our findings show that with one exception (the Ministry of Defense), it is the agency level that takes the most active role as project owner and initiator of governance arrangements. Admittedly, the ministries see themselves as owners of projects implemented by their subordinate agencies, and may be involved in various informal ways. The degree to which they are, seems to depend on, inter alia, project size, political risk, and the agency's experience and track record regarding project delivery.

The third research question was whether the governance frameworks adhere to the recommendations from the project governance literature. Our findings show that they largely do. They include stage-gate models with clearly defined phases, decision points, roles and responsibilities, and QA of decision documents. We are also quite impressed to see that the agencies regularly use stochastic cost estimation techniques as the basis for determining budgets and target costs, even for smaller projects. However, we also see room for improvements in some areas. Generally, the project owner role should be executed more actively and with a focus that is more strategic (the exception being the defense sector). The project sponsor is mostly located at the agency level and is often an individual at a fairly low level in the organization. Furthermore, in most cases, there is no role for the project sponsor in the crucial front-end phase, where project ideas arise and are selected for development into a concept. Thus, our findings seem to support those made by Olsson and Berg-Johansen (2016), who only observed the more narrowly oriented "project owner type 2" in Norwegian public projects. Furthermore, project boards are widely used, but more often than not they are not truly boards, but rather advisory groups, and they are normally established late in the project life-cycle, after the crucial choice of concept has been made.

The fourth research question was whether the schemes in the hierarchy (cabinet, ministry, agency) are internally consistent. Overall, the answer is yes. We have not found any obvious inconsistencies between governance arrangements on various levels in the hierarchy. Particularly, the schemes on the lower levels are well adapted to the cabinet's scheme and ensure that the largest projects are well prepared for the two control points. But again, the strategic and external perspective on project governance, which should be taken by the ministry level, is often missing or handled very informally.

The final research question was about the implications of our findings. Overall, it is our view that these project models have a somewhat narrow and internal focus, securing governance *of* projects, but not necessarily governance *through* projects, in Williams *et al.*'s (2010) terminology. At best, they ensure governance through projects in the agency's perspective, but the impact of the project normally goes beyond the agency. As discussed by Klakegg and Volden (2017), Norway has a strong democratic tradition, an egalitarian culture, and a high level of education, which makes a strong platform for organizing tasks as projects and for delegating authority downwards in the hierarchy. The government's introduction of the QA scheme was controversial in the beginning, and the introduction of formalized regulations from the ministry level would probably be controversial too. We still consider it a serious weakness that some ministries take such a limited role in the governance of projects, even when it comes to strategic activities such as project selection and portfolio management. It is our recommendation that the ministries should become more involved and actually play the role of "project owner type 1," formally or informally.

It should be noted that the research topic of this paper was the formal aspects of project governance. Müller (2017) introduced the term governmentality and were concerned with finding the optimal balance between formal and informal governance. The interviewees from the ministries in our study declared that their ministries had an informal dialogue with their agencies. This may very well be sufficient in some cases, but the interviews have also given us an impression that ministry involvement is ad hoc, not always early enough, and that many strategic decisions are, in practice, left to the agencies. A suggestion for future research could be to study these informal processes in more detail, preferably by following specific projects through the various phases, and revealing whether they actually compensate for the lack of formal processes and requirements.

Moreover, it is important to note that the ministries' governance of projects in subordinate agencies is part of their general governance of the agencies. The ministries set goals for the agencies and may give more or less detailed instructions with regard to, for example, activities and processes, depending on scope, risk, or political aspects. Traditionally, the ministries have not treated the agencies as project-based or governed them in terms of their projects, programs, and portfolios. The findings from our study suggest that they still do not. A relevant topic for future research could therefore be to take a closer look at how project governance could become a more integrated part of public governance.

As noted earlier in this paper, the top-level scheme in Norway and the other Scandinavian countries is very simple compared to, for example, in the UK, where the cabinet has introduced a more comprehensive model on top, including for example detailed processes, templates, and a common project management methodology. We have shown that a simple model on top does not necessarily imply that the governance framework is simple overall, since comprehensive governance models may be introduced by the lower levels. An interesting topic for future research could be to extend the international study conducted by Volden and Samset (2017b) to include all levels of the hierarchy, not only the top level. A more comprehensive governance scheme on top could be expected to result in harmonization of project practices across sectors and to strengthen public sector competence within project management and project governance. However, the result could also be a more bureaucratic system with less flexibility and autonomy available for the agencies.

Note

1. Including not only the largest projects for which external quality assurance is required, but also smaller investment projects.

References

Ahola, T., Ruuska, I., Artto, K. and Kujala, J. (2014), "What is project governance and what are its origins?", *International Journal of Project Management*, Vol. 32 No. 8, pp. 1321-1332.

Ansar, A., Flyvbjerg, B., Budzier, A. and Lunn, D. (2016), "Does infrastructure investments lead to economic growth or economic fragility? Evidence from China", *Oxford Review on Economic Policy*, Vol. 32 No. 3, pp. 360-390.

Association for Project Management (APM) (2011), *Directing Change: A Guide to Governance of Project Management*, Association for Project Management, Princes Risborough.

Berechman, J. and Wu, Q. (2006), "Cost overruns risk analysis in transportation infrastructure investments", UBC P3 Project Working Paper No. 2006-5, Phelps Centre for the Study of Government and Business, University of British Colombia, Vancouver BC.

Bevir, M. (2013), *A Theory of Governance*, Gaia Books, University of California, Berkeley, CA.

Cantarelli, C.C., van Wee, B., Molin, E.J.E. and Flyvbjerg, B. (2012), "Different cost performance: different determinants? The case of cost overruns in Dutch transport infrastructure projects", *Transport Policy*, Vol. 22, pp. 88-95.

Cantarelli, C.C., Flyvbjerg, B., Molin, E.J.E. and van Wee, B. (2010), "Cost overruns in large-scale transport infrastructure projects: explanations and their theoretical embeddedness", *European Journal of Transport Infrastructure Research*, Vol. 10 No. 1, pp. 5-18.

Crawford, L., Cooke-Davies, T., Hobbs, B., Labuschagne, L., Remington, K. and Ping, C. (2008), "Governance and support in the sponsoring of projects and programs", *Project Management Journal*, Vol. 39 No. 3, pp. S43-S55.

Eisenhardt, K. (1989), "Agency theory: an assessment and review", *The Academy of Management Review*, Vol. 14 No. 1, pp. 57-74.

Flyvbjerg, B. (2006), "Five misunderstandings about case-study research", *Qualitative Inquiry*, Vol. 12 No. 2, pp. 219-245.

Flyvbjerg, B. (2014), "What you should know about megaprojects and why: an overview", *Project Management Journal*, Vol. 45 No. 2, pp. 6-19.

Flyvbjerg, B. and Budzier, A. (2011), "Why your IT project may be riskier than you think", *Harvard Business Review*, Vol. 89 No. 9, pp. 23-25.

Flyvbjerg, B., Steward, A. and Budzier, A. (2016), *The Oxford Olympics Study 2016: Cost and Cost Overrun at the Games*, Saïd Business School Research Papers No. 2016-20, Saïd Business School, University of Oxford, Oxford.

Flyvbjerg, B., Bruzelius, N. and Rothengatter, W. (2003), *Megaprojects and Risk: An Anatomy of Ambition*, Cambridge University Press, Cambridge.

Flyvbjerg, B., Skamris Holm, M.K. and Buhl, S.L. (2003), "How common and how large are cost overruns in transport infrastructure projects?", *Transport Review*, Vol. 23 No. 1, pp. 71-88.

Haanæs, S., Holte, E. and Larsen, S.V. (2006), "Beslutningsunderlag og beslutninger i store statlige investeringsprosjekter", Concept Rapport No. 3, Concept-programmet, Norges teknisk-naturvitenskapelige universitet, Trondheim.

Hall, P. (1981), *Great Planning Disasters*, Penguin Books, Harmondsworth.

Helm, J. and Remington, K. (2005), "Effective project sponsorship: an evaluation of the role of the executive sponsor in complex infrastructure projects by senior project managers", *Project Management Journal*, Vol. 36 No. 3, pp. 51-61.

Hobbs, B. and Aubry, M. (2008), "An empirically grounded search for a typology of project management offices", *Project Management Journal*, Vol. 39 No. S1, pp. 569-582.

Klakegg, O.J. and Meistad, T. (2014), "Individual relations – the core line in project based organizations", paper presented at International Research Society for Public Management (IRSPM) Conference, Ottawa, April 9-11.

Klakegg, O.J. and Olsson, N. (2010), "An empirical illustration of public project ownership", *International Journal of Project Organization and Management*, Vol. 2 No. 1, pp. 16-39, available at: doi: http://dx.doi.org/10.1504/IJPOM.2010.03188

Klakegg, O.J. and Shannon, D. (2013), "Board of directors' responsibility for construction projects", paper presented at 7th Nordic Conference on Construction Economics and Organization, Trondheim, June 12-14.

Klakegg, O.J. and Volden, G.H. (2017), "Governance in public projects: the Norwegian case", in Müller, R. (Ed.), *Governance and Governmentality for Projects. Routledge Studies in Corporate Governance*, Routledge, New York, NY, pp. 129-155.

Klakegg, O.J., Williams, T. and Shiferaw, A.T. (2015), "Taming the 'trolls': major public projects in the making", *International Journal of Project Management*, Vol. 34 No. 3, pp. 282-296.

Kloppenborg, T., Manolis, C. and Tesch, D. (2009), "Successful project sponsor behaviour during initiation: an empirical investigation", *Journal of Managerial Issues*, Vol. 21 No. 1, pp. 140-159.

Love, P., Sing, C.-P., Wang, X., Irani, Z. and Thwala, D.W. (2012), "Overruns in transportation infrastructure projects", *Structure and Infrastructure Engineering*, Vol. 10 No. 2, pp. 141-159.

Lundberg, M., Jenpanitsup, A. and Pyddoke, R. (2011), "Cost overruns in Swedish transport projects", CTS Working Paper No. 2011:11, Centre for Transport Studies, Stockholm.

MacDonald, M. (2002), "Review of large public procurement in the UK", available at: www.parliament.vic.gov.au/images/stories/committees/paec/2010-11_Budget_Estimates/Extra_bits/Mott_McDonald_Flyvberg_Blake_Dawson_Waldron_studies.pdf (accessed January 2018).

McKinsey Global Institute (2013), *Infrastructure Productivity: How to Save $1 Trillion a Year*, McKinsey and Company.

Makovšek, D., Tominc, P. and Logožar, K. (2012), "A cost performance analysis of transport infrastructure construction in Slovenia", *Transportation*, Vol. 39 No. 1, pp. 197-2014.

Merrow, E.W. (2011), *Industrial Megaprojects: Concepts, Strategies and Practices for Success*, John Wiley & Sons, Hoboken, NJ.

Miller, R. and Hobbs, B. (2005), "Governance regimes for large complex projects", *Project Management Journal*, Vol. 36 No. 3, pp. 42-50.

Miller, R. and Lessard, D.R. (2000), *The Strategic Management of Large Engineering Projects*, MIT Press, Cambridge, MA.

Morris, P. (2009), "Implementing strategy through project management: the improtance of managing the project front-end", in Williams, T., Samset, K. and Sunnevåg, K.J. (Eds), *Making Essential Choices With Scant Information*, Palgrave Macmillan, Basingstoke, pp. 39-67.

Morris, P.W. (2013), *Reconstructing Project Management*, John Wiley & Sons, Chichester.

Morris, P.W.G. and Geraldi, J. (2011), "Managing the institutional context for projects", *Project Management Journal*, Vol. 42 No. 6, pp. 20-32.

Morris, P.W.G. and Hough, G.H. (1987), *The Anatomy of Major Projects*, John Wiley & Sons, Chichester.

Müller, R. (2009), *Project Governance: Fundamentals of Project Management*, Gower, Aldershot.

Müller, R. (2017), "Introduction", in Müller, R. (Ed.), *Governance and Governmentality for Projects: Enablers, Practices and Consequences*, Routledge, New York, NY, pp. 1-10.

Müller, R., Andersen, E., Klakegg, O.J. and Volden, G.H. (2017), "Governance institutions", in Müller, R. (Ed.), *Governance and Governmentality for Projects: Enabler, Practices and Consequences, Routledge Studies in Corporate Governance*, Routledge, New York, NY, pp. 51-67.

Müller, R., Pemsel, S. and Shao, J. (2014), "Organizational enablers for governance and governmentality of projects: a literature review", *International Journal of Project Management*, Vol. 32 No. 8, pp. 1309-1320.

Narayanan, V.K. and DeFillippi, R. (2012), "The influence of strategic context on project management systems: a senior management perspective", in Williams, T. and Samset, K. (Eds), *Project Governance: Getting Investments Right*, Palgrave Macmillan, Basingstoke, pp. 3-45.

Odeck, J. (2004), "Cost overruns in road construction – what are the sizes and determinants?", *Transport Policy*, Vol. 11 No. 1, pp. 43-53.

OECD (2015), *Towards a Framework for the Governance of Infrastructure*, OECD, Paris.

Olsson, N. and Berg-Johansen, G. (2016), "Aspects of project ownership in theory and practice", *Procedia Computer Science*, Vol. 100, pp. 790-795.

Pinto, J. (2006), "Organizational governance and project success: lessons from Boston's big dig", Concept international Symposium, Trondheim.

Project Management Institute (2013), *A Guide to the Project Management Body of Knowledge (PMBOK® Guide)*, 5th ed., Newtown Square, PA.

Samset, K. (2003), *Project Evaluation: Making Projects Succeed*, Tapir Akademisk, Trondheim.

Samset, K. and Volden, G.H. (2013), "Investing for impact: lessons with the Norwegian State project model and the first investment projects that have been subjected to external quality assurance", Concept Report No. 36. *ex ante* academic, Trondheim.

Samset, K. and Volden, G.H. (2015), "Front-end definition of projects: ten paradoxes and some reflections regarding project management and project governance", *International Journal of Project Management*, Vol. 34 No. 2, pp. 297-313.

Samset, K., Berg, B. and Klakegg, O.J. (2006), "Front-end governance of major public projects", paper presented at the EURAM 2006 Conference, Oslo, May 18.

Shenhar, A.J. (2004), "Strategic project leadership®: toward a strategic approach to project management", *R&D Management*, Vol. 34 No. 5, pp. 569-578.

Siemiatycki, M. (2015), *Cost Overruns in Infrastructure Projects: Patterns, Causes and Cures*, IMFG Perspectives 11, Institute on Municipal Finance & Governance, University of Toronto, Ontario.

Tirole, J. (2001), "Corporate governance", *Econometrica*, Vol. 69 No. 1, pp. 1-35.

Too, E. and Weaver, P. (2014), "The management of project management: a conceptual framework for project governance", *International Journal of Project Management*, Vol. 32 No. 8, pp. 1382-1394.

van Wee, B. (2007), "Large infrastructure projects: a review of the quality of demand forecasts and cost estimations", *Environment and Planning B: Urban Analytics and City Science*, Vol. 34 No. 4, pp. 611-625.

Volden, G.H. and Samset, K. (2017a), "Quality assurance in megaproject management: the Norwegian way", in Flyvbjerg, B. (Ed.), *The Oxford Handbook of Megaproject Management*, Oxford University Press, pp. 406-428.

Volden, G.H. and Samset, K. (2017b), "Governance of major public investment projects: principles and practices in six countries", *Project Management Journal*, Vol. 48 No. 3, pp. 90-109.

Williams, T. and Samset, K. (2010), "Issues in front-end decision making on projects", *Project Management Journal*, Vol. 41 No. 2, pp. 38-49.

Williams, T., Klakegg, O.J., Magnussen, O.M. and Glasspool, H. (2010), "An investigation of governance frameworks for public projects in Norway and the UK", *International Journal of Project Management*, Vol. 28 No. 1, pp. 40-50.

Yin, R. (2013), "Validity and generalization in future case study evaluations", *Evaluation*, Vol. 19 No. 3, pp. 321-332.

Zwikael, O. and Smyrk, J. (2012), "A general framework for gauging the performance of initiatives to enhance organizational value", *British Journal of Management*, Vol. 23, pp. S6-S22.

Further reading

Association for Project Management (APM) (2009), *Sponsoring Change. A Guide to the Governance Aspects of Project Sponsorship*, Association for Project Management, London.

About the authors

Gro Holst Volden holds the position of Research Director of the Concept Research Program on Front-end Management of Major Investment Projects, at the Norwegian University of Science and Technology. Her research is within project governance, public decision processes, and assessment and evaluation of major public investments. She in an Economist from the Norwegian School of Economics, and has a prior career as a Senior Advisor in the consulting industry and in the government bureaucracy. Volden is also the President of the Norwegian Evaluation Society. Gro Holst Volden is the corresponding author and can be contacted at: gro.holst.volden@ntnu.no

Bjorn Andersen is a Professor of Quality and Project Management at the Norwegian University of Science and Technology. He has authored/co-authored around 20 books and numerous papers for international journals and conferences, in total more than 250 publications. He has managed and been involved in several national and international research and implementation projects. He serves as the Director of Project Norway, is an Academic in the International Academy of Quality, is Co-editor of the *International Journal of Production Planning & Control* and a Reviewer for several other journals and conferences, and directs the NTNU master program in mechanical engineering.

Paper VII

Available online at www.sciencedirect.com

ScienceDirect

International Journal of Project Management 37 (2019) 549–564

International Journal of Project Management

www.elsevier.com/locate/ijproman

Assessing public projects' value for money: An empirical study of the usefulness of cost–benefit analyses in decision-making

Gro Holst Volden

Concept Research Program, Norwegian University of Science and Technology, 7491 Trondheim, Norway

Received 3 June 2018; received in revised form 3 February 2019; accepted 4 February 2019
Available online xxxx
Available online 23 March 2019

Abstract

Value for money, as measured by cost–benefit analyses (CBAs), is a crucial part of the business case for major public investment projects. However, the literature points to a range of challenges and weaknesses in CBAs that may cause their degree of usefulness in decision-making to be limited. The paper presents an empirical study of CBA practice in Norway, a country that has made considerable efforts to promote quality and accountability in CBAs of public projects. The research method is qualitative, based on a case study of 58 projects. The results indicate that the studied CBAs are largely of acceptable quality and heeded by decision-makers. Appraisal optimism seems to have been reduced by the introduction of external quality assurance of CBAs. However, there is need for a more consistent assessment of the non-monetized benefits, and distinguishing them from other decision perspectives such as the achievement of political goals. The paper offers a set of practical recommendations to increase CBA usefulness further.

Keywords: Project value; Project appraisal and evaluation; Cost–benefit analysis; Business case

1. Introduction

1.1. Projects ought to be good value for money

The project management community has increasingly shifted its attention beyond the 'iron triangle' of cost, time, and quality, to take a wider, strategic view of projects. Projects are implemented to deliver benefits and create value for users, the parent organization, and/or society at large (Morris, 2013; Samset and Volden, 2012). Accordingly, project governance has become an important issue in project research and practice. It refers to the processes, systems, and regulations that the financing party must have in place to ensure that relevant and viable projects are chosen and delivered efficiently (Müller, 2009; Volden and Samset, 2017b).

Williams and Samset (2010) refer to the choice of project concept as the most important decision that project owners make. The choice of concept ought to be approved on the basis of a business case, in which the expected benefits and strategic outcomes are described (Jenner, 2015). The business case provides a rationale for the preferred solution, and is therefore crucial for future benefits and cost management (Musawir et al., 2017; Serra and Kunc, 2015).

This paper focuses on the cost–benefit analysis (CBA) which is often a crucial part of the business case. The CBA concerns the relationship between resources invested and the benefits that can be achieved and is a tool to determine the project's value for money (i.e. whether it is profitable for society). Specifically, the aim of a CBA is to compute the net present value (NPV) of a project or various project alternatives, as defined by Eq. (1):

$$NPV = \sum_{t=0}^{N} \frac{B_t - C_t}{(1+i)^t} \quad (1)$$

where B is social benefit, C is social cost, i is social discount rate, t is time, and N is the period of analysis. It can be used to rank

E-mail address: gro.holst.volden@ntnu.no.

https://doi.org/10.1016/j.ijproman.2019.02.007
0263-7863/00

projects unambiguously (Boardman et al., 2011). The decision rule is to adopt a project if the NPV is positive, or in the case of several alternatives, to select the project with the highest NPV. Alternative criteria such as the benefit–cost ratio or internal rate of return can be applied too, but the NPV is normally recommended as a metric.

The CBA is particularly relevant for state-funded projects, as they are regarded in an overall national perspective, rather than the perspective of particular agencies, regions, or stakeholder groups. The benefits are interpreted in terms of the affected people's willingness to pay for them, and the costs are defined by the value of the alternative uses of the resources (Boardman et al., 2011).

The aim of the CBA is to be comprehensive in terms of the coverage of a project's impacts (Sager, 2013), and to monetize them as far as possible. Various techniques have been developed to elicit the willingness to pay (WTP) for non-market goods. However, remaining impacts that cannot be monetized must be described and presented in other ways, to enable decisions to be made as to whether they will be likely to improve or depreciate the NPV. In some cases, if analysts are unable or unwilling to attribute a monetary value to key benefits, they may be forced to apply cost-effectiveness analyses. In such cases, the intention is to minimize a ratio involving the benefit in physical units and monetary costs (e.g. cost per life saved). Unlike the CBA, the cost-effectiveness analysis does not make it possible for the analyst to conclude that the given project will contribute to social welfare (Boardman et al., 2011). It is thus a subordinate or second-best measure of value for money. Additionally, various multicriteria analyses are sometimes used, but they are not measures of value for money. In this paper we focus on value for money as measured by the CBA and not on project analysis in general.

A number of authors have highlighted the value for money perspective and the CBA (e.g. Jenner, 2015; Laursen and Svejvig, 2016; Terlizzi et al., 2017). Governments and professional project management bodies all require assurance of value for money, such as the Association for Project Management (2018), the (former) Office of Government Commerce (2009), and the Project Management Institute (2017). Volden and Samset (2017a) studied project governance frameworks in six OECD countries, and found that all of the frameworks highlighted the CBA in the front-end of projects. This is a dominant method of appraisal in the transport sector, for which many countries have developed guidelines (HEATCO, 2006; Mackie et al., 2014). Similarly, highlighting the CBA in the front-end has been used to assess development aid projects for decades, and is referred to as one of the World Bank's signature issues (World Bank, 2010). The appraisal method is also increasingly used in other sectors.

1.2. The research gap

However, the attention paid to the quality and utility of CBAs is limited in project research. The broad but fragmented literature on CBAs, which discusses a number of challenges and weaknesses, is rarely cited in project management and project governance literature. This is surprising, as we would normally expect that the quality of an analysis affects the extent to which CBAs are used, their recommendations followed and social benefits realized. We claim that it is not sufficient to require a CBA to be performed, but that also its usefulness must be ensured as part of project governance frameworks. A number of studies have documented the limited impact of CBAs on political decisions (e.g. Annema, 2013; Eliasson et al., 2015; Nyborg, 1998). For example, a review of World Bank projects shows that CBAs are rarely mentioned in policy documents, and that the percentage of projects justified following CBAs is declining (World Bank, 2010).

The explanations given in the literature are multifaceted and involve both analytical and political issues. For example, the World Bank report notes that only 54% of CBAs were of acceptable quality, but also that high-quality CBAs were often disregarded by decision-makers (World Bank, 2010).

In this paper we focus on the analytical issues in terms of the weaknesses that materialize in CBA reports. Other authors have focused on issues such as adverse incentives at the decision-making level that may result in the value for money aspect being played down when decisions are made (e.g. Sager, 2016). Decision-making in a democratic setting is inherently complex, frequently unpredictable, and influenced by other decision logics than just the rational economic ones. Therefore, as noted by Samset and Volden (2015), the greatest potential for improvement might be to strengthen the analytical processes.

1.3. This study

The aim of this study is to increase knowledge about the quality and usefulness of CBAs as basis for project selection. We take the perspective of the financing party (the true owner) who, in the case of public projects, is the entire society and its taxpayers, as represented by the Cabinet.

We define seven research questions (RQs) that together cover the main weaknesses in CBAs that have been discussed in academic literature (cf. Section 2). We want to learn about the relative prevalence of these weaknesses and to consider to what extent they reduce the quality and usefulness of analyses. The seventh and last research question, about whether CBA recommendations are actually followed (RQ7), is therefore of particular interest, and we consider it in relation to the other six questions. The seven questions are as follows:

RQ1: Are the CBAs consistent across projects with respect to which impacts are included, whether a valuation has been performed, and parameters and assumptions applied?
RQ2: Are non-monetized impacts assessed and presented consistently?
RQ3: Are associated uncertainties identified and presented?
RQ4: Are distributional impacts presented as supplementary information?
RQ5: Are CBAs unbiased? Specifically, is there a difference between CBAs done by project promoters and CBAs done by an independent party?
RQ6: Is transparency and clarity acceptable in the reports?
RQ7: Do decision-makers follow the advice presented in the CBAs?

To answer these research questions, we apply high-quality empirical data from Norway. Since 2005, CBAs have been compulsory in appraisals of the country's largest public investment projects under the Ministry of Finance's Quality Assurance (QA) scheme. The scheme is presented and discussed by Volden and Samset (2017b).

The QA scheme applies to public infrastructure projects that exceed an estimated threshold cost of NOK 750 million (USD 100 million). In those projects, external quality assurance (QA) of decision documents is required before the Cabinet makes its choice of project concept. As a basis for the external QA, the sectoral ministry or agency prepares a *conceptual appraisal* (CA) document. The CA is the business case and must include an assessment of needs and overall requirements, a possibility study that results in at least three alternative project concepts, including the zero-investment alternative, and a CBA of these concepts. The QAs are performed by private consultants contracted by the Ministry of Finance. The QA team should review the CA and thereafter present its own independent CBA, with alternatives ranked on the basis of their estimated value for money. This implies that for each project there will be two value for money assessments, one produced by the initiating ministry or agency and the other by the external quality assurer.

The QA team includes economists who are experts on CBA. Additionally, the ministries and agencies use highly qualified people to prepare the CBAs. The CA-QA process takes place at the same stage in all projects' life cycle, namely the end of the pre-study phase. The Norwegian Ministry of Finance has issued guidelines with a set of overall requirements for CBAs that we consider to be in line with best practice internationally (Finansdepartementet, 2005, 2014).

We considered Norway an interesting research case because of the efforts made to ensure that CBAs are of high quality. According to Flyvbjerg's (2006) categorization of case study research, Norway is a 'critical case' (here understood as an assumed best case). Our findings should be relevant beyond the Norwegian context, our thinking being that CBA weaknesses observed in this country, with a project governance scheme that requires high-quality and quality assured CBAs, will most likely also be a problem in countries without such a scheme. That said, there may be cultural and other differences between countries that influence project practices. In a case study, we must always present reservations concerning transferability of results across countries.

In Section 2 we present a review of the literature on weaknesses in CBAs. The review forms the basis for the framework of analysis applied to study the case CBAs. The framework is presented in Section 3, and a description of the study data and methodology is provided in Section 4. In Section 5, we present and discuss the findings with respect to each research question. Lastly, in Sections 6 and 7 we present our conclusions and recommendations, and discuss possibilities for further work.

2. Literature review

Today it is widely recognized that not only programs and portfolios, but also individual projects, should be linked to higher-order goals and strategies. The project management community has been increasingly concerned with how projects create value and reap benefits (Shenhar et al., 2001; Zwikael and Smyrk, 2012; Morris, 2013; Breese et al., 2015; Hjelmbrekke et al., 2017). Whereas some authors focus on the front-end, others discuss benefits management throughout the project life-cycle (e.g., Serra and Kunc, 2015; Musawir et al., 2017).

However, this part of the project management literature is still young. As noted by Laursen and Svejvig (2016) the definitions of project benefits and value are sometimes vague and depend on the perspective chosen. Baccarini (1999) suggested a distinction between two levels of project success, i.e. project management success, which concerns delivery, and product success, which concerns the outcome. Samset (2003) suggested a triple-level performance test concerning project outputs, first-order effects for users, and long-term effects for society. A similar chain of benefits has been suggested by Zwikael and Smyrk (2012) and Serra and Kunc (2015) and is also largely in line with PRINCE2®.[1] In the framework suggested by Zwikael and Smyrk (2012) it is also specified who should be responsible for project success on each level. The project manager is responsible for success at the operational level (project management success), the project owner is responsible for success at the tactical level (project ownership success) and the funder is responsible for success at the strategic level (project investment success).

In this paper we focus on the highest level of project success (i.e., project investment success, in Zwikael and Smyrk's terminology) where benefits and costs are compared to determine the effective 'return' on the investment. The CBA takes an overall societal perspective where all benefits and costs to affected parties nation-wide ought to be included, and (to the extent possible) translated into the monetary amount that people are willing to exchange. This is not the only possible definition of project investment success (as discussed further in Section 2.1) but at least it provides a very clear definition.

The project management community has not devoted much attention to the specificities of the CBA thus far, and we therefore had to search for other types of literature. The 'CBA literature' is large, with publications in transport sector journals as well as journals in economics, public policy and other social sciences.

Many weaknesses and challenges have originated in both theory and practice regarding the use of CBAs, to the extent that decision-makers do not find them useful or trustworthy. Such weaknesses may remain undisclosed due to the complexity and often low transparency of the methodology. In the following subsections we synthesize the literature on the various weaknesses in CBAs, which may explain decision-makers' lack of confidence in this metric. The literature is fragmented in the sense that different authors focus on entirely different issues. However, we suggest the following categorization of the weaknesses in CBAs: (1) criticism of the CBAs' normative fundament, (2) discussion of various measurement problems, and (3) challenges relating to appraisal optimism.

[1] Projects IN Controlled Environments, see www.axelos.com.

2.1. The CBA – Its normative fundament

The CBA is a powerful project evaluation tool, primarily because it is not based on political preferences, and therefore it can be characterized as a 'neutral tool' (van Wee and Rietvold, 2013). However, this strength is also a weakness because the CBA only recognizes people's preferences in their role as consumers. By contrast, analysis of people's preferences in their role as citizens may give a different result (Mouter and Chorus, 2016), as may the use of either planners' preferences or decision-makers' preferences (Mackie et al., 2014). Thus, the CBA is a framework for measuring efficiency, not equity, alignment with political goals, or any other definition of social desirability. Inevitably, the use of WTP implies that more weight is attached to high-income groups than to low-income groups (Nyborg, 2014). Furthermore, by focusing on the *aggregate* WTP, the CBA disregards the fact that some groups may be *worse off* after project completion than they were previously. The use of aggregate WTP is justified by the Kaldor-Hicks efficiency criterion, according to which a new resource allocation would be an improvement for society if the winners could *hypothetically* compensate the losers and still be better off. However, there is no requirement for such compensation to be given (Nyborg, 2014).

Thus, the CBA is of little help in cases in which the public sector has clear policy objectives that differ from consumers' preferences. Nyborg (1998) found this an important reason why some Norwegian politicians did not trust the CBA, with politicians on the left of the political axis being most sceptical. Mouter (2017) has reported similar responses from Dutch politicians.

A related critique is that the CBA systematically downplays the welfare of future generations. Decision-makers are increasingly concerned with investments' sustainability (Eskerod and Huemann, 2013; Haavaldsen et al., 2014), which requires a more holistic and long-term perspective than taken in CBAs. In particular, the use of a discount rate in CBAs implies that impacts on future generations have low worth today, and this weakness has been criticized by a number of authors (e.g. Ackerman, 2004; Næss, 2006; Pearce et al., 2006).

Some researchers have suggested that the CBA should be replaced by some form of multicriteria analysis that is based on the preferences of planners or decision-makers, at least in cases with moral dimensions (Browne and Ryan, 2011, van Wee, 2013). Others have noted that a multicriteria analysis has weaknesses too, which makes it more subjective and manipulable (Dobes and Bennett, 2009). In our view, both types of analysis can supplement each other, as they measure different things. For all projects that either directly or indirectly aim to contribute to economic growth, the CBA should at least be partly relevant.

The solution to this weakness most often recommended by authors is that all the costs and benefits should be presented in a disaggregated and transparent form that shows how they are distributed, not just their aggregated effect. When relevant, a separate overview and discussion of significant distributional impacts, both within and between generations, should be provided in the report. In that way, decision-makers would be able to decide for themselves whether the distributional impacts are acceptable. The CBA could also be included more systematically in a broader project evaluation framework that includes other perspectives than efficiency, such as the Five Case Model in the UK, in which the economic case is one of the five cases (HM Treasury, 2013). Another framework, one that has been very influential in evaluations of development assistance projects, comprises the five OECD-DAC criteria of efficiency, effectiveness, impact, relevance, and sustainability (Samset, 2003). A variant of the latter framework has been applied in ex post evaluations of Norwegian projects (Volden, 2018).

2.2. Measurement problems

Even if the ethical and normative premises on which the CBA rests were accepted, the credibility and usefulness of the results might be low due to various measurement problems (Atkins et al., 2017). At an early stage, information about the effects of a project is sparse and depends on many assumptions (Samset and Volden, 2015). Thus, an early CBA will have many sources of error, such as omitted impacts, forecasting errors, and valuation errors. Several studies have indicated that cost estimates and demand forecasts are highly inaccurate (i.e. Flyvbjerg et al., 2003; Kelly et al., 2015; Nicolaisen and Driscoll, 2014; van Wee, 2007). For example, Nicolaisen and Driscoll (2014) reviewed 12 studies conducted within the transport sector in various countries and concluded that traffic forecasts were unreliable, largely due to weaknesses in the model specifications, combined with low transparency, which made it difficult for others to observe what had been done.

Prediction and valuation of non-market goods such as health, safety, and the environment are a particular challenge. Different studies have revealed very different estimates of people's willingness to pay for such goods: for example, research conducted for a recently published doctoral thesis revealed huge variation in the estimates of the value of a statistical life (Elvik, 2017). It should also be noted that valuation methods differ in what they measure. For example, while stated preference (SP) methods are designed to capture the *total value*, revealed preference (RP) methods estimate only *use values* (Boardman et al., 2011). In many cases, inferior approaches that violate the principle of consumer sovereignty are used, such as implicit valuation, whereby analysts use the government's WTP as a proxy for the population's WTP. As discussed by Sager (2013) and by Mouter and Chorus (2016), a related challenge is that the population's preferences may be unstable, and the difference between consumer values and political opinions may be blurred.

Thus, it is crucial that the uncertainty involved in estimation is not downplayed (Flyvbjerg et al., 2003). Additionally, transparency is crucial: Wachs (1989) recommends that all details of the models and parameters should be available to anyone who might wish to replicate, verify, or merely critique the uses of the technical procedures. This implies that the

findings must be presented in a disaggregated form and not only as a summary indicator (Nyborg, 1998; Næss, 2006).

A further challenge is that the CBA is normally based on a partial equilibrium model and only measures direct effects. This is acceptable as long as other markets are competitive, but following the publication of the SACTRA report in the UK (Standing Advisory Committee on Trunk Road Assessment, 1999), attention has been paid to market imperfections that may mean that the full benefits of a transport investment fail to be included in the CBA. Some authors have indicated that such wider economic benefits may be considerable (Venables, 2007; Vickerman, 2008), while others have noted that they may also be negative (Næss et al., 2017; Small, 1999). Given that these impacts are not included in the NPV, they must be identified, discussed, and potentially quantified separately.

More generally, some impacts are inherently difficult to quantify and monetize. In particular, environmental effects are often substantially underestimated or ignored in practice, despite being possible to measure in principle (Ackerman, 2004; Browne and Ryan, 2011; Kelly et al., 2015; Næss et al., 2017). CBA textbooks and guides make it clear that non-monetized impacts must be identified, described, and balanced against the NPV, yet few textbooks give specific guidance on how this should be done. In practice, the treatment of non-monetized impacts tends to be random or politically driven as noted by some authors (e.g. Ackerman, 2004; Mackie and Preston, 1998).

2.3. Appraisal optimism

The third and last weakness of CBAs is that they are inherently at risk of bias and manipulation. For example, Mouter (2017) interviewed decision-makers who said that they knew how easy it was to affect results by 'shifting the buttons in the model' (Mouter, 2017, p. 1134). As noted by Wachs (1989), planning is not just analytical, and 'the most effective planner is sometimes the one who can cloak advocacy in the guise of scientific or technical rationality' (Wachs, 1989, p. 477).

Mackie and Preston (1998) list 21 sources of error and bias in transport project appraisals and conclude that appraisal optimism is one of the most important sources. Empirically, it has been shown that not only are CBAs inaccurate, but also they are often biased on the optimistic side (Flyvbjerg et al., 2003; Kelly et al., 2015; Nicolaisen and Driscoll, 2014; van Wee, 2007; World Bank, 2010).

Significant research has focused on explaining leaders' and entrepreneurs' optimism bias as a feature inherent in human behaviour. Such people are self-confident and tend to exaggerate their own abilities and control over a situation. While some authors describe this behaviour as unconscious (e.g. Lovallo and Kahneman, 2003), others argue that the persistence of bias is intentional and driven by a persistent excess demand for project finance (e.g. Bertisen and Davis, 2008). The persistence of bias can also be explained in terms of a principal–agent problem (Eisenhardt, 1989), such as when project promoters, who themselves are not responsible for funding, compete for discretionary grants from a limited budget (Samset and Volden, 2015). However, it is difficult to find conclusive empirical evidence of manipulation, as noted by Andersen et al. (2016).

A common recommendation to avoid appraisal optimism, whether or not it is intentional, is to ensure an outside view (Flyvbjerg, 2009; Lovallo and Kahneman, 2003; Mackie and Preston, 1998). This can be done by, for example, applying historical data (e.g. through reference class forecasting) and/or by having an independent third party perform or review the CBA. Additionally, systematic ex post evaluations should be performed to learn about the costs and benefits that can be expected (Flyvbjerg et al., 2003; Mackie and Preston, 1998; Volden, 2018).

Additionally, incentives for true speech must be in place. In this respect, Flyvbjerg et al. (2003) and Samset and Volden (2015) all recommend that project promoters are made accountable for financing, risk, and benefits realization, and that the appraisals are transparent and open to scrutiny. Mouter (2017) points out that the CBA is often complex and lacks transparency, which makes it particularly difficult to discover manipulation. More generally, an overall project governance framework that takes the risk of front-end agency problems into account should be in place.

3. Conceptual framework

We argue that the three strands of literature discussed in the previous section give rise to three broad explanations for why CBAs may not be considered useful by decision-makers. A simple conceptual framework is presented in Fig. 1.We have chosen 'CBA usefulness' as the main outcome variable. It is a multifaceted term that, in meaning, partly overlaps other terms such as trustworthiness, validity, and credibility (see Patton, 1999, and Scriven, 2015, for a discussion of criteria of merit by which analyses and evaluations ought to be evaluated). Since the CBA is specifically intended for decision support, CBA usefulness is considered from decision-makers' perspective. To some extent, the assessment of CBA usefulness will be subjective and depend on each decision-maker's preferences, competencies, and other abilities, but our focus is on assessments with which most decision-makers are likely to agree.

In line with the three categories of weaknesses of CBAs presented in Sections 2.1–2.3, we argue that CBA usefulness is threatened when (1) the analysis is too narrow in terms of relevant aspects being included in the business case (only the CBA alone), (2) the analysis is inconsistent, incomplete, and uncertainties are underestimated, and (3) the analysis is biased on the part of the analyst. By contrast, CBA usefulness is high when these weaknesses are not present.

The next step is to develop a framework for the empirical analysis, based on the conceptual framework in Fig. 1. In practice, the relative significance of the weaknesses in CBAs is largely unknown, as is the extent to which CBAs adhere to the recommendations provided in the literature to avoid or mitigate the weaknesses. To date, few empirical studies have

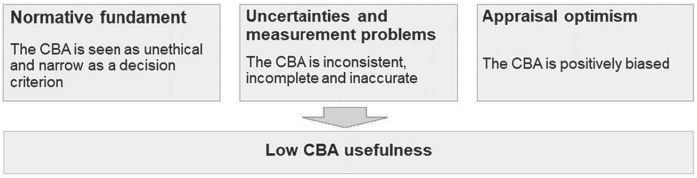


Fig. 1. Three types of weaknesses that lead to low CBA usefulness – a simple conceptual framework.

systematically reviewed CBA reports with respect to their overall relevance, quality, and credibility. This raises the question of whether it is possible for governments, through guidelines, quality standards, and other governance mechanisms, to ensure that CBAs are of high quality and useful to decision-makers. An interesting case is a recent study of the quality of CBAs of public projects in the UK (Atkins et al., 2017), in which the authors mainly focus on the second and third categories of weaknesses discussed above. The UK has taken steps to improve project competencies in central government and has introduced various governance arrangements to improve project performance (Volden and Samset, 2017a). Atkins et al. (2017) find that the CBAs are largely of acceptable quality, but that some challenges remain, the most important of which concern the lack of consistency across projects, and poor transparency and communication. They are also concerned about possible bias in the cost estimates, especially in cases in which decisions have been based on early estimates.

We draw on the most essential recommendations provided in literature, which, if adhered to, could increase CBA usefulness. Authors who criticize the normative foundations of the CBA (cf. Section 2.1) typically recommend that value for money assessments are supplemented by analyses of the project's impact on, for example, equity and sustainability. Those who discuss measurement problems (cf. Section 2.2) recommend a certain level of standardization, proper treatment of non-monerized impacts and uncertainty analyses. Lastly, those who are worried about appraisal optimism (cf. Section 2.3) recommend an outside view, and measures to ensure accountability. Common to all of the aforementioned three groups of authors is that they recommend transparent CBAs. Fig. 2 shows our framework for the empirical analysis, including the seven research questions presented in Section 1.

The use of the CBA in practice, understood as adherence to its recommendations, is a relevant indicator of CBA usefulness and is applied in this study (RQ7). We expect, ceteris paribus, that a CBA is more often adhered to when it is of high quality. However, it should be noted that adherence is not a perfect indicator of usefulness. As noted by Scriven (2015), there may be a number of reasons for lack of adherence to the results of a high-quality analysis. A thorough treatment of these issues would lead us beyond the analytical process and into political decision-making. Hence, for the purpose of this study, we merely assume that an instrumental decision logic or the 'rational ideal' is applied on the part of decision-makers (Samset and Christensen, 2017) and therefore disregard problems on the decision-making level, such as self-interest, the practice of 'horse trading', positioning, and power.

The final step in the outcome chain would be 'realized value for money'. This, too, would be an interesting indicator (although similar caution is required). Unfortunately, we do not have access to ex post data, and therefore this is not a topic of the empirical study.

4. Methodology

The empirical part of this study is largely qualitative, with the purpose of exploring, describing, and evaluating CBA practice within the Norwegian QA scheme. It is a multiple-case study of 58 Norwegian projects, based on a document review, interviews, and a review of the decisions made by the Cabinet. Although we refer to the cases as 'projects', all of the investments are studied in their early phases, in which they

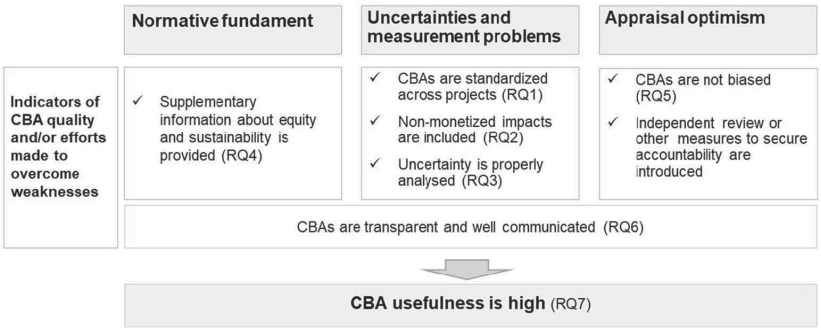


Fig. 2. Framework of analysis.

exist only conceptually. The Cabinet might chose the zero-investment alternative, in which case the project proposal will be rejected. Since very few of the projects have been completed, no information is available that can be used to determine the accuracy of the CBAs.

It should be noted that although the main unit of analysis is the project, we present some of the findings at 'CBA report' level (since most of the projects have two CBAs). At a higher level, one could consider Norway as a case, since all of the projects have been through CBAs in their front-end phase as part of the Norwegian QA scheme. However, this study is not an evaluation of the scheme but rather a study of CBA practice in a relatively large number of case projects, all of which belong to this (supposedly) favourable context.

The seven research questions listed in Section 1 were disaggregated into 25 subquestions that were more specific and contained indicators for the review of documents, as shown in Table 1. Some subquestions may contribute to answering more than one research question (RQ). However, the analysis was also inductive and open for exploring and describing other patterns and relationships that were revealed in the process.

Our main data source was the CA and QA reports for the 58 projects, which constituted the total population of projects that underwent CA and QA in the period 2005–2014, and are thus representative of projects in all of the major sectors that undergo QA in Norway. Currently, the transport sector has the largest number of projects, with most QAs performed on road projects. Other major categories are building construction, defence, and ICT projects.[2] The projects varied in size, complexity, purpose, and stakeholders involved, but in general they were the largest state-funded infrastructure projects in Norway in the period (Table 2).

For five of the projects (three of them within defence), the CA document was exempt from public access. For these projects, we only had access to the QA report and the presentation of the CA results discussed therein. Thus, we had access to a total of 111 CBA reports for our 58 projects.

The CA-QA process is followed by an administrative and political process in government. We established the status of all projects as of 2016, after the choice of project concept had been made by the Cabinet. To do this, we conducted a broad investigation of government documents, with particular focus on White Papers, to establish Parliament's ultimate choice of concept.

Additionally, we held semi-structured interviews with 26 key informants, all of whom were highly experienced within the field of CBA and had been involved in one or more of the studied projects. We considered that the interviews provided us with a deeper understanding, and since they were conducted after the document reviews, we were able to present some key findings and ask the interviewees for comments on them. Ten interviewees were senior ministry officers who commissioned CBAs from agencies, consultants, and quality assurers. They represented the decision-making level in this context. The other 16 interviewees were experts from the agencies and the QA teams and represented the persons who conducted the analyses. The interview guides were structured around the seven research questions, and the interviewees were invited to talk freely, based on their own experiences. It should be noted that the data collected from the interviews did not concern particular

[2] Some sectors are exempt from the Ministry of Finance's scheme, but have their own, similar schemes, such as the energy and petroleum sector, and the hospital sector. These are not included in the study.

Table 1
Subquestions applied for the review of CBAs.

RQ		Subquestion
RQ1	1	Describe the impacts included.
RQ1, RQ2	2	How are impacts treated (especially monetized or not).
RQ1, RQ6	3	Key assumptions and parameter values used to estimate the NPV (according to a pre-established list).
RQ1, RQ5	4	What is QAs reaction to CBA structure in CA? describe deviations between the two CBAs.
RQ2	5	Analyst's interpretation of the non-monetized impacts ('economic effect' or other).
RQ2	6	Methodology and measurement scale used to assess non-monetized impacts.
RQ2	7	Comprehensive analysis of non-monetized impacts? (pages used in the report)
RQ2	8	Comprehensive analysis of non-monetized impacts? (researcher's judgement)
RQ2	9	Non-monetized impacts – whose judgement? (e.g. experts, stakeholders, decision-maker).
RQ3	10	Type of risk analysis conducted, if any.
RQ3	11	Comprehensive risk analysis (researcher's judgement)? Capital cost, benefits, non-monetized separately.
RQ4	12	Distributional impacts or other considerations included along with the CBA.
RQ4	13	Comprehensive distributional analysis (researcher's judgement)?
RQ4	14	Distributional/other decision criteria clearly separated from CBA (researcher's judgement)?
All	15	Are the CA and QA in agreement on the recommendation?
RQ4, RQ5	16	Sign (and value?) of NPV of recommended alternative.
RQ4, RQ5	17	Is the recommended alternative the one with highest NPV?
RQ5	18	Is the zero option recommended?
RQ6	19	Overall level of transparency (researcher's judgement).
RQ6	20	Are models used to simulate impacts?
RQ6	21	If so, are the models explained? (reference to manuals, model version, etc.)
RQ6	22	Does the report include an executive summary?
RQ6	23	Is the report written in a non-technical language? (researcher's judgement)
RQ7	24	Status of the project as of today.
RQ5, RQ7	25	Whose advice is followed, CA or QA?

Table 2
Projects included in the research.

Projects included (sector)	N = 58
Road	20
Railway	5
Other transport (sea, coast, mixed)	11
Building	8
Defence	5
ICT	4
Sports event	3
Other	2

projects, but rather the general practice in central government. Each interview lasted 1–2 h.

A large Excel spreadsheet was applied, in which facts, assessments, notes from the document reviews as well as the interview transcripts were combined in the coding process. A list of the most interesting topics, counts, and possible relationships was continuously revised as we went through the material. The resulting themes and categories were not too different from the initial ones. The findings also included a number of categorizations, counting of occurrences, and cross-tabulation. In particular, the responses to subquestions 15 and 25, about whether the QA approved the CA and whose advice was followed by the decision-makers, were compared with various quality indicators. The results were also cross-tabulated against background variables such as project type.

All of the steps in the coding process gave considerable room for the researcher's own judgement, which might give rise to concerns about subjectivity and potential bias in our results. An important mechanism used to secure reliability and validity was the consultation of reliable sources of information. We used high quality, publicly available documents, as well as interviewees who had first-hand experience of CBA practice. The interviews were transcribed and the interviewees were subsequently given the opportunity to read and comment on the transcription. Furthermore, the use of different sources (i.e. document reviews *and* interviews, and interviewees with different perspectives) to illuminate each RQ, proved useful for revealing any inconsistencies in the data. The coding and analysis were also discussed with fellow researchers.

5. Presentation and discussion of findings

5.1. CBAs are comprehensive and partly standardized (RQ1)

Our overall assessment based on the document review is that most of the CBAs are relatively comprehensive, and that appraisals of similar types of projects generally include the same impact categories. In particular, payable costs, including both the capital cost and the maintenance and operating cost, are thoroughly estimated in most cases. Some benefits are monetized, most notably payable revenues, time savings, other consumer benefits, and in some cases also impacts on health and safety and the environment. Other impacts are treated as non-monetized impacts in the framework. Overall, only about half of the CBAs (45% of CAs, 55% of QAs) monetize all or the most important impacts. The degree of monetization varies across sectors, but even for road projects, less than 80% of the CBAs monetize all or the most important impacts. Thus, non-monetized impacts play a key role in the studied analyses.

Further, the CBAs of road and rail projects are more standardized than the CBAs of other project types. For example, whereas some CBAs of building projects only present and discuss first-order effects for users (e.g. users of a museum, university, or prison), others discuss long-term, wider benefits, such as improved national competitiveness due to better research and education. The interviewees reported that they were often unsure about whether and how to treat indirect, long-term impacts, for which no guidelines exist. Generally, the level of standardization regarding the non-monetized impacts is low. We return to this problem in Section 5.2.

Some quality assurers claim that the CAs are overly 'creative' with regard to the benefits included. This is particularly the case for non-monetized benefits. Table 3 shows the most common changes made by QAs relative to the CAs. The good news is that the largest category of changes is 'No or minor changes'. There are no clear sector differences. It can also be shown that 'No or minor changes' is correlated with QAs approving the final recommendation, cf. subquestion 15.

The calculation of an NPV is normally based on a number of parameters and assumptions, and an overview of some them is given in Table 4. Although it should be possible to vary most parameters due to, for example, local variation in people's WTP, it seems that the observed variation is somewhat higher than expected. For example, there seems to be much confusion about the discount rate and how it should vary according to systematic risk. Similarly, the degree to which real price adjustment is applied seems arbitrary. Some sectors (e.g. transport) have their own CBA guidelines that specify key parameters and values, implying that practice is more consistent in these CBAs. None of the CBAs included independent valuation studies to obtain exact WTPs.

Prior to 2014, hardly any parameters had been fixed as compulsory in the national guidelines issued by the Ministry of Finance, with the exception of the marginal cost of funds. Since 2014, some additional parameters have been fixed, most notably the discount rate and the value of a statistical life. In our view, this has led to a more consistent practice across CBAs, and should have been considered for other parameters too, most notably the social cost of carbon.

5.2. Inconsistent handling of non-monetized impacts (RQ2)

Non-monetized impacts are often essential in the CBAs. However, their interpretation is sometimes unclear and arbitrary, especially in the CAs. Some findings from the document review are presented in Table 5. On the one hand, the ministries and agencies seem to put more efforts into the analysis of non-monetized impacts than do the quality assurers, but on the other hand, they have a less clear understanding of what those impacts actually measure. Many CAs tend to mix economic impacts with goal achievement and other

Table 3
Changes in CBA structure. QA compared with the CA for the same project (most important change registered) (N = 58).

Type of change	Number	%
No change or minor change	17	29
Impact categories removed	13	22
Impact categories added	8	14
More impacts monetized (formerly non-monetized)	3	5
Impossible to compare due to different approach	12	21
No information	5	9
Total	58	100

Table 4
Selected parameters applied in the CBAs (N = 111).

Parameters	Practice observed
Marginal cost of public funds	0.2 (fixed by the Ministry of Finance)
Discount rate	Varies within the range 2–5%, later fixed at 4% and declines over time
Value of a statistical life	Varies in the range NOK 15–35 million, later fixed at NOK 30 million
Value of time	In most cases, average wage is used for business travel, but lower for leisure (in the transport sectors, based on a Norwegian SP study)
Method for calculating residual value	Large variations. Linear depreciation, market valuation, NPV of remaining net benefit flows, or set to 0
Real price adjustment	Large variations. Applied by some sectors, only for some impacts
Social cost of carbon	Varies within the range NOK 110–400 per ton, later an increasing price path is introduced in some sectors

considerations when presenting non-monetized impacts. Political and strategic considerations at various levels (e.g. agency, sector, region, or a stakeholder group) that extend far beyond consumer preferences are frequently brought into the discussion of whether the projects are good value for money. In our view, this is a serious weakness, that may lead to wrong conclusions.

Not only the interpretation, but also the choice of measurement scales varies considerably (e.g. cardinal, ordinal, or purely qualitative). Most CBAs of road projects apply the road agency's recommended framework for assessing five types of negative effects on nature and the environment, which are summarized in terms of 'plusses and minuses' on a scale ranging from −4 to +4. CBAs of other project types have a less systematic approach. Some quality assurers have introduced their own frameworks for analysing non-monetized impacts, but these frameworks are not consistent.

We consider that the documentation of the non-monetized impacts is sufficient in less than half the CBAs (cf. subquestion 8).

Table 5
Selected findings relating to non-monetized impacts in CBAs, sorted by CAs (N = 53) and QAs (N = 58).

Indicator	All (%)	CAs (%)	QAs (%)
Interpretation/perspective (researcher's understanding)			
Economic impact	56	34	77
Goal achievement, mixed or unclear	44	66	23
	100	100	100
Methodology			
Qualitative	22	21	23
'Plusses and minuses'	54	46	64
Other scoring or ranking	24	33	13
	100	100	100
Comprehensiveness			
Average % of CBA (in terms of page numbers)	22	27	17
Well documented (researchers' judgement), % 'yes'	45	53	36

Generally, the data sources used, the people involved, and the principles for valuation, are not well documented. For example, information about whose judgement they are based on is not provided in many cases. Moreover, in general, the development of these impacts over time is not discussed. There are no obvious differences between sectors or project types.

Interestingly, a comprehensive treatment of the non-monetized impacts in the CA is *not* correlated with QAs approving the final recommendation. Only when CAs apply the same interpretation of non-monetized impacts as the QA, they are more likely to agree on the final recommendation, and vice versa. This is supported by the interviews and indicates that quality assurers tend to be suspicious about a thorough discussion of non-monetized impacts that extend beyond an economic interpretation.

Interviewees from ministries and agencies acknowledged that performing the non-monetized part of the CBA is difficult. One interviewee said, 'In our sector [defence] we often discuss the achievement of military goals rather than socio-economic benefits. I guess we need better guidance on how to distinguish between a *multiple-criteria analysis* and a CBA.' By contrast, the quality assurers are more loyal to the economic perspective.

5.3. Uncertainty thoroughly assessed for capital cost, but to a lesser extent for other impacts (RQ3)

Our document review included an assessment of major uncertainties relating to costs and benefits, and how these were assessed and presented. Generally, the studied CBAs were more concerned about risks to the capital cost than risks to benefits and other long-term impacts. The reason probably lies in the QA scheme itself, which requires that stochastic estimation techniques are applied to estimate the capital cost, but there are no such requirements for other impacts. Overall, capital cost uncertainties are well handled in the studied CBAs. Uncertainties relating to other impacts are more varied and often superficial. About 60% of the CBAs (CAs and QAs alike) report sensitivity tests, but such tests are often simple and only focus on one or two parameters. One analyst said, 'We have strict deadlines, and sensitivity testing is just one of the things that we don't have time for.' Uncertainties relating to non-monetized impacts are rarely discussed in the CBAs. In our view, more attention should be paid to uncertainties in all impacts, not just capital cost.

The combination of uncertainties and irreversible investments that gives rise to quasi-option values (Boardman et al., 2011) is discussed briefly and qualitatively in some of the QA reports. Quasi-option values are typically higher in the zero-investment alternative, and in some cases this has been used by quality assurers as an argument for postponing the investment decision.

Overall, we consider that about two-thirds of the CBAs as acceptable with regard to identifying and analysing risk (cf. subquestion 11). QAs perform far better than do CAs (74% acceptable versus 47%). Interestingly, when a CA is in the 'acceptable' category, the QA approves the final

recommendation more often. This indicates that QAs recognize a good uncertainty analysis as a crucial quality indicator of the CBA.

5.4. Other considerations are not clearly distinguished from value for money (RQ4)

Overall, 47% of CAs present other decision criteria (goal achievement, distributional analyses etc.) along with the CBA, whereas only 5% of the QAs do the same (cf. subquestion 14). We do not find any clear sectoral differences. Generally, the discussion of distributional impacts is rather superficial, and in most cases not sufficiently comprehensive. Immediate effects are discussed more often than are long-term distributional effects. For example, impacts on future generations are hardly mentioned in any of the reports. An equally worrying observation is that when such other considerations are included in the report, they are in many cases not clearly separated from the value for money perspective.

As discussed in Section 5.2, benefits for specific groups or regions are often discussed in the CBAs as if they were net economic benefits to the country, although they may be a matter of redistribution. This explains the failure to report distributional impacts in many of the CBAs, particularly the CAs. They are already reported as benefits (but the corresponding negative impacts for other groups are not presented). By contrast, the quality assurers mention that their primary focus is on value for money, and some seem to ignore decision-makers' need for supplementary information altogether. Cross-tabulations show that CAs that present a broad and holistic decision base, correlates with QAs *not* approving their recommendations.

It should be noted that the distinction between wider economic benefits and pure distributional effects (i.e. economic effects that are most likely to be offset elsewhere) is not always clear. Our interviewees confirmed that performing this part of the analysis is challenging, and that more research and better guidance is welcome.

5.5. Appraisal optimism has been avoided for NPV estimation, but may influence the CBA in other ways (RQ5)

Although not always openly stated, there is commonly a preferred project alternative from the agency's perspective. One of the consultants stated: 'Everyone knows which concept the CA is hoping for, and it is always the most expensive one.' This raises the question of whether the CAs are biased in favour of a preferred alternative.

In the absence of ex post data, we compared the CBAs done by agency and quality assurer, in the knowledge that the latter party was independent of the project. It should be noted that the quality assurers may introduce new combined alternatives or adjustments to existing alternatives, for example to make the zero investment alternative more realistic, which implies that the sets of project alternatives assessed in the two reports are not identical. Therefore, instead of pairwise comparisons of alternatives, we studied the characteristics of *each party's highest ranked alternative*.

Generally, the QAs disagree with the CA recommendations, either partly or fully, in the majority of projects (33 out of 58). We have already mentioned that QAs seem to 'reward' CAs for having an appropriate CBA structure and for including a comprehensive uncertainty analysis, but *not* for comprehensive analyses of non-monetized impacts or for presenting a broad decision base. We also found that there are no striking sectoral differences: if anything, there seems to be slightly less disagreement about defence projects. Next, we focus our discussion on the extent to which CAs are systematically more optimistic about the projects' value for money. Specifically, in the knowledge that QAs put much weight on the NPV, one could suspect that the CAs present a biased NPV.

From Table 6, it can be seen that the CAs recommended project alternatives with a negative or zero NPV in 75% of the cases, whereas the corresponding percentage for the QAs is slightly lower (64%). Thus, it is apparent that the ministries and agencies are not concerned about promoting projects with a negative NPV. Rather, these findings may indicate that the NPV is not manipulated to make projects appear more profitable.

It should be noted that in our review of parameters and assumptions (cf. subquestion 3), we also looked for systematic differences between the CAs and the QAs. In this case, too, we did not find any clear indications that the CAs applied more optimistic parameters. Generally, practice seemed to vary as much between different quality assurers as between quality assurers and ministries and agencies.

However, we cannot exclude the possibility that CAs are biased in terms of the non-monetized impacts, or by excluding or systematically downgrading the simplest and less costly alternatives. As shown in the lower part of Table 6, CAs recommend the alternative with the highest or least negative NPV less often than do QAs. CAs hardly ever recommend the zero alternative. One group of projects that attracted our attention is those for which CA recommends an alternative with negative NPV and the QA recommends an alternative with positive NPV (10 projects). In each of these cases, the QA either preferred a less costly alternative, or downscaled the alternative recommended by the CA, thus turning a negative NPV into a positive one.

The findings presented in Table 6 also demonstrate the emphasis that ministries and agencies, and to some extent quality assurers put on the non-monetized impacts, which are

Table 6
Characteristics of the recommended project alternative (N = 58 projects).

Indicator	All (%)	CAs (%)	QAs (%)
Sign of NPV in recommended alternative			
Positive	30	25	36
Negative or zero	70	75	64
	100	100	100
The recommended alternative has the highest/least negative NPV, % of the CBAs	55	44	66
Zero alternative recommended, % of the CBAs	11	3	19

considered to outweigh a negative NPV in the majority of cases. In light of the emphasis put on those impacts, the inconsistent interpretation and treatment of such impacts is worrying (as discussed previously). Furthermore, there are indications that the quality assurers do not scrutinize this part of the CA in the same way as they scrutinize the NPV. One interviewed quality assurer said, 'I guess the agencies realize that any attempts to cheat with numbers will be revealed. It is easier to get away with the qualitative assessments.' The interviewees from the agencies denied that they had manipulated the data. Rather, they accused quality assurers of ignoring important non-monetized benefits. The interviewees who were decision-makers stated that they felt uncertain about how to interpret the reports and which party to believe when the CA and QA differed. First and foremost, they considered it important to be able to trust the quality of the CBAs. Some referred to the QA reports as helpful for determining the quality of the CAs, but one interviewee said he would have liked the QA reports to be 'reviewed by independent experts too'.

5.6. Transparency and communication acceptable, but could be improved (RQ6)

Transparency and clear communication are crucial to ensure CBA usefulness. Overall, we judge the level of transparency as acceptable (cf. subquestion 19) in c.80% of the studied CBAs, meaning that they are documented in sufficient detail, either in the main report or in an appendix. However, many reports could have been improved. Key parameters, such as the discount rate, price level, and period of analysis, are not always explicitly stated; for example, 12% of the CBAs do not include information about the discount rate used. Generally, the QAs are more transparent than are the CAs. There is also a tendency for the more transparent CBAs to have been produced by inexperienced agencies than by, for example, the road and rail agencies, possibly because they lack a standard framework and therefore need to explain every step of their analysis.

Traffic models and impact models are frequently used by the transport agencies, and some consultants have developed their own economic models that produce inputs to the CBAs. These models are not always well explained in the reports, and often appear as black boxes. Even experts in the agencies find the models difficult to understand, as exemplified by one interviewee, who said, 'The result of traffic simulations depends on so many detailed assumptions about the new road, such as curvature, width, velocity, etc. It is impossible to understand everything. You just have to trust the model.' One quality assurer admitted that he often took the traffic estimates from the agencies' models for granted, because it was impossible to verify them. By contrast, interviewees from ministries/agencies accused some consultant of treating their own models as business secrets.

Economic impacts are often presented in an aggregate form in the CBAs. For example, road projects normally generate a range of emissions to air (NOx, CO2, N2O, and local air pollution in the form of particulate matter). These are commonly presented in the reports as 'environmental costs', which obscures their individual impacts.

Furthermore, in all projects, a large number of project-specific assumptions will have to be set by the analyst. These are not always well explained in the CBAs. One example is the assumption made about toll fees on new roads in Norway, which may affect consumer benefits significantly. In two-thirds of the road project CAs, it is assumed there are no user fees, and hardly any of those CAs include an explanation of the reasons behind this assumption. The QA reports are therefore useful because they may question key assumptions. They may agree or disagree with the ministries and agencies, but their discussions will nevertheless add useful information for decision-makers. We only find a slightly positive correlation between the transparency in CAs and the QAs approving the final recommendation.

In many CBAs, technical language is used, and the reports are long: reports with 100 pages or more are common. This is relevant in terms of accessibility because decision-makers normally face constraints in terms of their expertise and time. The majority of CBAs (95% of QAs and 63% of CAs) include a summary. However, most of these summaries are short and rather superficial. In our view, only c.10% of the reports include a sufficiently informative summary that cover all major impacts (whether monetized or not), uncertainties, distributional impacts and/or other considerations, and key assumptions on which the results are based.

The interviewed decision-makers confirmed that they often found it difficult to understand the complexity of CBAs. They also confirmed that they thought summaries should be more comprehensive.

5.7. Decision-makers found CBAs more useful when approved by an independent party (RQ7)

The ultimate test of whether decision-makers' find CBAs useful, is the extent to which they follow the recommendations in the reports. Certainly, other concerns than value for money may affect public investment decisions, and traditionally the CBA has not been very influential in public project decision-making in Norway. However, it is important to note that the CBA follows an assessment of public needs and strategies, implying that the shortlisted alternatives are all considered relevant to these strategies. We therefore expect political decision-makers to follow the ranking based on value for money at least to some extent, given that they have confidence in the analyses.

Overall, in the majority of cases (c.80%), the Cabinet has chosen to go-ahead with either one conceptual alternative or, in a few cases, several conceptual alternatives to be developed further into a major construction project. Only in c.20% of the cases is the zero alternative selected or the project put on hold or withdrawn. There are no clear differences between project types. We did a large number of cross-tabulations to shed light on how CBA quality might have influenced decisions. The following findings are worth mentioning. A low degree of monetization does not seem to reduce adherence. Rather,

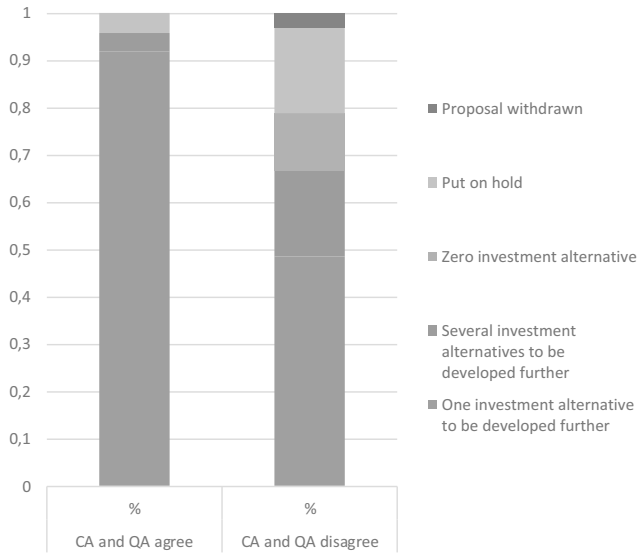


Fig. 3. Cabinet decisions, based on 58 projects, of which for 25 the two CBAs agree and for 33 they disagree with key recommendations. Percentages for each of the two groups.

decision-makers' adherence seem to be higher when the CBAs include comprehensive analyses of the non-monetized impacts and the distributional impacts, and they prefer reports that present a broad decision base that includes more than value for money. There is *no* correlation between adherence and the sign of the NPV in the recommended alternative, which is another indication that decision-makers care about the non-monetized impacts. By contrast, comprehensive risk analysis is not correlated with adherence. This is partly in contrast to the quality indicators that QAs seem to emphasize in their assessment of the CAs.

One finding that attracted our attention was that when CA recommendations (based on both NPV *and* the non-monetized impacts) were approved by the QAs, decision-makers' adherence was substantially higher. The distinction between cases in which ministry/agency and quality assurer agreed on the project ranking and cases in which they disagreed, is shown in Fig. 3. The Cabinet has followed the recommendation in 92% of the cases in which the two CBAs are in agreement. By contrast, when they are not in agreement, the Cabinet has made a clear choice of concept in only 48% of cases, often in line with CA recommendations. In the remaining 52% of cases, the Cabinet has chosen either multiple alternatives or no investment (the latter often in line with QA recommendations) or has put the decision on hold. In one case, a sports event, the proposal was withdrawn following a very critical QA report. These findings suggest that decision-makers care about more than just value for money. They also suggest that CBAs are heeded and that a critical QA can make decision-makers stop and reconsider the case.

Additionally, we asked interviewees to comment explicitly on the perceived usefulness of the CBAs. The majority, especially those who were decision-makers, found the CBAs useful 'given that they are of high quality'. One interviewee said, 'The existence of two CBAs that come to the same conclusion is a strong indicator of quality.' Another interviewee, a consultant, stated that 'In some cases, politicians need an excuse for rejecting a hopeless project, and a critical QA

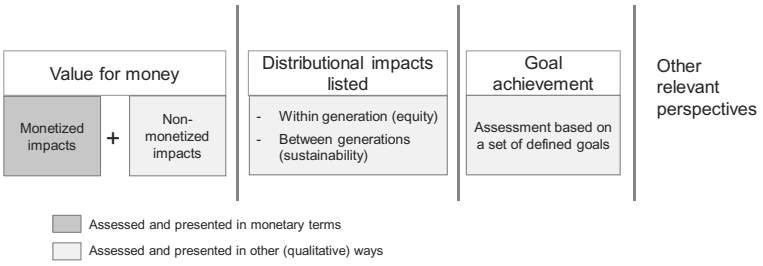


Fig. 4. Suggested early-phase business case – different decision perspectives (not to be added).

report can be that excuse.' Also, one interviewee noted that the QA scheme itself might discourage agencies from coming forward with poor proposals in the first place. However, another interviewee reminded us that decision-makers are not obliged to follow the advice from CBAs, and said, 'It is nice to know a project's value for money, but we cannot make politics only based on that.'

6. Conclusions

A CBA offers a clearly defined interpretation of project success, as may be formally required in relation to public project selection. However, challenges and weaknesses in CBAs may be overlooked, which implies that decision-makers may not find them useful and trustworthy. We have studied the usefulness of CBAs produced as part of compulsory appraisals of major infrastructure projects. Two types of CBAs are done for major public projects in Norway, one by the initiating ministry/agency and one by external quality assurers. Both types of CBAs rank the project alternatives based on their estimated value for money. With a few exceptions, they are openly available to researchers as well as to members of the wider public.

We expected, and found, that the studied CBAs would be and are largely of good quality. In particular, the use of independent quality assurers is normally considered a means to reduce the risk of appraisal optimism. Also, the risk of inconsistent, incomplete, and inaccurate estimates should be limited, given the time and resources spent on the analyses and the considerable expertise involved. Thus, the study of a 'critical case' (Flyvbjerg, 2006) should be useful to explore the potential for overcoming any CBA weaknesses, and to identify weaknesses that are more difficult to avoid than are others.

6.1. CBAs are heeded by decision-makers

A key finding from our research is that decision-makers consider CBAs a vital part of the business case for infrastructure project proposals. This was found through direct measurement (interviews) as well as indirectly (revealed adherence to recommendations). This contrasts with the role of CBAs before the QA scheme was introduced in 2005. In the past, if CBAs were produced at all, they rarely affected public project decision-making in Norway (Nyborg, 1998). Generally, we find that the ministries and agencies invest considerable resources in their CBAs today. This is in line with findings from an earlier study (Volden and Andersen, 2018), which demonstrates that the QA scheme has led to strong efforts in ministries and agencies to strengthen their project competencies and governance models at the agency level. However, we would like to make it clear that we have not proved an effect of the QA scheme as such.

We find that the Cabinet has almost always approved a project proposal if it was recommended as good value for money by the ministry/agency, *and* endorsed by the quality assurer. However, if a project proposal was recommended by the ministry/agency, but not endorsed by the quality assurer, the Cabinet was more likely to have rejected it or reconsidered it. This is a clear indication that the CBAs are heeded by decision-makers. Furthermore, the interviewed decision-makers explicitly stated that they considered the use of two CBAs was a stronger decision base than the use of just one CBA. This finding is in line with literature on appraisal optimism that recommends an external view on the appraisal and planning of a project (Flyvbjerg, 2009; Lovallo and Kahneman, 2003; Mackie and Preston, 1998).

Our findings indicate that appraisal optimism has largely been avoided in NPV estimation (i.e. the third category of weaknesses in CBAs, cf. Fig. 1). Ministries and agencies generally do not estimate NPVs as positive more often than do quality assurers. The fact that an external review will be performed seems to have a disciplining effect on ministries and agencies. However, we cannot exclude the possibility that CAs deliberately downgrade or exclude 'cheap alternatives' in some cases.

Furthermore, the comprehensiveness and consistency of analyses is largely at an acceptable level in the studied CBAs (cf. the second category of weaknesses). This also applies to transparency, which is essential to reveal all three types of weaknesses in CBAs and to increase decision-makers' understanding of the analyses. Thus, the situation in Norway is somewhat more encouraging than that found in the UK by Atkins et al. (2017), where inconsistency, poor transparency, and communication were serious weaknesses in project appraisals. Similarly, Annema (2013) found that transparency in Dutch CBAs was generally poor, despite the introduction of a new CBA guide that had led to other improvements. An explanation may be the requirement that QA reports in Norway should be openly available to the public. Nevertheless, there is potential for improvement in the Norwegian CBAs with regard to consistency and to uncertainty assessments and transparency.

To summarize, the following research questions have all largely been answered with a 'yes' response or at least a 'to an acceptable extent' response: RQ1 about CBA consistency across projects, RQ3 about uncertainties being identified and presented, RQ5 about unbiased estimates, and RQ6 about transparency and clarity. This may, in turn, explain why RQ7 about decision-makers' adherence to CBA recommendations, can also be answered with a conditional affirmative.

6.2. Non-monetized impacts need a clearer definition and more systematic treatment, distinguished from considerations beyond the project's value for money

Two remaining weaknesses in CBAs require attention. First, RQ2 about whether the non-monetized impacts are handled consistently has been answered negatively. Second, the answer to RQ4, about distributional impacts and other considerations, is that such issues are being presented and discussed in CAs, but they are often mixed with the value for money assessments. The former finding is much in line with the findings of Ackerman (2004) and of Mackie and Preston (1998), whereas the latter finding has not been studied systematically, to our knowledge.

It should be noted that the two weaknesses are related. There may be many pros and cons relating to the project beyond value for money. Our findings confirm that decision-makers do care about information beyond value for money assessments. However, when included, such other considerations are often incorrectly referred to as non-monetized impacts and 'added' to the NPV. This creates confusion for decision-makers, who cannot be sure what has been measured (i.e., whether value for money or some other confounded criterion).

One explanation for such observed weaknesses is that the non-monetized part of a CBA is a difficult topic - a fact that is neglected in CBA textbooks and guidelines. However, differences between ministries/agencies and quality assurers may also indicate opportunism. This means that ministries and agencies may deliberately overestimate the non-monetized impacts by including benefits that are not true economic benefits, and they could do this in the knowledge that it would be more difficult for the quality assurers to disprove qualitative arguments than quantitative arguments. If that is the case, the problem of appraisal optimism in CBAs may be present after all, although in another form than expected.

Clearly, methodological improvements as well as guidelines for assessing non-monetized impacts are required. Additionally, quality assurers must take such impacts seriously. Assessments of non-monetized impacts ought to be guided by the question of *whether they are likely to improve or deteriorate the NPV*, not by some other valuation principle (such as whether they are in line with a set of political goals). Admittedly, the distinction between consumer preferences and other perspectives is not easy in practice, but this is also a challenge in monetization (Sager, 2013; Mouter and Chorus, 2016). If we allow for arbitrary interpretations of the non-monetized impacts, the pricing versus non-pricing decision could become an opportunistic one.

As noted by Laursen and Svejvig (2016), the definition of 'value' in projects is often vague and may depend on the perspective taken. The great advantage with value for money as defined by the CBA is the clarity. Therefore, it is important to accept that definition in practice, whether impacts are monetized or not monetized. The great advantage with value for money as defined by the CBA is the clarity. The disadvantage is that only efficiency aspects are covered. We believe the definition of CBA should be accepted, whether impacts are monetized or not monetized. However, with a narrow interpretation of non-monetized impacts, it is even more crucial to balance value for money against other perspectives or interpretations of social value. Not only should each project alternative's distributional impacts be presented as part of the business case, but we suggest that also each project alternative's achievement of relevant goals and strategies is assessed and presented. Goals and strategies may overlap with value for money, which would typically be the case when goals are related to national economic development. In other cases, goals and strategies may be better aligned with distributional considerations, and thus in conflict with value for money considerations. For example, goals could be defined for the well-being of specific groups or regions, environmental sustainability or other considerations not well covered by the CBA (cf. Section 2.1). Basically, goals and strategies could be related to anything that political decision-makers care about.

Admittedly, goal alignment is already checked for the shortlisted alternatives in a CA, but some alternatives will often score higher than do others, which may be relevant for project selection. We think it is important that the three (or more) perspectives are presented separately, as shown in Fig. 4 by the thick lines between them. Thus, it is clear that although the monetized and non-monetized impacts should be added to assess the project's value for money, the different decision perspectives should not be added. Instead, any conflicts between the perspectives should be identified, and the final balancing between them ought to be done by the decision-makers. Should there be no conflicts, this will normally be highly relevant and useful information too. The framework constitutes a holistic business case that can easily be expanded to fit with an early-phase version of the Five Case Model applied in the UK (HM Treasury, 2013) or with the OECD-DAC criteria (Volden, 2018). This topic is worthy of more attention from the research community as well as from governments.

6.3. Recommendations

The findings from our research have provided the basis for a set of practical recommendations to increase CBA usefulness. The target group for these recommendations is project owners and senior officers who are responsible for project governance frameworks. Although the studied projects are public ones, we believe that many of the following recommendations are relevant to private sector organizations too.

1. A number of perspectives beyond value for money may be relevant to decision-makers. We suggest these perspectives are defined by decision-makers in advance and included in the business case. In our study, high-quality CBAs were often presented alone, forming a business case that was too narrow.
2. An important purpose of a CBA is to assess a number of alternative solutions to the problem at hand. Not only large construction projects, but also simple and low-cost solutions. One should be aware that project promoters may not have the right incentives to include the latter type of alternatives.
3. Completeness and consistency are important quality criteria, which comprise, for example, the impact categories included, the extent to which impacts are monetized, and the choice of parameter values. Although all projects are unique, our findings indicate that there is room for more standardization.
4. Possible errors and uncertainties need to be identified and presented as part of the CBA, to the extent that they can affect the ranking and recommendations.
5. The Non-monetized impacts are as relevant as the monetized impacts. They should not be ignored (as some, highly

experienced, analysts tended to do, in this study), nor should they be overvalued or mixed with other perspectives than the value for money perspective.
6. Measures to prevent optimism bias on the part of project promoters are recommended. Relevant measures such as transparency and external quality assurance of reports, seemed to work well in the studied projects.
7. Although not found to be a problem in this study, analyst competence and qualifications are key.
8. Understandability and communication (meaning, for example, the use of simple language and a readily available summary) are important aspects of transparency in reports, and relevant to decision-makers who are not CBA experts.

7. Limitations and further work

The use of case projects from a single country has some limitations. Therefore, broader conclusions cannot be drawn on the basis of our findings. In particular, as highlighted in the governance literature, a project governance scheme ought to be adapted to a specific context. The experiences gained from the application of the Norwegian QA scheme may not be transferable to other countries. An interesting topic for further study would be a systematic comparison of CBA practices in countries that have introduced independent review of CBAs.

Further, for the sake of simplicity, we have assumed that decisions are based on an instrumental decision logic, and we have not considered adverse incentives on the decision-making level. The true potential for improving decisions through better CBAs would probably be moderated by various conditions at the decision-making level. An extended model ought to be established to take this into account.

Additionally, it should be noted that we have studied CBAs in an early project phase. It remains to be seen whether the projects will actually be good value for money after they have been implemented. The selected project alternative needs to be developed further in a detailed planning process before the project is implemented. In that phase, there is a risk of cost escalation, and the realization of intended impacts has to be ensured through active cost and benefits management. An interesting topic for further research would therefore be to follow the projects throughout subsequent phases, and to perform updated CBAs *in medias res* as well as *ex post*, to learn whether the agencies manage to retain their focus on producing value for money.

Funding

The work was supported by the Concept Research Program at the Norwegian University of Science and Technology, which in turn is funded by the Norwegian Ministry of Finance.

Declaration of interest

No conflicts of interest.

Acknowledgements

The author would like to thank her colleagues Ole Jonny Klakegg, Heidi Bull-Berg and Ola Lædre for interesting discussions and useful comments to an earlier draft. Further, she would like to thank the editor and the anonymous reviewers for their contributions to improving the paper in the review process.

References

Ackerman, F., 2004. Priceless benefits, costly mistakes: what's wrong with cost-benefit analysis. Post-autistic Econ. Rev. 25.

Andersen, B.S., Samset, K., Welde, M., 2016. Low estimates – high stakes: underestimation of costs at the front-end of projects. Int. J. Manag. Proj. Bus. 9 (1), 171–193.

Annema, J.A., 2013. The use of CBA in decision-making on mega-projects: Empirical evidence. In: Priemus, H., van Wee, B. (Eds.), International Handbook on Mega-Projects. Edward Elgar, Cheltenham, UK, pp. 291–313.

Association for Project Management, 2018. Resources. https://www.apm.org.uk/resources/ (retrieved April 2018).

Atkins, G., Davies, N., Kidney Bishop, T., 2017. How to Value Infrastructure: Improving Cost Benefit Analysis. Institute for Government, London, UK.

Baccarini, D., 1999. The logical framework method for defining project success. Proj. Manag. J. 30 (4), 25–32.

Bertisen, J., Davis, G.A., 2008. Bias and error in mine project capital cost estimation. Eng. Econ. 53 (2), 118–139.

Boardman, A., Greenberg, D., Vining, A., Weimer, D., 2011. Cost-Benefit Analysis. 4th ed. Pearson.

Breese, R., Jenner, S., Serra, C.E.M., Thorp, J., 2015. Benefits management. Lost and found in translation. Int. J. Proj. Manag. 33, 1438–1451.

Browne, D., Ryan, L., 2011. Comparative analysis of evaluation techniques for transport policies. Environ. Impact Assess. Rev. 31 (3), 226–233.

Dobes, L., Bennett, J., 2009. Multi-criteria analysis: 'good enough' for government work? Agenda 16 (3).

Eisenhardt, K., 1989. Agency theory: an assessment and review. Acad. Manag. Rev. 14 (1), 57–74.

Eliasson, J., Börjesson, M., Odeck, J., Welde, M., 2015. Does benefit–cost efficiency influence transport investment decisions? J. Transport Econ. Pol. 49 (3), 377–396.

Elvik, R., 2017. The Value of Life: The Rise and Fall of a Scientific Research Programme. Doctoral theses at NTNU 2017:340. NTNU, Trondheim, Norway.

Eskerod, P., Huemann, M., 2013. Sustainable development and project stakeholder management: what standards say. Int. J. Manag. Proj. Bus. 6 (1), 36–50.

Finansdepartementet, 2005. Veileder i samfunnsøkonomiske analyser. Regjeringen, Oslo, Norway.

Finansdepartementet, 2014. Prinsipper og krav ved utarbeidelse av samfunnsøkonomiske analyser mv. Rundskriv R-109/14. Regjeringen, Oslo, Norway.

Flyvbjerg, B., 2006. Five misunderstandings about case-study research. Qual. Inq. 12 (2), 219–245.

Flyvbjerg, B., 2009. Survival of the unfittest: why the worst infrastructure gets built—and what we can do about it. Oxf. Rev. Econ. Policy 25 (3), 344–367.

Flyvbjerg, B., Bruzelius, N., Rothengatter, W., 2003. Megaprojects and Risk: An Anatomy of Ambition. Cambridge University Press, Cambridge, UK.

Haavaldsen, T., Lædre, O., Volden, G.H., Lohne, J., 2014. On the concept of sustainability – assessing the sustainability of large public infrastructure investment projects. Int. J. Sustain. Eng. 7 (1), 2–12.

HEATCO, 2006. Deliverable 5: Proposal for Harmonised Guidelines. http://www.kbsz.hu/dokumentumok/20070411_0.2-HEATCO_D5.pdf (retrieved 24 Sept 2018).

Hjelmbrekke, H., Klakegg, O.J., Lohne, J., 2017. Governing value creation in construction project: a new model. Int. J. Manag. Proj. Bus. 10 (1), 60–83.
HM Treasury, 2013. Public Sector Business Cases Using the Five Case Model (Green Book Supplementary Guidance on Developing Public Value from Spending Proposals).
Jenner, S., 2015. Why do projects 'fail' and more to the point what can we do about it? The case for disciplined, 'fast and frugal' decision-making. PM World J. 4 (3).
Kelly, C., Laird, J., Constantini, S., Richards, P., Carbajo, J., Nellthorp, J., 2015. Ex post appraisal: what lessons can be learnt from EU cohesion funded transport projects. Transp. Policy 37, 83–91.
Laursen, M., Svejvig, P., 2016. Taking stock of project value creation: a structured literature review with future directions for research and practice. Int. J. Proj. Manag. 34 (4), 736–747.
Lovallo, D., Kahneman, D., 2003. Delusions of success: how optimism undermines executives' decisions. Harv. Bus. Rev. 81 (7), 56–63.
Mackie, P., Preston, J., 1998. Twenty-one sources of error and bias in transport project appraisal. Transp. Policy 5, 1–7.
Mackie, P., Worsley, T., Eliasson, J., 2014. Transport appraisal revisited. Res. Transp. Econ. 47, 3–18.
Morris, P.W.G., 2013. Reconstructing project management reprised: a knowledge perspective. Proj. Manag. J. 44 (5), 6–23.
Mouter, N., 2017. Dutch politicians' use of cost-benefit analysis. Transportation 44, 1127–1145.
Mouter, N., Chorus, C., 2016. Value of time – a citizen perspective. Transp. Res. A 91, 317–329.
Müller, R., 2009. Project Governance: Fundamentals of Project Management. Gower, New York, NY.
Musawir, A., Serra, C., Zwikael, O., Ali, I., 2017. Project governance, benefit management, and project success: towards a framework for supporting organizational strategy implementation. Int. J. Proj. Manag. 35 (8), 1658–1672.
Næss, P., 2006. Cost-benefit analysis of transportation investments: neither critical nor realistic. J. Crit. Real. 5 (1), 32–60.
Næss, P., Volden, G.H., Odeck, J., Richardson, T., 2017. Neglected and Underestimated Negative Impacts of Transport Investments. Concept Report No. 54. Ex ante Academic Publisher, Trondheim, Norway.
Nicolaisen, M.S., Driscoll, P.A., 2014. Ex-post evaluations of demand forecast accuracy: a literature review. Transp. Rev. 34 (4), 540–557.
Nyborg, K., 1998. Some politicians' use of cost-benefit analysis. Public Choice 95, 381–401.
Nyborg, K., 2014. Project evaluation with democratic decision-making: what does cost-benefit analysis really measure? Ecol. Econ. 196, 124–131.
Office of Government Commerce, 2009. Managing Successful Projects with PRINCE2 (PRINCE2™). 5th ed. TSO (The Stationery Office), London, UK.
Patton, M.Q., 1999. Enhancing the quality and credibility of qualitative analysis. Health Serv. Res. J. 34 (5 Pt 2), 1189–1208.
Pearce, D., Atkinson, G., Mourato, P., 2006. Cost-benefit Analysis and the Environment: Recent Developments. OECD, Paris, France.
Project Management Institute (Ed.), 2017. A Guide to the Project Management Body of Knowledge (PMBOK® Guide), 6th ed Project Management Institute, Newton Square, PA.
Sager, T., 2013. The comprehensiveness dilemma of cost-benefit analysis. Eur. J. Transp. Infrastruct. Res. 13 (3), 169–183.
Sager, T., 2016. Why don't cost-benefit results count for more? The case of Norwegian road investment priorities. Urban Plan. Transp. Res. 4 (1), 101–121.
Samset, K., 2003. Project Evaluation: Making Projects Succeed. Tapir Academic Press. Trondheim, Norway.
Samset, K., Christensen, T., 2017. Ex ante project evaluation and the complexity of early decision-making. Public Org. Rev. 17 (1), 1–17.
Samset, K., Volden, G.H., 2012. The proposal. In: Williams, T., Samset, K. (Eds.), Project Governance: Getting Investments Right. Palgrave Macmillan, Basingstoke, UK, pp. 46–80.
Samset, K., Volden, G.H., 2015. Front-end definition of projects: ten paradoxes and some reflections regarding project management and project governance. Int. J. Proj. Manag. 34 (2), 297–313.
Scriven, M., 2015. Key Evaluation Checklist (KEC). http://michaelscriven.info/papersandpublications.html (retrieved 23rd January 2019).
Serra, C.E.M., Kunc, M., 2015. Benefits realisation management and its influence on project success and on the execution of business strategies. Int. J. Proj. Manag. 33 (1), 53–66.
Shenhar, A.J., Dvir, D., Levy, O., Maltz, A.C., 2001. Project success: a multidimensional strategic concept. Long Range Plan. 34, 699–725.
Small, K., 1999. Project evaluation. In: Gómez-Ibáñez, J., Tye, W.B., Winston, C. (Eds.), Essays in Transport Economics and Policy: A Handbook in Honor of John R. Meyer. Brookings Institution, Washington, DC.
Standing Advisory Committee on Trunk Road Assessment (SACTRA), 1999. Transport and the Economy. TSO, Norwich, UK.
Terlizzi, M.A., Albertin, A.L., de Oliveira, H.R., de Moraes, C., 2017. IT benefits management in financial institutions: practices and barriers. Int. J. Proj. Manag. 35 (5), 763–782.
Venables, A.J., 2007. Evaluating urban transport improvements: cost-benefit analysis in the presence of agglomeration and income taxation. J. Transp. Econ. Pol. 41 (2), 173–188.
Vickerman, R., 2008. Transit investments and economic development. Res. Transp. Econ. 23 (1), 107–115.
Volden, G.H., 2018. Public project success as seen in a broad perspective: lessons from a meta-evaluation of 20 infrastructure projects in Norway. Eval. Prog. Plan. 69, 109–117.
Volden, G.H., Andersen, B., 2018. The hierarchy of public project governance frameworks: an empirical study of principles and practices in Norwegian ministries and agencies. Int. J. Manag. Proj. Bus. 11 (1), 174–198.
Volden, G.H., Samset, K., 2017a. Governance of major public investment projects: principles and practices in six countries. Proj. Manag. J. 48 (3), 90–108.
Volden, G.H., Samset, K., 2017b. Quality assurance in megaproject management: The Norwegian way. In: Flyvbjerg, B. (Ed.), The Oxford Handbook of Megaproject Management. Oxford University Press, Oxford, UK.
Wachs, M., 1989. When planners lie with numbers. J. Am. Plan. Assoc. 55 (4), 476–479.
van Wee, B., 2007. Large infrastructure projects: a review of the quality of demand forecasts and cost estimations. Environ. Plan. B 34, 611–625.
van Wee, B., 2013. Ethics and the ex ante evaluation of mega-projects, in: Priemus, H., van Wee, B. (Eds.), International Handbook on Mega-Projects. Edward Elgar, Cheltenham, UK, 356–381.
van Wee, B., Rietvold, P., 2013. CBA: Ex ante evaluation of mega-projects, in: Priemus, H., van Wee, B. (Eds.), International Handbook on Mega-Projects. Edward Elgar, Cheltenham, UK, 269–291.
Williams, T., Samset, K., 2010. Issues in front-end decision making on projects. Proj. Manag. J. 41 (2), 38–49.
World Bank, 2010. Cost-Benefit Analysis in World Bank Projects. Independent Evaluation Group, World Bank, Washington, DC.
Zwikael, O., Smyrk, J., 2012. A general framework for gauging the performance of initiatives to enhance organizational value. Br. J. Manag. 23, 6–22.

Paper VIII

Paper IX

Evaluation and Program Planning 69 (2018) 109–117

Contents lists available at ScienceDirect

Evaluation and Program Planning

journal homepage: www.elsevier.com/locate/evalprogplan

Public project success as seen in a broad perspective. Lessons from a meta-evaluation of 20 infrastructure projects in Norway

Gro Holst Volden

Norwegian University of Science and Technology, 7491 Trondheim, Norway

ARTICLE INFO

Keywords:
Project evaluation
Project success
Evaluation framework
OECD-DAC criteria

ABSTRACT

Infrastructure projects in developed countries are rarely evaluated ex-post. Despite their number and scope, our knowledge about their various impacts is surprisingly limited. The paper argues that such projects must be assessed in a broad perspective that includes both operational, tactical and strategic aspects, and unintended as well as intended effects. A generic six-criteria evaluation framework is suggested, inspired by a framework frequently used to evaluate development assistance projects. It is tested on 20 Norwegian projects from various sectors (transport, defence, ICT, buildings). The results indicate that the majority of projects were successful, especially in operational terms, possibly because they underwent external quality assurance up-front. It is argued that applying this type of standardized framework provides a good basis for comparison and learning across sectors. It is suggested that evaluations should be conducted with the aim of promoting accountability, building knowledge about infrastructure projects, and continuously improve the tools, methods and governance arrangements used in the front-end of project development.

1. Introduction

The purpose of this study is (1) to demonstrate the importance of and need for a broad evaluation approach to measure success in large infrastructure projects, and (2) to test an evaluation methodology that is commonly applied in projects and undertakings in low-income countries, but now on projects in a more complex context in a high-income country.

1.1. Broad evaluation of public projects

Governments in high-income countries invest vast amounts of funds each year in infrastructure and other large public projects, such as roads and railways, public buildings, defence acquisitions and ICT infrastructure. The number and scale of such projects grow over time (Flyvbjerg, 2014). Even in a small country such as Norway, annual investments in large public projects amount to USD 6 billion per year not including petroleum sector investments (Norwegian Ministry of Finance, 2015).

Samset (2003) argues that in order to be truly successful, public investment projects must not only perform well operationally, but also tactically and strategically. Correspondingly, Baccarini (1999) defines two levels of project success, i.e. project management success (which concerns delivery), and product success (which concerns the outcome). However, whereas operational project success is highlighted by practitioners as well as academics (the problem of cost overruns being particularly well documented in the literature, cf. Morris & Hough, 1991; Flyvbjerg, Skamris Holm, & Buhl, 2003; van Wee, 2007), tactical and strategic success is often ignored, possibly because it challenges the way analysts think and has political aspects to it (Samset & Christensen, 2017).

Although Norway, as many other OECD countries, has been assigned a high level of evaluation maturity in national government (Jacob, Speer, & Furubo, 2015), systematic evaluations of public investment projects with respect to their outcomes are rarely conducted (Samset & Christensen, 2017; Rambøll & Agenda Kaupang, 2016). One exception is the transport sector, where benefit-cost analyses are performed to documents the projects' value-for-money (not so much ex-post, but before projects are submitted for government approval). However, many authors argue that benefit-cost efficiency is too narrow as measure of projects' tactical and strategic success (House, 2000; Heinzerling & Ackerman, 2002). This view is supported by the fact that benefit-cost efficiency rarely affects the priority ranking of road projects in Norway, which implies that decision-makers pursue other goals (Nyborg, 1998; Eliasson, Börjesson, Odeck, & Welde, 2015). Project success is clearly multi-faceted, and an evaluation can only be relevant to various stakeholders if it comprises a broader set of criteria.

This paper presents a generic framework for broad evaluations of large public projects. It is inspired by the criteria recommended by the Organisation for Economic Cooperation and Development's

E-mail address: gro.holst.volden@ntnu.no.

https://doi.org/10.1016/j.evalprogplan.2018.04.008
Received 22 September 2017; Received in revised form 14 March 2018; Accepted 29 April 2018
Available online 30 April 2018

Development Assistance Committee (OECD-DAC), which are much used in development assistance. The present study aims to demonstrate that the criteria are well-suited for infrastructure projects in industrial countries as well.

1.2. The case of Norway

Several authors have highlighted the crucial role of the front-end phase of projects (Williams & Samset, 2010; Morris, 2013; Samset & Volden, 2015). Many aspects that later create problems are typically present already at the project definition stage. In public projects, the Government as ultimate project owner should ensure the necessary quality-at-entry of project proposals and plans. This was done in Norway year 2000, when a scheme requiring external quality assurance of the decision basis was introduced for projects with an estimated investment cost exceeding USD 90 million. The scheme includes: (1) quality assurance of the choice of concept before the Cabinet decision to start a pre-project, and (2) subsequent quality assurance of the project management basis and cost estimate before the project is submitted to Parliament for approval and funding. Quality assurance is performed by external experts that are pre-qualified by the Ministry of Finance (Volden & Samset, 2017).

As of today more than 200 projects have been subject to quality assurance up-front, of which some 90 have so far been completed. There is evidence that this has improved the Norwegian Government's basis for decisions regarding major public investments (Kvalheim, Christensen, Samset, & Volden, 2015) and that the projects keep within their budgets (Welde, 2017). Nevertheless, the projects should also be evaluated ex-post, to verify how they actually perform in a broad perspective. In this study we test the suggested OECD-DAC evaluation framework on 20 Norwegian projects that were quality-assured in their front-end phase. The findings regarding these projects' performance are interesting in themselves, but the main purpose of this paper is to discuss the experiences with the evaluation framework and the evaluation process, as basis for improving and consolidating them.

The framework was first tested on four projects and the results presented to the Norwegian Ministry of Finance. The subsequent 16 evaluations included in this study were conducted on request from the Ministry. Ex-post evaluation has thus already become an integrated part of the project governance scheme and is likely to be used to further improve the quality assurance scheme. With time, a database is built, which allows for quantitative analyses of success at different levels across sectors and project types.

The paper starts with a presentation of the theoretical framing and the chosen evaluation framework. Then we present our methodology and data (the 20 projects), before we provide a synthesis of the findings in terms of the projects' success on various levels, and a discussion of the experiences with the evaluation framework and how it has been applied. Finally we offer some conclusions and discuss future extensions of the study.

2. Theoretical framing

Evaluation is the systematic investigation of the effectiveness of a project or other intervention. An evaluation requires evaluation expertise and rigorous application of scientific methods, while at the same time being focused on solving practical problems and being useful to project sponsors, decision-makers and other stakeholders (Rossi, Lipsey, & Freeman, 2004).

Evaluation became particularly relevant in the U.S. in the 1960s associated with the Kennedy and Johnson administrations and the social programs implemented at the time. Its aim was to learn from successes and failures and improve forward planning. It spread subsequently to other countries and different sectors, particularly to international development aid, where the effectiveness of investments and policy was contested.

Evaluation may be conducted at different stages during a project's lifetime. Each stage raises different questions to be answered, and correspondingly different evaluation approaches are needed. This would involve the assessment of i) the need for the project, ii) project design and logic/theory, iii) the implementation of the project, iv) its outcome or impact (i.e., what it has actually achieved), and v) its cost and efficiency (Rossi et al., 2004).

All projects are explicitly or implicitly based on an assumed set of causal relationships between inputs, project activities, outputs, and outcome. Several authors argue the merits of using this so-called logic model (McLaughlin & Jordan, 1999; Samset, 2003), also referred to as the programme theory (Chen, 1990; Weiss, 1997; Rogers, Petrosino, Huebner, & Hacsi, 2000) as representation of the project to help visualize important aspects, and especially when preparing for an evaluation. It helps clarify for all stakeholders: the definition of the problem, the overarching goals, and the capacity and outputs of the project (McLaughlin & Jordan, 1999). Further, looking at the different components of a project in relation to the overall societal objective, it allows for illumination of potential misalignments. Experience has shown that projects' logic is often unclear (Karlsen & Jentoft, 2013) and that goal hierarchies are characterized by a range of errors (Samset, Andersen, & Austeng, 2014). A critical assessment of the project's logic model might enable the evaluator to reveal a weak or faulty logic before any empirical evidence has been gathered (Brousselle & Champagne, 2011). In recent years, new versions of theory-based evaluation have emerged, such as realistic evaluation (Pawson & Tilley, 1997) and the theory of change (ToC) (Connell & Kubisch, 1998; Sullivan & Stewart, 2006).

3. A six-criteria evaluation framework

The chosen evaluation framework in the Norwegian context is presented below. As a general requirement, an overall evaluation framework ought to measure the success of projects in a broad perspective. It should be flexible enough to accommodate all types of projects, and sufficiently standardized to allow for comparisons between projects.

The starting point of each project evaluation should be the mapping of the logic model. The logical framework methodology is used, which focuses on the hierarchy of agreed goals, and identifies external risks on each level (Samset, 2003). The methodology was originally developed for USAID (Rosenberg, Posner, & Hanley, 1970), but its use spread rapidly by the UN, to aid administrations in a number of countries, and later to the OECD and the EU Commission.

In the Norwegian quality-assured projects, a logic model in the form of a goal hierarchy already exists, but it must be checked for consistency, and if necessary upgraded by the evaluator. When possible, the evaluator should also thoroughly investigate the goodness of the underlying theories (i.e. apply a truly theory-based approach), using existing literature and expert statements. The resulting model should be on the form illustrated in Fig. 1.

The overall evaluation criteria should be developed from the logic model. Since projects are de facto established to fulfil a certain purpose (Project Management Institute, 2013), one must ask *whether the intended*

Fig. 1. The logic model for a project.

results have been attained. Operational, tactical as well as strategic goal achievement should be assessed (i.e. output, outcome and societal objective). Furthermore, one should study any *side effects* that can be attributed to the project, i.e. impacts beyond those defined by the project owner (Sartorius, 1991; Gasper, 2000; Bakewell & Garbutt, 2005). Finally, it is important to note that public funding is a limited resource, and we must ensure that the funds are spent wisely. Therefore, regardless of whether efficiency and value-for-money are stated as project goals, an evaluation should always ask about this.

A standardized set of five evaluation criteria is much used, by the UN and other institutions and aid organizations, and has been endorsed by the OECD-DAC (OECD, 1991, 2002). Evaluation according to this framework highlights i) the need for the project (relevance), ii) whether the uses of resources and time are reasonable (efficiency), iii) whether stated goals are achieved (effectiveness), iv) what other positive or negative effects may occur because of the project (other impacts), and v) whether the positive effects persist after the conclusion of the project (sustainability).

As noted by Picciotto (2013), development projects are not so different from projects in developed countries. The five criteria reflect hard-won lessons of experience, and have by and large replaced prior approaches that focused only on inputs and outputs. They can be used equally at project, programme and policy level, and are aligned with the results-oriented stance favoured by most countries. Other sectors have introduced variants of the criteria (see, for example, European Commission (2013) concerning socio-economic development in Europe; ALNAP (2006) concerning humanitarian projects; and European Commission (2015) concerning regulations). A thorough review of the five criteria, which was performed by a group of professional evaluators (Chianza, 2008), concluded that the criteria work well and in particular that they satisfy Michael Scriven's Key Evaluation Checklist for programme evaluation (see Scriven, 2015 for the most recent version). However, Chianza suggested some improvements, the most important being to widen the interpretation of some criteria, such as relevance and sustainability not only covering the project owner perspective, and to define efficiency not only in the narrow sense (cost and time efficient delivery of the project), but also in terms of value-for-money. We agree with Chianza, and as a consequence have chosen to include benefit-cost efficiency as a sixth criterion of the model. For such economic analyses, we follow the standard method, as presented by, for example, Boardman, Greenberg, Vining, & Weimer, 2011. An implication is that the efficiency criterion focuses only on cost and time efficient project delivery.

Our definitions of the six criteria, and the level of success which they represent, are presented in Table 1. Their relation to the logic model is illustrated in Fig. 2.

The purpose of our evaluations is to give an *overall* picture of public project success. With budget limitations, we cannot be too ambitious regarding the methodological rigour when responding to each criterion. Experimental designs are rarely realistic for any of the evaluation criteria, and certainly not for the strategic ones. Rather, we must use simplified evaluation designs, economic data collection methods, and triangulation between various data sources and methods of analysis, quantitative as well as qualitative, to ensure validity and reliability. This is discussed further in the next section.

4. Research setting and research methodology

4.1. The projects

This study regards the evaluation of Norwegian public investment projects that have been subject to formal quality assurance before being approved for funding. In total, more than 200 major projects have been through the government's quality-at-entry scheme since the year 2000, representing primarily roads (53%), buildings (18%), railway (9%), ICT (11%) and defence (9%). Since the subsequent detailed planning and implementation period of such large projects is extensive, only 90 projects have been completed so far. Of these, 40 have been in operation for at least five years, and thus considered ready for evaluation.

A total of 20 projects was evaluated during 2012–2017 and are included in this meta-evaluation: eight road projects, five buildings, three railway projects, two ICT projects, and two defence projects (Table 2). The projects were chosen in chronological order, and constitute a relatively representative picture of quality-assured projects in their operational phase (50%). In addition to the sample projects, Table 2 shows which evaluators were involved. They represent consultancies in Norway and Sweden, and researchers from the Concept Research Programme, all considered independent of the projects and the implementing agencies.

4.2. The evaluation process

The six criteria framework does not guarantee high-quality evaluation by titself. In addition, evaluation skills, independent evaluators, appropriate data collection and analysis methods, etc. is required.

Each evaluation followed a defined process, which consisted of six steps, based on Samset's (2003) project evaluation textbook and also aligned with Michael Scriven's Key Evaluation Checklist (2015). In *Step 1* the Concept Research Programme selected the project to be evaluated, and sought acceptance from the responsible ministry (e.g. the Ministry of Transportation in the case of road projects). The ministries could, in principle, decline, but none of them did. A contact person in the ministry (and its subordinate agency when relevant) was identified.

In *Step 2* the evaluation team was established, usually following a public call. Researchers participated in some evaluation teams in order to gather experience in the use of the model. The team consisted of three or four people, all with good evaluation skills and knowledge of the sector. The scope was set to approximately three person-months of work per evaluation, depending somewhat on the project's complexity and availability of data.

Table 1
Definitions of the six evaluation criteria.

Level of success	Evaluation criterion	Definition
Operational	Efficiency	This criterion concerns project implementation and outputs in terms of cost, time and quality, and how economically the project organization has converted inputs into outputs.
Tactical	Effectiveness	This concerns whether the agreed outcome has been obtained and to what extent the project has contributed to this outcome.
Strategic	Other impacts	This includes all consequences beyond the agreed outcome (i.e. side effects) that can be attributed as the result of the project, positive and negative, short term and long term, for different stakeholders.
	Relevance	A project is relevant if there is a need for what the project delivers. Project relevance is measured in relation to national political priorities, but also stakeholders' preferences. It is essential to bring conflicts of interest to light as part of the evaluation.
	Sustainability	A project is sustainable if its benefits are likely to persist throughout its lifetime. This usually requires that the total impacts (financial, environmental and social) are acceptable in the long run.
	Benefit-cost efficiency	This should be measured in terms of total willingness to pay in relation to cost, or secondarily in terms of outcome in relation to cost (i.e. cost-effectiveness).

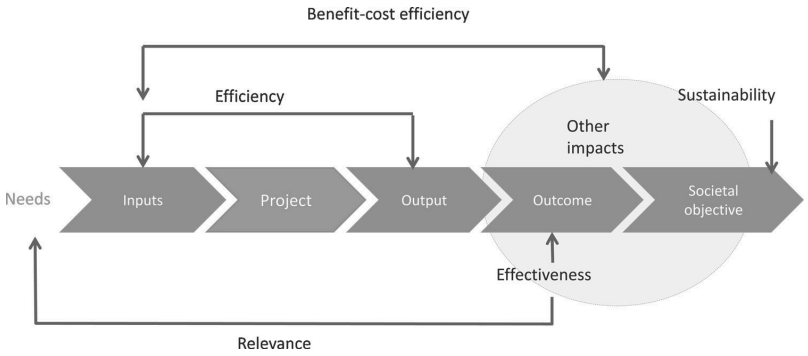


Fig. 2. The six evaluation criteria shown in relation to the logic model.

In *Step 3* the evaluation team reviewed and, if necessary, adjusted the logic model. Then it operationalized each of the six criteria by selecting more specific indicators or evaluation questions to be answered.

In *Step 4* the evaluation was carried out by collecting and analysing data, and answering the evaluation questions by combining different data sources and methods. We leaned on a number of authors who have suggested the mixing of methods to improve rigour in evaluations of complex interventions (see for example Ton, 2012; Green et al., 2017; Yin, 2013a,b), and the use of so-called rapid evaluation methods when faced with restricted budgets and timelines (World Bank, 2004; Bamberger, Rugh, Church, and Fort, 2004; Samset, 2003). As mentioned above, the use of existing literature and expert statements to assess the goodness of the programme theory was used as a supplementary approach to strengthen the validity of findings. Detecting impacts beyond the intended effects normally requires a wide, inductive and multidisciplinary approach.

In *Step 5* the evaluation team summarized its assessment for each criterion by setting a score between 1 and 6, where 1 is failure and 6 is highly successful. Score 4 should be awarded when the result for the relevant criterion is acceptable, but not an over-achievement. An overall guideline for score-setting was prepared in advance, to assist the evaluation teams.

Step 6. The final report was made public and distributed. The report and key results were stored in a database that is openly available to the public, (www.ntnu.edu/concept/evaluation-reports). The ministry and responsible agency were encouraged to follow up the results internally.

4.3. Meta-evaluation

This paper represents a meta-evaluation based on the findings and lessons learned from the 20 first evaluations. The term meta-evaluation is ambiguous. Generally, it implies that original analyses of data become the objects of a new analysis on a higher level (Glass, 1976). The much used UK Magenta Book (HM Treasury, 2011) primarily refers to meta-evaluation as a synthesis of a number of related evaluations, with the purpose of providing some estimate of the average or combined effect. This interpretation is close to what Yin (2013a) defines as cross-case synthesis. On the other hand, Scriven (2015) refers to meta-evaluation as an evaluation of one or more evaluations in order to identify their strengths, limitations and other uses, against a set of quality standards. A similar interpretation is suggested by Stufflebeam (2010) who distinguishes between three groups of standards: technical adequacy, utility and cost-effectiveness.

The present study applies the OECD's definition which includes both the above-mentioned interpretations: meta-evaluations are here defined as "*evaluations designed to aggregate findings from a series of evaluations. It can also be used to denote the evaluation of an evaluation to judge its quality and/or assess the performance of the evaluators*" (OECD, 2002, p. 26, our underlinings).

First, we present an aggregation and synthesis of the findings from the 20 separate evaluations, to establish the success of Norwegian investment projects. Second, we evaluate the evaluations themselves, including the suitability of the methodological framework.

Table 2
Key information relating to the sample projects. Sorted according to the year of evaluation.

No.	Project name	Sector	Start (yr)	End (yr)	Eval. (yr)	Evaluator
1	Customs area, Svinesund	Building	2004	2005	2012	SINTEF; Concept
2	Asker–Sandvika	Rail	2001	2005	2012	VTI
3	E18 Momarken–Sekkelsten	Road	2005	2007	2012	Concept
4	Skjold class MTB	Defence	2003	2013	2012	Scanteam; Concept
5	E6 Riksgr.–Svingenskogen	Road	2002	2004	2014	COWI
6	Svalbard Research Park	Building	2003	2005	2014	Concept
7	Lofoten fixed link	Road	2003	2007	2014	UiN; Nordlandsforskning AS
8	Eiksund fixed link	Road	2003	2008	2014	Menon
9	NAV ICT Basis	ICT	2006	2010	2014	NIBR
10	Østfold University College	Building	2003	2006	2015	SINTEF; Concept
11	E16 Kløfta–Nybakk	Road	2005	2007	2015	Urbanet
12	Rv 519 Finnfast fixed link	Road	2005	2007	2015	Menon
13	Sandnes–Stavanger	Rail	2005	2009	2015	Oslo Economics; Atkins
14	Military area. Østlandet	Defence	2002	2012	2015	Prokonsult AS
15	Perform	ICT	2008	2012	2015	Menon; Vivento
16	Halden Prison	Building	2006	2010	2016	Oslo Economics; Tyrilistiftelsen; Sweco
17	New Opera House. Oslo	Building	2005	2008	2016	HR Prosjekt
18	E6 Svingenskogen–Åsgård	Road	2005	2008	2016	Menon; Concept
19	E6 Åsgard–Halmstad	Road	2004	2005	2016	Menon; Concept
20	Gevingåsen railway tunnel	Rail	2009	2011	2017	Concept; SINTEF

The main data source for the first part was the 20 evaluation reports. We coded and summarized the assessments done by the evaluation teams, based on a set of questions prepared for the study. It included, for each evaluation criterion, the overall score as well as a range of more detailed indicators (e.g. for efficiency, it involved time, cost and quality, respectively). Different measurement scales were used for different aspects, including inter alia the number scale used for score-setting, binary variables (achieved/not achieved, etc.) and qualitative descriptions. Accordingly, aggregation of findings across projects was partly quantitative and partly qualitative. The scores awarded on each criterion were particularly useful since they allowed us to compute averages and see how they differed across sectors (although the number of projects was too small for statistical testing).

The second part of the meta-evaluation consisted of registration and assessments of various quality aspects of the evaluations themselves as well as the evaluation teams. Here we used a separate set of questions which was based on common advice from the project management and evaluation literature, and more specific recommendations concerning the OECD-DAC criteria. The questions included, inter alia, whether the criteria were interpreted and applied as intended, rigour in design, triangulation, adequate and efficient use of data sources, unbiased score-setting, use of a benchmark, and a sufficiently broad (multi-disciplinary) evaluation team. Ministries' and agencies' actual use of the findings in their planning of new projects was not included. Again, different measurement scales were used, many of them binary or qualitative. Our assessments were unavoidably subjective, but meticulously explained and documented.

Furthermore, a focus group meeting was held with 11 experienced evaluators, all of whom had participated in one or more evaluations. The group was confronted with preliminary findings and assessments, and the participants were given the opportunity to comment on and share their experiences and assessments of the model, the process and the need for guidance. The focus group meeting largely confirmed the picture presented by us, but assessments and conclusions were updated on some points.

Finally, it should be noted that the author also drew on her own experiences from the process of commissioning and following-up of the 20 evaluations, as well as participating in three of them.

5. Findings: public project success

In this section, we present and discuss the aggregated findings.

Table 3 summarizes the scoring results from the 20 evaluations. The overall picture is very positive, with average scores between 4 and 5 for all criteria, and the highest scores for efficiency and effectiveness.

Efficiency concerns the project's operational success. As shown in Table 3, 19 out of 20 projects scored 4 or better on this criterion. In total, 15 projects were completed with a final cost below the approved cost frame, 15 were completed within the time frame, and 16 were considered to meet high quality standards. The projects were also largely well organized and managed. The findings relating to cost control have been confirmed in a study that included 78 completed projects that had been through external quality assurance (Welde, 2017). In some projects, deviations from the cost frame were considerable, in both positive and negative direction. In some of these, the evaluator suggested that the cost frame was not realistic (typically because market uncertainty was underestimated). This indicates that the quality-at-entry scheme may not serve as guarantee for realistic budgets.

With regard to tactical project success, as measured by *effectiveness*, 18 projects scored 4 or better, which means that most of the projects' outcomes were in accordance with plans. Three projects received top scores, i.e. two road projects that realized very large time savings and reductions in accident levels, and one ICT project that generated major benefits in terms of time savings for the agency and improved quality of the services for the users.

The other four evaluation criteria express strategic project success. The evaluators concluded that the projects performed acceptably in these dimensions too. All 20 scored 4 or better on at least one strategic criterion.18 projects scored 4 or better on *other impacts*. Negative side effects were identified in few cases, and several projects generated positive side effects. One project, the New Opera House in Oslo, received top score for its very positive contribution to urban development, while some projects could have done more to avoid negative side effects. 17 projects scored 4 or better on *relevance*. This implies that most were considered solutions to real problems. However, some involved conflicts of interest (i.e. they were not equally relevant in all perspectives). This was the case in regards the Lofoten fixed link, a road project that connected a remote region to areas that were more urban, but left the neighbouring remote region even more isolated. 19 out of 20 projects scored 4 or better on *sustainability*. This implies that project benefits were largely expected to continue in a sustained number of years. However, the projects had to be sustainable in *all* aspects (i.e. financial, environmental and social) to be assigned a top score. For example, one project (defence) scored very low, because growing operational and maintenance costs had made it financially unsustainable. The projects scored slightly lower on *benefit-cost efficiency*. In total, 13 out of 20 projects scored 4 or better. The five most profitable projects were all road projects in urban areas.

In summary, most of the projects were considered successful in more than one aspect, and especially in operational terms. There appears to be some correlation between the scores for the various criteria. This is not surprising, since a well-thought-out and carefully planned project will normally be successful in several respects. However, there may also be conflicts, for example when some of the projects scored high on relevance and sustainability, and lower on benefit-cost efficiency. All three railway projects were well aligned with the government's strategy for sustainable transport. But with passenger numbers that were lower than estimated by the time of evaluation, and a relatively high capital cost, the value-for-money was considered low. We agree that not all projects can or should be 'profitable', but one should at least consider whether a simpler solution, still with acceptable goal achievement, would substantially improve value-for-money. Similarly, a project with high effectiveness but negative side effects should perhaps have been redesigned, such that it got a better overall score.

One interesting observation is that many of the projects that scored high in tactical and strategic terms were not aimed at specific stakeholder groups or regions, but rather followed from national political objectives. This supports earlier findings that when specific stakeholder groups manage to mobilize government funding for 'their' project, the project may turn out to be less relevant from the perspective of the wider society − a phenomenon known as perverse incentives in project

Table 3
Evaluation results (N = 20).

	Efficiency	Effectiveness	Other impacts	Relevance	Sustainability	Benefit-cost eff.
Average score	4.7	4.7	4.4	4.6	4.6	4.2
Median score	5	5	4	5	5	4
Min.–Max. score	3–6	3–6	3–6	3–6	2–6	2–6
4 or better (no. of projects)	19	18	18	17	19	13
5 or better (no. of projects)	13	12	8	13	13	9

Table 4
Evaluation results: average per sector.

	No. of projects	Efficiency	Effectiveness	Other impacts	Relevance	Sustainability	Benefit-cost eff.
Building	5	5.4	4.2	4.6	4.6	4.8	3.8
Defence	2	4.5	4.5	4.5	4.5	3.5	3.5
ICT	2	5.0	5.5	4.5	4.0	5.5	4.0
Railway	3	4.3	3.3	4.0	4.7	4.7	2.7
Road	8	4.4	5.3	4.3	4.6	4.5	5.3

selection (see Volden & Samset, 2015).

Although the number of evaluated projects is low, the results indicate some interesting sectoral differences, as shown in Table 4. In the following we will briefly comment on the three largest project categories (i.e., including at least three projects). The *building projects* performed excellently in operational terms, implying that they were delivered within time and budget and with a high quality. Some of these buildings were awarded architectural prizes. However, they scored slightly lower tactically and strategically, some of them with outcomes that had still not been realized several years after completion. Some of the projects had ambitious goals, such as to 'co-locate departments A and B and realize professional synergies', and should have devoted more attention to the fact that it takes more than a building to obtain such goals. The *railway projects* were closely aligned with government strategies for a 'green shift' in transport, and thus were considered relevant and sustainable. However, they scored low on benefit-cost efficiency and on effectiveness. As for building projects, ambitious goals usually require more than the physical infrastructure, and this should have been given more attention than was done in the sampled projects. The *road projects* scored high on most criteria, but somewhat lower on efficiency and other impacts. Road projects experienced the biggest cost overruns, but also the biggest cost underruns, implying that the Public Roads Agency should make efforts to obtain more accurate estimates. And, that more attention should be paid to side effects in the planning phase.

Overall, the findings concerning project success are very positive, which seems to conflict the public discourse and studies that demonstrate a low level of success in public projects. For instance, Flyvbjerg, Garbuio, and Lovallo, (2009) used the expression "*over budget, over time, over and over again*" and explained the widespread problem of cost overruns by delusion and deception. We think that caution should be made when referring to public projects as generally unsuccessful. Firstly, the media as well as the academic literature have mostly been concerned with cost control, which is only one aspect of project success. Secondly, as noted by Love, Smith, Simpson, Regan and Olatunji, (2015), different empirical studies of cost control come to very different conclusions, depending, inter alia, on the point of reference from which a cost overrun is measured (those that find the largest overruns typically compare with early and uncertain estimates). That said, we believe that the 20 Norwegian projects stand out as successful, which can be explained, at least to some extent, by the quality-at-entry requirements which ensure that they are thoroughly planned and reviewed before being submitted to Parliament for approval and funding. It should be noted that the remedy suggested by Flyvbjerg et al. (2009) to avoid delusion and deception, was to take an *outside view* on project planning and estimation, which is exactly what the quality-at-entry scheme does.

On the other hand we cannot eliminate the risk that some scores are positively biased. Wiig and Holm-Hansen (2014) found that *positive* side effects were mentioned more often than negative ones in evaluations of development assistance projects. They suggested that evaluators may be reluctant to criticise without hard evidence on which to base their criticisms, but willing to mention a positive issue that has the same level of uncertainty. While acknowledging the general risk of positive bias, we believe it is moderate to low in the 20 evaluations. The main reason is that the evaluations were organized by a third party, and all the teams were entirely independent of the projects.

However, another and more pertinent matter is how to ensure that the scores are well calibrated across projects. The scores were set by a different evaluation team in each case, based on the team's subjective assessments. As will be discussed in the next section, we believe there is a need for clearer guidelines for score-setting. In particular, the level of ambition inherent in the goal hierarchy must be taken into account when deciding on the score for effectiveness. We suspect that different levels of ambition related to project goals may explain some of the sectoral differences when it comes to effectiveness.

The findings of the evaluations should be useful for the purpose of accountability as well as for learning and improvement. Although each evaluation was limited in time and resources, it may identify major risks and problems that should be examined in more detail by the responsible ministry. Furthermore, the findings should provide input to the appraisal and planning processes of future investment projects funded by national government. It should be noted again that the sample of projects is not statistically significant, thus any attempts to generalize findings should be regarded as preliminary and tentative. Over time, when the database of evaluated projects is larger, it should be examined whether the patterns described above still hold.

6. Lessons learned about the evaluation framework and procedure

In the second part of the meta-evaluation, we look at the evaluations themselves, at how they were implemented, and at the framework's suitability for the purpose. The 20 evaluations have provided useful experiences and a basis for consolidating the model and practice.

6.1. General experiences of the model

The evaluators agreed that the chosen framework worked well, and that the six criteria covered the main aspects of public project success. Some noted that the strategic criteria, other impacts, relevance and sustainability, were 'eye-openers'. Knowing that pure economic evaluations are often considered by decision-makers to be too narrow (Nyborg, 1998), our evaluators agreed that the six-criteria framework should be more relevant.

The process of disaggregating the criteria into specific indicators and then aggregating the findings to provide answers to each criterion, provides a good balance between the need for standardization and flexibility. The evaluations have converged more and more into a common form, and their quality has improved over time.

However, we also see some challenges related to the interpretation of other impacts, relevance and sustainability. Some evaluators treated these strategic aspects superficially when realizing that they could not be measured in quantitative terms. Others interpreted them too narrowly, in the same way as found by Chianza (2008). Others, still, confused them with benefit-cost efficiency, downplaying, for example, environmental, social and ethical concerns. The explanation may be that many of the evaluation teams were dominated by economists. The evaluators confirmed that they were uncertain and wanted more guidance on the interpretation of these criteria.

6.2. Methodological rigour versus available resources

The evaluations consist of six criteria, each of which requires proper treatment. At the same time, the budget and time available implies clear limitations with regard to scope and methodological rigour. A timely question is therefore whether it is possible to give a professionally sound assessment of all six criteria. Experience so far suggests that the answer is yes. The 20 evaluations are largely of an acceptable quality. Admittedly, they are 'rapid' and the scores are sometimes uncertain, but this is not uncommon in evaluations, and must be accepted as long as the choice of methods and limitations are communicated. These findings support those of Samset (2003), Bamberger et al. (2004) and others who have argued that it is possible to conduct evaluations of acceptable quality under budget, time and data constraints.

Furthermore, the framework is flexible, and in individual cases, evaluators may spend more resources on a particular criterion, while treating other criteria more superficially. That has happened, for example, in cases of large deviations from the cost frame, which implied a need to look more closely into efficiency.

6.3. Methods for data collection and analysis

Different evaluation teams mostly chose the same methods for data collection and of analysis. For efficiency, they typically used data from project reports, interviews and benchmarking of cost data with similar projects. For effectiveness, they used time-series data for outcome indicators (often including comparison groups, such as similar geographic regions without the new infrastructure) and interviews with a wide range of stakeholders. For the strategic criteria, the evaluators used a combination of different sources, predominantly qualitative ones. For benefit-cost efficiency, they used all existing data and a set of assumptions and price tags. All evaluations included site visits.

It is costly to collect primary data, and evaluators must therefore prioritize carefully. Few of the evaluators for the studied 20 projects had done extensive outcome evaluations including control groups. Generally, they had chosen simple and informal methods.

The quality of the evaluations rests strongly on the ability to use a broad approach with a wide range of sources and methods. Most the evaluators did use triangulation to an acceptable extent, but some focused too much on quantitative data and the experiment as the gold standard. For example, one evaluation report devoted more space to discussing the difficulties of quantifying benefit-cost efficiency than to describing it with alternative data.

6.4. The project logic and uncritical evaluators

Reference data, in terms of descriptions of the goal hierarchy or logic model, existed in all projects, but often had weaknesses, which is not uncommon in project evaluations (cf. Samset et al., 2014). The quality-at-entry scheme requires each project to have a defined goal hierarchy. Despite this, more often than not, there were problems such as a missing causal logic or the wrong level of ambition (too high or too low). The evaluators handled this problem in quite different ways. Some re-established the logic model *as it should be*, as basis for their evaluation, while others made only minor or no adjustments. Although evaluators should not 'overrule' the formally agreed goals, we think they should interpret what the project is expected to do and then take this into account when setting scores. A project should not be awarded a score of 6 if its goal was trivial, and likewise, it should not be awarded a score of 1 if the goal was unattainable.

Few, if any, of the evaluators chose a truly theory-based approach to evaluation, which was somewhat surprising. Scholarly literature as well as past evaluations might have been helpful when deciding whether certain changes were likely to be an effect of the project (i.e. the attribution problem). Hatling, Damman, and Halvorsen (2016) concluded that a commonly used assumption in 'co-location' projects, namely that they will automatically generate synergies between residents of the building, is not well grounded in the literature. Kaplan and Garrett (2005) mentioned that a common example of theory failure is to assume that a new technology or infrastructure will make people change their habits without additional measures, such as training and financial incentives. A review of the programme theory could have revealed such a failure, and may similarly reveal redundant project components.

6.5. The evaluation teams

It was required that evaluation teams had no relation to the projects they evaluated. Furthermore, that they had expertise within evaluation, economics and project management, and some knowledge of the sector and type of project. As noted by Scriven (2015), an evaluation team must be broad and represent different perspectives and disciplines, as this is essential for comprehensive and balanced assessments. In our view, the latter was not always satisfied in the studied evaluations, as some of the teams consisted primarily of economists. Only 12 out of the 20 evaluations were performed by sufficiently broad teams. By contrast, 19 out of 20 had high or very high levels of expertise within economics.

6.6. Score-setting – the need for common guidelines

Score-setting was an essential part of the studied evaluations. Our findings indicate that the use of scores is valuable for drawing lessons across projects and sectors. However, experience suggests that efforts should be made to ensure that results are well calibrated. When scores are set by different teams, they may interpret and use the scale differently. A relevant question is whether we could have applied a more objective quantitative summary measure, where scores are obtained from the application of an algorithm that brings the same result independently from the evaluation team (see for instance Chiesa & Frattini, 2007). Unfortunately, we think the answer is no. As long as the framework is used to evaluate different projects with different types of outcomes, different stakeholders, etc., subjective judgements, regarding the choice of indicators as well as score-setting, cannot be avoided. Instead, we believe the solution to ensure calibrated results is clearer guidelines for evaluators. This should be seen in relation to the above-mentioned need for common interpretations of the strategic evaluation criteria, and the need to adjust the goals so that the levels of ambition are realistic across projects.

We also think it is important to be open about the level of uncertainty when it comes to score-setting, as are some evaluators actually, for example, by assigning 'high', 'medium' or 'low' uncertainty next to each score.

7. Conclusions

In Norway, ministries and agencies with large investment projects have become quite good at appraisal and planning. Since the year 2000, the project decision documents have gone through external quality assurance. The assumption is that this will also contribute to improved project performance. However, ex-post evaluations of government investment projects are still rare. Worsley (2014) referred to ex-post evaluations as "the weak link" in the assessment process for transport projects in OECD countries. This is perhaps not surprising. In contrast to, for example, health or educational programs, an infrastructure project cannot be implemented stepwise. Therefore, it could be argued that whereas good planning is crucial, ex-post evaluation is a waste of time and resources. However, that would be an erroneous conclusion because there is much to learn from one project to another, both within and between sectors. Given the poor reputation of public projects in high-income countries in general (Flyvbjerg et al., 2003), the potential to improve project practices is considerable. So is the potential to improve project planning, governance and the quality-at-entry scheme

itself. Evaluation should be based on the project's logic model, as recommended by several authors in the extant evaluation literature (see for instance Samset, 2003). It should ask not only about economic aspects, but take a broad and multifaceted view on project success. In their most recent economic survey of Norway, the OECD (2017), focusing primarily on transport projects, suggests that ex-post evaluation of projects are conducted more systematically, and that a broad framework is applied, to strengthen scope, accuracy and credibility. We have applied a generic evaluation framework inspired by the one recommended by the OECD-DAC for the evaluation of development assistance projects and programs.

A key finding in this study is that most of the projects were rather successful, as considered by the evaluators. This contrasts the public discourse and studies, by Flyvbjerg and others that demonstrate a low level of success in public projects. The 20 projects were highly successful in operational terms, and somewhat more varied in tactical and strategic terms. Some projects scored high on relevance and sustainability, but low on benefit-cost efficiency, and vice versa. This type of deviance needs to be communicated to project owners and various stakeholders, who might have conflicting views on the weighting of the criteria. The evaluations thus provide a basis for discussing whether a better balance between different concerns could have been possible. The possibility to compare, and learn across different sectors, is also considered useful. Some sectors are better at cost control, others at benefits realization, and still others at sustainability, etc.

Our conclusion is that the evaluation results by and large provide a realistic picture of the projects' success. Although the degree of success may seem very high, there is no reason to believe that there is a serious bias on behalf of the evaluators. It is a sample of the country's largest investment projects, which have been through a particularly comprehensive analytical and political process up-front, before they were approved individually by the country's highest authority, i.e. the Parliament.

The evaluators' experiences of the evaluation framework were largely positive. This time evaluation is not limited to aspects of project management success, which has traditionally been the main focus in the project management community. Neither is the framework limited to benefit-cost efficiency, which is normally the main focus in the transport sector. (Other sectors rarely conduct evaluations at all). Instead, the six criteria cover intended and unintended effects alike, goal-oriented and efficiency perspectives, and explicitly raise questions about the long-term viability. Also, this meta-evaluation revealed some improvement points, and the lessons learned should result in a set of requirements and guidelines for future evaluations, regarding how the teams should be put together, how the criteria should be understood, and clear, common principles for score-setting.

One lesson to be drawn is that the evaluation format used in development projects in low-income countries (LIC) is also well-suited in high-income countries (HIC). The reason is that there is no fundamental difference between investment projects in the two types of countries. All projects are implemented to have an impact, and evaluations should be useful for planners, beneficiaries, sponsors and other stakeholders alike. The main difference may be that HICs pay particular attention to projects' value-for-money as measured by the benefit-cost analysis, while in development projects social and ethical justifications may weigh heavier for donors and recipient countries. This has been taken into account in the Norwegian context by expanding the evaluation format with a separate assessment of benefit-cost efficiency.

In evaluations of development assistance projects, the trend in recent years has been to perform larger, strategic and often thematic evaluations, and not only focus on individual projects (OECD, 2016). This approach should be considered for public investment projects in high-income countries too, and we think that our project evaluations would provide useful input to such a broader topic.

In addition to the improved evaluation framework, ministries and agencies need to see the benefits of the evaluations and their learning potential. It is still too early to determine whether these 20 evaluations have led to improved practices, but this will be an interesting topic for future studies. It is well-known that it is more difficult to obtain learning and improvements when evaluations are initiated and conducted by an external party than from internal reviews. However, Reichborn-Kjennerud and Vabo (2017), who studied Norwegian ministries and agencies' ability to learn from performance audits, concluded rather positively. They found that audit reports were often used for improvements in planning and management systems, provided that the reports were found to be relevant, of good quality and sufficiently balanced.

Over time, hopefully, a large number of project evaluations will be produced corresponding to this framework. One ambition is to further improve their quality and ensure that scoring will become better calibrated over time. Since the projects in each sector have similar outcomes, allowing for rather standardized measures, the resulting evaluation database would then provide a valuable basis for robust practices and better determinants of government investment projects' success.

Funding

The work was supported by the Concept Research Programme at the Norwegian University of Science and Technology, which in turn is funded by the Norwegian Ministry of Finance.

Declaration of interest

No conflicts of interest.

Acknowledgements

The author would like to thank all the involved evaluators who contributed to the 20 evaluations as well as in focus group discussions. A special thanks to professor Knut Samset for great inspiration, interesting discussions and useful comments to an earlier draft.

References

ALNAP (2006). *Evaluating humanitarian action using the OECD-DAC criteria: An ALNAP guide for humanitarian agencies.* London: Overseas Development Institute.

Baccarini, D. (1999). The logical framework method for defining project success. *Project Management Institute, 30*(4), 25–32.

Bakewell, O., & Garbutt, A. (2005). *The use and abuse of the logical framework approach: A review of international development NGOs' experiences* [A report for Sida].

Bamberger, M., Rugh, J., Church, M., & Fort, L. (2004). Shoestring evaluation: Designing impact evaluations under budget, time and data constraints. *American Journal of Evaluation, 25*(1), 5–37.

Boardman, A., Greenberg, D., Vining, A., & Weimer, D. (2011). *Cost-benefit analysis: Concepts and practice* (4th ed.). Upper Saddle River, NJ: Pearson.

Brousselle, A., & Champagne, F. (2011). Program theory evaluation: Logic analysis. *Evaluation and Program Planning, 34*(1), 69–78.

Chen, H. T. (1990). *Theory-driven evaluations.* Thousand Oaks, CA: Sage.

Chianza, T. (2008). The OECD/DAC criteria for international development interventions: An assessment and ideas for improvement. *Journal of MultiDisciplinary Evaluation, 5*(9), 41–51.

Chiesa, V., & Frattini, F. (2007). Exploring the differences in performance measurement between research and development: Evidence from a multiple case study. *R&D Management, 37*(4), 283–301.

Connell, J. P., & Kubisch, A. C. (1998). Applying a theory of change approach to the evaluation of comprehensive community initiatives: Progress, prospects and problems. In K. Fulbright-Anderson, A. C. Kubisch, & J. P. Connell (Eds.). *New approaches to evaluating community initiatives. vol. II theory, measurement and analysis* (pp. 15–44). Washington D.C: Aspen Institute.

Eliasson, J., Börjesson, M., Odeck, J., & Welde, M. (2015). Does benefit-cost efficiency influence transport investment decisions? *Journal of Transport Economics and Policy, 49*(3), 377–396.

European Commission (2013). *EVALSED: The resource for the evaluation of socio-Economic development –Evaluation guide.* [Downloadable from: http://ec.europa.eu/regional_policy/en/information/publications/evaluations-guidance-documents/2013/evalsed-the-resource-for-the-evaluation-of-socio-economic-development-evaluation-guide. (Accessed 26 May 2017)].

European Commission (2015). *Guidelines on evaluation and fitness checks.* [http://ec.europa.eu/smart-regulation/guidelines/ug_chap6_en.htm. (Accessed 26 May 2017)].

Flyvbjerg, B., Skamris Holm, M. K., & Buhl, S. L. (2003). How common and how large are cost overruns in transport infrastructure projects? *Transport Review, 23*(1), 71–88.
Flyvbjerg, B., Garbuio, M., & Lovallo, D. (2009). Delusion and deception in large infrastructure projects: Two models for explaining and preventing executive disaster. *California Management Review, 51*(2), 170–194.
Flyvbjerg, B. (2014). What you should know about megaprojects and why: An overview. *Project Management Journal, 45*(2), 6–19.
Gasper, D. (2000). Evaluating the 'logical framework approach' towards learning-oriented development evaluation. *Public Administration and Development, 20*, 17–28.
Glass, G. (1976). Primary, secondary and meta-Analysis of research. *Educational Researcher, 5*.
Green, J., Roberts, H., Petticrew, M., Steinbach, R., Goodman, A., Jones, A., et al. (2017). Integrating quasi-experimental and inductive designs in evaluation: A case study of the impact of free bus travel on public health. *Evaluation, 21*(4), 391–406.
HM Treasury (2011). *The magenta book. guidance for evaluation*.
Hatling, M., Damman, S., & Halvorsen, T. (2016). *Samlokaliseringseffekter – hva sier litteraturen? En oversikt over resultater fra relevante studier*Trondheim: SINTEF [SINTEF Report no A27652].
Heinzerling, L., & Ackerman, F. (2002). *Pricing the priceless: Cost benefit analysis of environmental protection.* Washington, D.C: Georgetown Environmental Law and Policy Institute.
House, E. (2000). The limits of cost benefit evaluation. *Evaluation, 6*(1), 79–86.
Jacob, S., Speer, S., & Furubo, J. E. (2015). The institutionalization of evaluation matters: Updating the International Atlas of Evaluation 10 years later. *Evaluation, 21*(1), 6–31.
Kaplan, S., & Garrett, K. (2005). The use of logic models by community-based initiatives. *Evaluation and Program Planning, 28*, 167–172.
Karlsen, T., & Jentoft, N. (2013). Programteori versus samfunnsvitenskapelig teori. In A. Halvorsen, E. L. Madsen, & N. Jentoft (Eds.). *Evaluering: Tradisjoner, praksis og mangfold* (pp. 164–180). Bergen: Fagbokforlaget.
Kvalheim, E. V., Christensen, T., Samset, K., & Volden, G. H. (2015). *Har regjeringen fått et bedre beslutningsgrunnlag? Om effekten av å innføre konseptvalgutredning (KVU) og ekstern kvalitetssikring (KS1 og KS2) for store statlige investeringsprosjekter. [Concept working paper].* [https://www.ntnu.no/documents/1261860271/1262021752/4200+Regjeringens+beslutningsgrunnlag+sluttrapport.pdf/ff529615-f4c4-4b9d-a561-1d4790f7faea. (Accessed 26 May 2017)].
Love, P. E. D., Smith, J., Simpson, I., Regan, M., & Olatunji, O. (2015). Understanding the landscape of overruns in transport infrastructure projects. *Environment and Planning B: Planning and Design, 42* [490–209.
McLaughlin, J. A., & Jordan, G. B. (1999). Logic models: A tool for telling your programs performance story. *Evaluation and Program Planning, 22*(1), 65–72.
Morris, P. W. G., & Hough, G. H. (1991). *The anatomy of major projects: A study of the reality of project management.* Chichester, UK: John Wiley & Sons.
Morris, P. W. (2013). *Reconstructing project management.* Chichester, UK: John Wiley & Sons.
Norwegian Ministry of Finance (2015). Meld. St. 3 (2015-2016). Statsrekneskapen 2015.
Nyborg, K. (1998). Some Norwegian politicians' use of cost-benefit analysis. *Public Choice, 95*, 381–401.
OECD (1991). *DAC principles for evaluation of development assistance.* Paris: OECD, Development Assistance Committee.
OECD (2002). *Glossary of key terms in evaluation and results based management.* Paris: OECD, Development Assistance Committee.
OECD (2016). *Evaluation systems in development Co-operation.* Paris: OECD [2016 Review].
OECD (2017). *OECD economy surveys. Norway 2018.* Paris: OECD.
Pawson, R., & Tilley, N. (1997). An introduction to scientific realist evaluation. In E. Chelimsky, & W. Shadish (Eds.). *Evaluation for the 21 st century: A handbook* (pp. 405–418). Thousand Oaks, CA: Sage.
Picciotto, R. (2013). The logic of development effectiveness: Is it time for the broader evaluation community to take notice? *Evaluation, 19*(2), 155–170.
Project Management Institute (2013). *A Guide to the project management body of knowledge (PMBOK® Guide)* (5th ed.). PMI Inc.
Rambøll, & Agenda Kaupang (2016). *Bruk av evaluering i statlig styring* [Report R9392. https://evalueringsportalen.no/evaluering/bruk-av-evaluering-i-statlig-styring/Rapport%20DF%C3%98%20Bruk%20av%20evaluering%20i%20statlig%20styring.pdf/@@inline. (Accessed 26 May 2017)].
Reichborn-Kjennerud, K., & Vabo, S. I. (2017). Performance audit as a contributor to change and improvement in public administration. *Evaluation, 23*(1), 6–23.
Rogers, P., Petrosino, A., Huebner, T., & Hacsi, T. (2000). Program theory evaluation: Practice: promise and problems. *New Directions for Evaluation, 87*, 5–13.
Rosenberg, L., Posner, L. D., & Hanley, E. J. (1970). *Project evaluation and the project appraisal reporting system* [Final report submitted to the USAID, by Fry consultants incorporated].
Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach* (7th ed.). Thousand Oaks: Sage.
Samset, K., & Christensen, T. (2017). Ex ante project evaluation and the complexity of early decision-making. *Public Organization Review, 17*(1), 1–17.
Samset, K., & Volden, G. H. (2015). Front-end definition of projects: Ten paradoxes and some reflections regarding project management and project governance. *International Journal of Project Management, 34*(2), 297–313.
Samset, K., Andersen, B., & Austeng, K. (2014). To which extent do projects explore the opportunity space? A study of conceptual appraisals and the choice of conceptual solutions. *International Journal of Managing Projects in Business, 7*(3), 473–492.
Samset, K. (2003). *Project evaluation: Making projects succeed.* Trondheim: Tapir Akademisk Forlag.
Sartorius, R. H. (1991). The logical framework approach to project design and management. *Evaluation Practice, 12*(2), 139–147.
Scriven, M. (2015). *Key evaluation checklist (KEC).* [http://michaelscriven.info/papersandpublications.html].
Stufflebeam, D. (2010). Meta-evaluation. *Journal of Multi-Disciplinary Evaluation, 7*(15), 99–158.
Sullivan, H., & Stewart, M. (2006). Who owns the theory of change. *Evaluation, 12*(2), 179–199.
Ton, G. (2012). The mixing of methods. A three-step process for improving rigour in impact evaluations. *Evaluation, 18*(1), 5–25.
van Wee, B. (2007). Large infrastructure projects: A review of the quality of demand forecasts and cost estimations. *Environment and Planning B: Planning and Design, 34*, 611–625.
Volden, G. H., & Samset, K. (2015). Perverse incentives in the front-end: Public funding and counterproductive projects. *Paper at IRNOP conference*.
Volden, G. H., & Samset, K. (2017). Quality assurance in megaproject management: The Norwegian way. In B. Flyvbjerg (Ed.). *The oxford handbook of megaproject management* (pp. 406–427). Oxford: Oxford University Press.
Weiss, C. (1997). Theory-based evaluation: Past present, and future. *New Direction for Evaluation, 76*, 41–55.
Welde, M. (2017). *Kostnadskontroll i store statlige investeringer underlagt ordningen med ekstern kvalitetssikring*Trondheim: Ex Ante Academic Publisher [Concept Report No 51].
Wiig, H., & Holm-Hansen, J. (2014). *Unintended effects in evaluations of norwegian aid: A desk study*Oslo: Norwegian Agency for Development Cooperation [NORAD] [Norad Report 2/2014].
Williams, T., & Samset, K. (2010). Issues in front-end decision making on projects. *Project Management Journal, 41*, 38–49.
World Bank (2004). *Monitoring and evaluation: Some tools, methods and approaches.* Washington D.C: World Bank.
Worsley, T. (2014). *Ex post assessment of transport investments and policy interventions: Prerequisites for ex-post assessments and methodological challenges. ITS roundtable summary and conclusions.* [International transport Foundation Discussion Paper 2014–19. https://www.econstor.eu/bitstream/10419/109153/1/818314141.pdf. (Accesssed 26 May 2017).
Yin, R. (2013a). *Case study research Design and methods.* Sage.
Yin, R. (2013b). Validity and generalization in future case study evaluations. *Evaluation, 19*(3), 321–332.

Gro Holst Volden holds the position of research director of the Concept Research Programme on Front-end Management of Major Investment Projects, at the Norwegian University of Science and Technology. Her research is within project governance, public decision processes, and appraisal and evaluation of major public investments. She in an economist from the Norwegian School of Economics, and has a prior career as a senior advisor in the consulting industry and in the government agency for financial management. Volden is currently president of the Norwegian Evaluation Society.