Measuring efficiency and effectiveness through ex-post evaluation: Case studies of Norwegian transport projects

Morten Welde, Gro Holst Volden
NTNU – Norwegian University of Science and Technology, Department of Civil and Environmental Engineering, Trondheim, Norway

1. Introduction

Governments throughout the world expend large amounts on the planning and implementation of transport infrastructure, usually organised as large projects. Most countries have comprehensive appraisal methods for costs and benefits, and most have national guidelines for cost–benefit analysis (CBA) (Mackie et al., 2014).

The aim of CBA is to maximise economic efficiency, which means the maximisation of social welfare given alternative project rankings. This is normally referred to in the literature as *allocative efficiency* or *Pareto efficiency*, namely the allocation of goods in cases when no other allocation can make at least one person better off without making anyone else worse off. In practice, actual Pareto efficiency would result in society forgoing many policies that offer positive net benefits. CBA is thus based on the Kaldor-Hicks criterion, which states that a policy should be adopted if, and only if, those who would benefit would be able to compensate fully those who would lose and still be better off. Following this principle would mean that we only adopt policies that have positive net benefits (Boardman et al., 2006 p. 31). However, there are no requirements for such compensation to be given.

In practice, decision-makers pursue different policies that may or may not be in accordance with the maximisation of social welfare. Projects may be tailored to achieve specific goals, such as equality, national security, greenhouse gas reduction, expenditure constraints, and political feasibility. There may, for example, be good reasons to work towards Vision Zero in the transport system (see, e.g., Trafikverket, 2017), with no fatalities or serious injuries related to road traffic, even if the costs of realising the policy may be higher than the benefits measured by the total willingness to pay. Realising Vision Zero may not be efficient in economic terms, but it may be the right thing to do from a political or ethical perspective.

Achieving such goals may support economic efficiency, but not necessarily. Issues such as the distribution of wealth and protecting the environment may be desirable goals but they have little or no value in an economic appraisal. Many countries have therefore adopted a business case approach that is aimed at capturing all reasons for carrying out a project and thereby help decision-makers to ensure that a proposed initiative not only provides value for money but also is in line with relevant policies. *Effectiveness* measures the ability to achieve a specific goal or output (Yu, 2008; Førsund, 2017). Whereas economic efficiency may be regarded as the most important success criterion from a societal point of view, this is an aggregated parameter. There are large variations in how success is defined and interpreted, both ex-ante and ex-post (Samset, 2003).

The purpose of this paper is to demonstrate that project success may be multifaceted and that the success of projects could be assessed through a broad framework that has been applied in the context of ex-post evaluation. We argue that although economic efficiency is an important general

success criterion for transport projects, there may be other and equally important measures of success. The paper briefly presents the results of evaluations of 12 road and rail projects that illustrate the benefits of a broad approach to ex-post evaluation.

The paper proceeds as follows. Section 2 includes a brief review of the literature on ex-ante appraisal and ex-post evaluation. In Section 3, we discuss the usefulness of CBA in project selection and argue that if CBA is not used ex-ante, ex-post evaluation should be based on a broader approach. Section 4 presents the framework for ex-post evaluation used in Norway. In Section 5 we present the evaluated projects and the results of the evaluations. Lastly, in Section 6 we summarise the paper and present some conclusions.

2. Evaluation of transport projects

Transport projects have long time horizons. The costs of implementation up front are usually large, but the results will have an impact on users and society over many years. Today, we benefit from investments made many decades ago or, even more than 100 years ago. The estimation of the effects of projects therefore requires some form of prediction or ex-ante appraisal.

The appraisal of projects normally includes a social CBA, in which the purpose is to inform decision-makers about transport projects' estimated value for money. In CBA, the idea is to determine relative weights of different types of benefits through citizens' preferences, as opposed to, for example, decision-makers' or planners' preferences (Eliasson, 2014). As CBA measures the relative economic efficiency of projects, it is a potentially useful tool for ranking projects.

However, CBA results are often questioned because they may depend on uncertain assumptions about the future and on methodologically uncertain valuations of costs and benefits (Börjesson et al., 2014). Mackie and Preston (1998) listed 21 sources of error and bias in appraisals, and concluded that appraisal optimism is the greatest risk in transport investment analysis. Appraisal optimism happens because scheme promoters may, deliberately or unwittingly, bias the appraisal. Hence, the real outcome of projects may not as positive as that presented ex-ante.

One of the main avenues through which bias could enter the appraisal is the traffic forecasts. If real traffic levels deviate significantly from forecasts, this will ultimately affect the estimated economic benefits and, potentially, the ranking of projects. The consequences of inaccurate traffic forecasts depend on the context within which the new facility is built. In uncongested conditions, underestimated traffic will imply underestimated economic benefits. If congestion is or will be a problem during the appraisal period, underestimated traffic may imply a shorter period of relief from congestion and hence an overestimation of benefits. Despite the crucial role of traffic forecasts, ex-post studies are relatively rare. Nicolaisen and Driscoll (2014) surveyed 12 studies of forecast accuracy in road and rail projects in different countries from the 1970s to the present. They found that the mean inaccuracy for road projects was typically positive, indicating that more demand than expected materialised after the projects had been completed. By contrast, the mean inaccuracy for rail projects was negative. The authors concluded that the relatively large range within which traffic forecasts fall represents a challenge for the use of travel demand forecasts as decision support.

The investment cost is typically the parameter that attracts the most attention throughout both the appraisal and the implementation phase of projects. Cost overruns attract considerable interest in most countries. Flyvbjerg et al. (2002) showed that overruns are a problem across countries and that they have been throughout history. Odeck (2017) reviewed 48 studies from different continents and found that the average overrun was 34 per cent, but that recent studies had showed improved cost performance compared with earlier studies. All other things being equal, this will reduce the net benefits from projects.

Although a number of studies have documented that crucial input parameters in appraisals are inaccurate and that this may bias the decision-making process, comprehensive ex-post evaluation of transport projects is rare. We use more resources on how we think a scheme might perform than on demonstrating how it actually has performed. Our knowledge of projects' economic efficiency, effectiveness, and other impacts is limited. The International Transport Forum (ITF) referred to the lack of meaningful ex-post evaluations as the weak link in the assessment of transport infrastructure and policy (International Transport Forum, 2017).

There are several reasons for the lack of meaningful ex-post evaluations. 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. Personnel, organisations, and systems change over time, and data become more difficult to obtain on time. Hence, it may take a long time for the full impacts to be seen. It can take 10–15 years from project initiation to completion, and in most countries the appraisal period covers up to 60 years. This means that successful evaluations, in which results are used and fed into the appraisal process, require a long-term commitment, a strong culture of evaluation, and effective governance (Department for Transport, 2015). Nicolaisen (2012) suggested a further reason why ex-post evaluation may be difficult, namely the lack of transparency. He found it difficult to obtain information about model specifications, key assumptions, and data sources, thus making it difficult to subject appraisals to critical scrutiny.

Nicolaisen and Driscoll (2016) reviewed the literature and found that although there are a number of examples of ex-post reviews of projects carried out by transport authorities, auditing authorities, or lending institutions, there is a striking lack of standardised methods for conducting ex-post evaluation at the national level, which thus inhibits comparability and learning. A noticeable exception is the Post Opening Project Evaluation (POPE) developed by Highways England (HE) in the UK, which is undertaken for all of the HE's major schemes. POPE studies are undertaken for each major scheme one year and five years after opening. The purpose is to assess whether the schemes have delivered value for money and whether traffic forecasts and cost estimates have been accurate. The POPE studies also assess whether schemes have achieved their objectives. The latest summary report (Highways England, 2016) shows that most schemes achieved their objectives, traffic forecasts were accurate, cost performance was acceptable, and that most schemes offered good value for money. In addition to the evaluation of major schemes, the UK Department for Transport requires evaluation of locally delivered projects with central funding (Atkins and AECOM, 2014).

Similarly, the Norwegian Public Roads Administration (Statens vegvesen) subjects 3–5 road projects to ex-post CBA annually. The purpose is to verify whether the estimated costs and the quantified benefits deviate from real outcomes. Thus far, the results indicate that the original CBAs have been based on conservative estimates, as 20 out of 25 projects showed an improvement in net present values compared with the original analyses. The main reason for higher benefits was that the rate of traffic growth had been higher than forecasted and that the frequency of accidents had been lower. The majority of the projects were completed with lower construction costs than estimated. Kjerkreit and Odeck (2015) thus concluded that the CBAs presented to decision-makers were not over-optimistic but rather erred on the pessimistic side.

In common with the UK and Norway, France has a framework for ex-post evaluation through its 'permanent observatories', which are established by law and used to collect data to facilitate detailed evaluations of major transport schemes. For the largest projects, an observatory is established at the same time the project is approved, so that 'before' data can be collected. The purposes of the evaluations are to identify discrepancies between forecasts and actual outcomes, to assess value for money in completed schemes, and to provide feedback to improve the methods for ex-ante appraisal. The main lessons learnt thus far are that cost overruns and inaccurate traffic forecasts are common in rail projects (International Transport Forum, 2017; Meunier and Welde, 2017).

3. The usefulness of CBA as an ex-post evaluation tool

In a perfect world, CBA would capture all relevant effects of projects. However, the distribution of effects between groups, and the uncertainty of measuring and valuing impacts for a long time into the future have resulted in many questioning the role of CBA as an appraisal tool (Jones et al, 2014). Several studies have shown that CBA results seem to have limited impact on project selection, even in countries claiming to put a large weight on appraisal results (Jansson and Nilsson, 1989; Nilsson, 1991; Fridstrøm and Elvik, 1997; Odeck, 1996, 2010; Nellthorp and Mackie, 2000; Laird et al., 2012). Political decision-makers do not always trust the normative basis for CBA and may use the results opportunistically to the extent that they support their own opinions (Nyborg, 1998; Sager and Ravlum, 2005; Mouter, 2017). This is especially true for Norway, where Eliasson et al. (2015) found that project selection was uncorrelated with both the benefits and costs included in the CBA, even when the initial rankings were carried out by planners and not by politicians. There is thus ample evidence to conclude that even if CBA is useful as decision support, it is not necessarily comprehensive in that that decision-makers may have aspirations beyond those that can be captured by the CBA.

The use and status of CBA varies between countries (Annema et al., 2017). In Norway, projects have been routinely implemented with poor or negative value for money. In the national transport plan for 2018–2029 that was presented in the Norwegian parliament in the spring of 2017, the net effects of the projects planned to start during the period were estimated to be minus 179 billion NOK[1]. While topography and geographical population patterns make it harder to achieve positive CBAs in Norway than in most other western countries, this also suggests that the decision-makers

[1] 1 NOK ≈ 0.10 EUR as of April 2018.

value other effects than those included in the CBA or that they may value the elements in the CBA more than the net outcome.

In reality, both project promoters and decision-makers are often concerned with a range of issues that may or may not be included in the CBA.

Hence, the effectiveness of projects in terms of the achievement of stated outcome goals may be a more important success criterion than economic efficiency, as measured by the net present value or other metrics. Effectiveness measures the degree to which something is successful in producing a desired result. Projects are designed to produce a range of benefits, which may be summarised through the CBA, but in order to gauge the success of projects ex-post, more comprehensive evaluation is needed.

4. Framework and data

To capture a broader range of issues relevant to decision-makers and other stakeholders, we have evaluated Norwegian transport projects ex-post by using a generic framework that consists of six evaluation criteria, which together cover different aspects of project success.

For a presentation of the evaluation framework and a more thorough discussion of the motivation for applying it to public infrastructure projects in high-income countries such as Norway, see the published literature by Samset (2003), Volden (2017) and Volden and Samset (2017a). The original six evaluation criteria have been restructured into four criteria for the purpose of this article. The evaluations are based on a goal-oriented framework originally developed by the OECD Development Assistance Committee and commonly referred to as the OECD Development Assistance Committee's evaluation criteria or logframe evaluation (OECD Development Assistance Committee, 1991). The European Investment Bank uses a similar approach in its evaluations of transport projects (European Investment Bank, 2005). In a logframe evaluation, the project is described in terms of a chain of goals on various levels, from output goals (that concern the delivery of the project), to outcome goals (also referred to as target benefits by Chih and Zwikael, 2015), and finally the societal objective, which describes a desired societal development. A goal-oriented framework is well suited for projects, which by definition are phenomena that are limited in time and scope, especially when the projects are well defined with formally agreed-upon goals on several levels, as is normally the case for transport investment projects. The framework is generic and therefore applicable to all types of investment projects, not only those within transport.

The CBA measures economic efficiency. Ideally, this should capture all relevant effects, but as we have seen, its practical use suggests that other issues may be equally or more important than the net present value. For road and rail projects, effectiveness may be measured by the achievement of typical outcome goals for accessibility, time savings, safety, security, environmental impacts, or job creation. These are normally elements in the CBA, but their monetised effects may not always exceed the costs of implementation. Nonetheless, decision-makers and other stakeholders may regard a project as successful if it delivers the stated outcome goals. Instead of focusing the ex-post evaluation solely on the CBA, which decision-makers may not care much about, our evaluation framework is aimed at measuring issues that are relevant to both users and decision-makers. In addition to efficiency and effectiveness, the evaluation framework measures cost performance, and

assesses long-term strategic success. The latter includes (1) relevance (public support and alignment with political-normative objectives and requirements), (2) other impacts (short-term and long-term) beyond the intended effects, and (3) sustainability, which includes economic, environmental, and social issues in the long term (see Haavaldsen et al., 2014 for a definition of sustainability). The criteria as used in the ex-post evaluation of transport projects in Norway are summarised in Table 1.

Table 1: Evaluation criteria used in ex-post evaluations

Economic efficiency	- Did the project deliver value for money as summarised by the net present value or benefit-cost ratio?
Effectiveness	- Were the stated outcome goals achieved? - To what extent did the project contribute to the goal achievement?
Cost performance	- To what degree was the project delivered within the agreed scope and budget?
Strategic success	- Relevance: Is the project aligned with the government's strategic objectives for the transport sector? - Is it in line with the needs and priorities of different stakeholders? - Other impacts: Has the project had any other positive or negative impacts other than those planned? - Sustainability: Are the positive effects derived from the project likely to continue - and the negative effects acceptable – in the long-run?

Although the OECD criteria are widely accepted and implemented in evaluations of international development projects and increasingly in other areas (Chianca, 2008), their use in transport has been limited. However, the increasingly broad strategic scope of transport investments suggests that ex-post evaluations should adopt a similar perspective. The evaluation criteria listed in Table 1 are broad and must be operationalised to fit individual projects. For example, political support and public acceptance are typically issues that need to be covered in the assessment of strategic success.

The evaluation framework is managed by the Concept Research Programme,[2] which is financed by the Norwegian Ministry of Finance and does research on large government investment projects that have been subjected to the Ministry of Finance's Quality Assurance regime. The Concept Research Programme is independent from the responsible agencies and is located at the Norwegian University of Science and Technology. In addition to the transport projects presented in this paper, the framework has been tested on a larger set of projects from various sectors (transport, defence, ICT, and buildings) (Volden, 2017).

[2] See https://www.ntnu.edu/web/concept/ for details.

Large government investment projects in Norway are subjected to a governance framework that is similar across sectors. Projects are developed in stages and in cases when the reasons for implementation are well documented before project approval (Volden and Samset, 2017b). The business case for each project describes the intended output, user benefits, and strategic objectives, and provides a benchmark that the results can be mapped against.

The evaluation process is identical for all evaluations and is as follows:

1) A project is selected. Projects eligible for ex-post evaluation have been subject to the Ministry of Finance's Quality Assurance (QA) regime, implying external scrutiny of business cases. They must have an estimated cost above EUR 55 million and the resulting infrastructure must have been in operation for at least three years.
2) A multidisciplinary evaluation team is established to carry out each evaluation. The budgets usually allows for approximately three person-months in total.
3) The team reviews and, if necessary, adjusts the goal structure of the project so that results can be compared with project goals, and then breaks down the evaluation criteria into more specific evaluation questions and indicators.
4) The team collects and analyses data that can provide answers to the evaluation questions and indicators.
5) The team summarizes its assessment for each of the criteria on a scale ranging from 1 to 6. Score-setting is based on common guidelines for the evaluators (Concept, 2017).
6) The result is a report of usually 60-100 pages plus appendices.

The evaluations are based on both quantitative and qualitative sources. In Norway the transport agencies regularly collect a wide range of quantitative data that are available to researchers. Although data on construction costs, time savings and traffic safety effects may be relatively readily available, other effects may take longer to materialise or be less well documented. Quantitative data are always supplemented with interviews and observations. Triangulation of information allows data validation and gives more insight into the issues that are investigated.

The evaluation criteria listed in Table 1 can partly be measured quantitatively, but they also depend on the assessment of different evaluators. This requires personal judgement, which could be regarded as potential source of bias. However, supported by the guidelines for evaluation and experiences from past evaluations, the variation in score setting for identical results is gradually reduced.

There are many advantages of systematic ex-post evaluation based on a common framework. It provides knowledge of actual scheme performance. Although the operational phase of transport projects lasts decades, the first years after opening provide much greater certainty about costs and benefits than can be estimated before project approval and completion. It is increasingly often acknowledged that quality at entry through comprehensive ex-ante appraisal is crucial to project success. The quality of cost estimates, traffic forecasts, and other assumptions will undoubtedly improve as ex-post evaluation is integrated into a project governance framework. A final advantage is that the evaluations include non-monetised impacts and unintended effects that may not have been included in the original CBA.

Thus far, 12 transport projects have been evaluated using the evaluation framework. Together they should be reasonably representative of large road and rail projects carried out in Norway during the last 15 years and subjected to external QA. The evaluated projects are presented in Table 2 (see Volden and Samset, 2017a, and the references therein for details on the individual evaluations).

Table 2: The evaluated projects

Project name	Type of project		Size	Opening year
	Description	Budget at the time of decision to build (million EUR)	Length (km)	Year
E6 Riksgrensen-Svingenskogen	Bridge and highway dualling	90	5.0	2005
E18 Momarken-Sekkelsten	Highway dualling	55	6.2	2007
E16 Kløfta-Nybakk	Highway dualling	70	10.5	2007
E10 Lofast	New road	125	29.9	2007
E6 Svingenskogen-Åsgård	Highway dualling	30	33.0	2008
E6 Åsgård-Halmstad	Highway dualling	95	11.0	2008
Rv 653 Eiksundsambandet	Subsea road tunnel	95	13.7	2008
Fv 519 Finnfast	Subsea road tunnel	60	8.4	2009
Sandvika-Asker	Doubling of a single railway line	460	9.5	2005
Stavanger-Sandnes	Doubling of a single railway line	210	14.5	2009
Barkåker-Tønsberg	Doubling of a single railway line	165	5.8	2011
Gevingåsen	Railway tunnel	88	5.4	2011

5. Evaluation results

The results of the evaluated projects, with scores from 1 (unacceptable), through 4 (acceptable) to 6 (excellent), are presented in Table 3. The top eight projects are road projects and the bottom four are rail projects. The performance of each project was assessed in the individual evaluation reports and the scores provided by the evaluators based on the common guidelines.

Table 3: Summary of evaluation results

	Economic efficiency	**Effectiveness**	**Cost performance**	**Strategic success**
E6 Riksgrensen-Svingenskogen	6	6	5	5
E18 Momarken-Sekkelsten	6	5	3	4

E16 Kløfta-Nybakk	4	5	4	4
E10 Lofast	2	5	4	3
E6 Svingenskogen-Åsgård	6	5	4	4
E6 Åsgård-Halmstad	6	5	5	4
Rv 653 Eiksundsambandet	5	6	5	5
Fv 519 Finnfast	6	5	5	5
Sandvika-Asker	2	2	5	5
Stavanger-Sandnes	2	4	4	4
Barkåker-Tønsberg	2	3	4	3
Gevingåsen tunnel	2	3	4	4

5.1 Economic efficiency

The purpose of the economic appraisal is to calculate the economic efficiency through a social CBA. As discussed above, although Norway uses large resources on CBAs throughout the early appraisal of projects, the practical impact on project ranking has been limited. Projects with positive net present values have not been more likely to be selected than projects with negative value for money (Eliasson et al., 2015). In our sample, only four projects were estimated to have given value for money ex-ante, and this confirms that decision-makers may have had other reasons for approving projects than economic efficiency alone.

It should be noted that the ex-post economic assessment is not as thorough as the ex-ante assessment. Due to resource constraints, the ex-post methodology is simpler and does not include all effects. However, it does include the main effects and provides an accurate estimate of whether the projects have provided value for money.

The limited use of CBA in decision-making may have had an unintended consequence of pessimism bias. As planners have little incentive to exaggerate benefits, CBAs of road projects have been cautious, and benefits have been underestimated rather than overestimated. In line with the results of Kjerkreit and Odeck (2015), the majority of the road projects that have been evaluated to date have had better value for money than estimated ex-ante. The majority of the road projects had estimated net benefit–cost ratios of 0.5–2.0. Five projects were estimated to have negative net present values ex-post, while only two of the eight projects were estimated to be profitable ex-ante. An interesting result is that all four rail projects had very low value for money, and that their CBAs ex-ante seem to have been positively biased. When significant increases in passenger numbers and reductions in travel times have failed to materialise, this has resulted in the costs being been realised but not the benefits.

The main reason why the road projects have performed better is higher traffic volumes, lower accident frequencies, and less severe accidents than forecasted. By contrast, the rail projects have struggled to deliver their user benefits. So far, increases in passenger numbers, time savings, train

frequencies, and other planned effects have failed to materialise. This is in line with findings in ex-post evaluations of European high-speed rail projects, which in general have been found to be lower value for money than predicted ex-ante (de Ruz, 2012; Crozet, 2013) and that mode substitutions from aircraft, car, and coach have been modest (Givoni and Dobruszkes, 2013).

5.2 Effectiveness

Effectiveness measures the extent to which the project managed to achieve its stated goals. For road projects, typical short-term goals are reduced travel times and improved traffic safety. Some projects, in which reductions in generalised costs are large, may additionally be aimed at regional development. Rail projects are justified on the basis of planned user effects, but also aim at modal shifts from car to rail and at reduced emissions.

All of the evaluated road projects achieved most of their goals, but some of the goals were not very ambitious. Given a certain traffic volume, reduced travel time is easily achieved once a road with increased capacity is opened. In some cases, where ferries have been replaced by tunnels or bridges, travel times have been reduced by up to 60 minutes. With regard to the wider benefits (agglomeration, increased competition, and better functioning labour markets), the projects show varying results. One of the subsea road tunnels, Rv 563 Eiksundsambandet, has clear indications of wider benefits. A possible explanation could be that the project connected two sufficiently large regions with different business structures, such that they complemented each other.

The rail projects were less effective in achieving their goals, but their goals were also more ambitious, since goal achievement relied not only on further investment in the rail network, but also on the development in other markets, such as road improvements and the price and availability of parking. The evaluations concluded that some of the ex-ante assumptions in the rail projects had been unrealistic. The rail authority had an infrastructure perspective rather than a user-benefit perspective. Instead of focusing on what investments and which changes were needed to realise benefits for users and society, the rail authority focused on the physical infrastructure and simply assumed that the benefits would be realised immediately after the new facility opened. The ex-post evaluations found that, due to a lack of railway switches, passing loops, or trains, the goals had largely not been achieved. However, once these measures are in place, as they are planned to be sometime in the future, the railway projects will probably deliver as originally planned.

5.3 Cost performance

During the last 15-20 years, Norway has taken significant steps towards improved estimate accuracy and project management. Cost estimates are prepared using stochastic cost estimation and estimates are scrutinised by external consultants before the Norwegian parliament approves the final budget. These improved practices have led to better cost performance. . Supporting evidence is provided by our evaluations. Of the 12 projects, 8 were completed below budget,[3] which was typically set at or just below the P85 percentile. The P85 is an estimate of the project cost based on an 85 per cent probability that the cost will not be exceeded. Half the projects were delivered at or below the P50. Additionally, most of the projects were delivered on time and with the agreed quality. While optimism bias seems to have played a role in the estimation of benefits of the rail

[3] The sample mean percentage error, which is the measure of cost performance used in most international studies, was minus 7.3 per cent.

projects, the cost performance of rail projects was identical to that of the road projects. Four out of the five rail projects were completed on or below budget. Similar results have been found in other studies of the cost performance of Norwegian projects that have been subjected to external QA (Odeck et al., 2015; Welde, 2017). Figure 1 shows the distribution of final costs to budgets in the 12 projects.

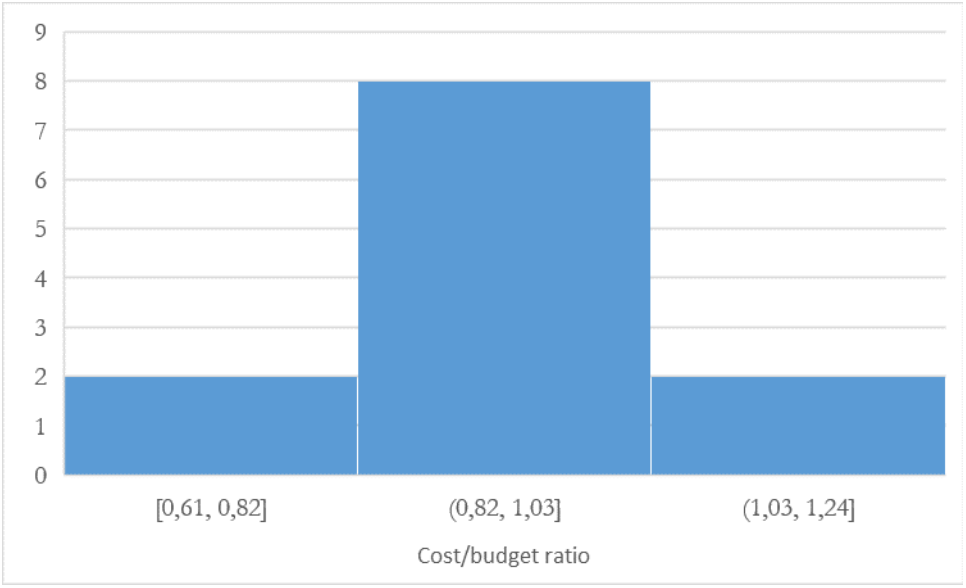

Figure 1: Distribution of final costs to budgets for 12 projects

From our evaluations of individual projects and from previous studies of cost performance with larger samples (Odeck, 2014; Odeck et al., 2015; Welde, 2017), it appears that cost estimates in Norwegian transport projects are not systematically underestimated, as has been suggested internationally. The majority of the evaluated projects had final costs slightly below budget. However, as transport projects go through several stages before formal project approval, they are well developed once the final budget is approved. This means that cost escalation, as is common in many transport projects, may have incurred in the front end (i.e. before the formal decision to build) (Welde and Odeck, 2017).

5.4 Strategic success

Although short-term user effects are an important argument for transport projects, such projects should have a long-term justification by supporting strategic objectives and should be aligned with the government's strategy for the rest of society. The projects should also be sustainable by being able to deliver benefits long into the future. There is, for example, no point in transport investment to relieve congestion if traffic growth will quickly fill the spare capacity. Furthermore, there should not be any significant negative external effects to society, the environment, and other non-users of the new facility. In the evaluations, the assessment of strategic success comprised three subcriteria: relevance, other impacts, and sustainability.

Relevance: The evaluations revealed that the projects have been largely successful from a strategic perspective, thus indicating that there had been a need for the effects that the projects produced. The Norwegian government is working to upgrade the parts of the road network that is considered substandard and will thereby improve accessibility and promote economic growth. Since 2008, there have been huge improvements in traffic safety (since 2000, the number of road deaths in

Norway has reduced from 350 per year to ca. 100 per year) and road and rail improvements are an important measure for realising the long-term vision of zero road deaths.

Other impacts: Transport projects may have negative external effects. Many of these are not monetised and included in the economic appraisals. Some are monetised in the CBA, but given a very low calculation price that is far from reflecting the political goals, for example related to reduction in greenhouse gas emissions. Generally, the impacts beyond the intended effects were not considered to be substantial, although negative side-effects have been reported for several road projects such as increased pollution, impacts on farmland and landscape, and increased car dependency.

Sustainability: There is no reason to expect any significant reduction in traffic levels, but most Norwegian roads outside major cities have sufficient capacity to cater for future traffic growth. However, from a climate perspective, road traffic at present levels could be considered a challenge, yet providing public transport in rural areas is not possible and would not necessarily be a more environmentally sustainable alternative.

6. Concluding remarks

Cost-benefit analysis is an important tool for appraisals of transport projects, but the actual selection of projects is not always based on its results Firstly, the estimates of the various costs and benefits may be inaccurate, implying that ex-post analyses are needed both to determine whether deviations are biased, and as a basis for improving estimation tools and methods. Secondly, the CBA may be too narrow as measure of strategic project success. An indication of the latter is that the actual selection of projects is not always based on its results. In reality, decision-makers take a range of issues, both monetised and non-monetised, into consideration when making their decisions. Effectiveness, or the ability to achieve specific goals, may in some cases be more important than economic efficiency. This implies that evaluation frameworks (regardless whether they are applied ex-ante or ex-post) should include multiple success criteria. Most countries use large resources on the planning and appraisal of projects. Many studies have gauged the accuracy of estimates of monetised impacts used in appraisals, but comprehensive ex-post evaluations aimed at assessing the success of projects are not common. Countries where ex-post evaluation is mandatory focus on the monetised effects and may thus fail to capture effects that are not included in economic appraisals.

In this paper, we have argued that ex-post evaluation is an important tool for providing insights into the performance and outcomes of transport infrastructure and decision-making, and for informing the public. As discrepancies between estimated and real costs, traffic levels, and other elements in the appraisal of projects remain large, there is a need to analyse the causes in order to improve ex-ante methodologies. The imbalance between what we think will happen and what we know has happened can only be corrected through systematic before-and-after studies.

The best way to increase our knowledge of actual scheme performance is through ex-post evaluations that assess the success of projects through a framework that includes traditional economic efficiency, but also includes the wider social and environmental impacts, land use

changes, and impacts on regional development. Broadening evaluations can strengthen scope, accuracy, and credibility (OECD, 2018).

We have presented a framework for evaluating large government projects in Norway. The framework includes assessments of economic efficiency, effectiveness, cost performance, and alignment with strategic objectives. Practice can be improved through systematic evaluation in which the results are mapped against predictions and goals are formulated in the original business cases, and then fed into the appraisal of future projects. The evaluation is arms' lengths from policy, thus ensuring that different measures of project performance are considered.

The evaluation framework presented in this paper is to a greater extent aimed at addressing the aspirations and concerns of decision-makers than can be summarised in the CBA. It recognises that success is multi-dimensional and that partial measures such as cost performance or economic efficiency may be too narrow for projects with multiple stakeholders and long time-horizons.

Ex-post evaluation can be tailored to assess how projects perform and how this could help the ex-ante appraisal process. We have not presented detailed results of the individual evaluations in this paper, nor have we discussed the practical challenges of carrying out the evaluations. Documentation of ex-ante assumptions, developing sound counterfactual scenarios, and uncertainty with regards to future development are issues that should be considered in the further development of the evaluation framework.

The results of the evaluations show that the minority of the projects perform well on all dimensions, but even projects that represented inefficient use of resources may have acceptable results measured on other dimensions.

Figure 2 shows the balance between value for money and strategic success (the black dots represent road projects and the grey dots rail projects). Ideally, all projects should sit in the upper right corner of the graph, and many do. However, Figure 2 does illustrate a paradox. Road investment may be less expensive and provide better value for money than rail investment, according to conventional CBA methodology. For all that, there is political consensus that rail transport must play an important role in a future low-emission society.

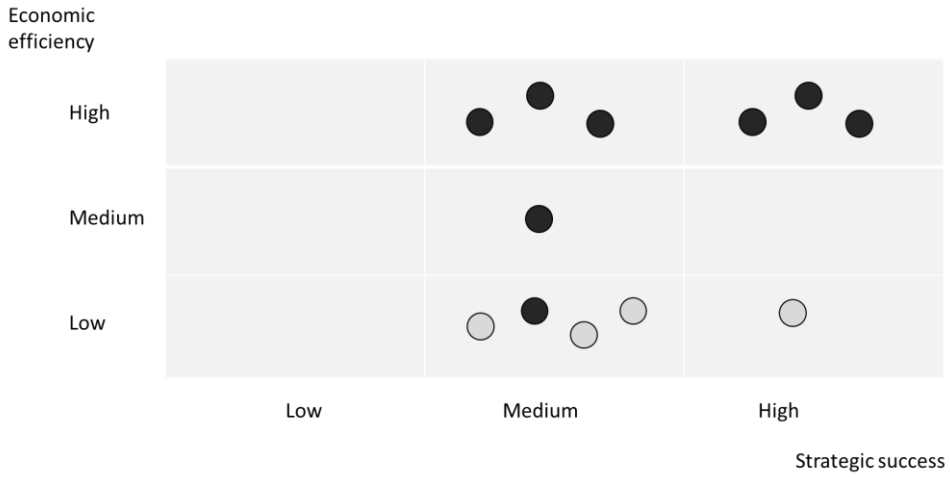

Figure 2: Combinations of value for money and strategic relevance ex-post

Norway is among a few countries that routinely carry out ex-post evaluations of implemented projects. The evaluation framework presented in this paper has originated in academia, but is now supported by the Ministry of Finance and is being used for projects in different sectors. The results can help to improve practice in future projects. As suggested by Samset and Christensen (2015), the early appraisal of an investment case should apply essentially the same evaluation criteria as used in ex-post evaluation. In a governance framework with formalised requirements for front-end documentation, there is potential for making ex-post evaluation an integrated part of the assessment process for transport investment and policy. By systematising the results of case-by-case evaluations into a formalised governance framework, this could provide a better way into improved ex-ante appraisal practices.

The evaluations presented in this paper were all carried out some five years after the respective projects opened. However, transport infrastructure is a long-term investment and it may take a long time and a combination of several measures before the total effects feed fully into the wider economy. Thus, a topic for future research should be to follow the projects further into their operational phase.

References

Annema, J.A., Frenken, K., Koopmans, C. and Kroesen, M., 2017. Relating cost-benefit analysis results with transport project decisions in the Netherlands. *Letters in Spatial and Resource Sciences*, 10 (1), 109-127.

Atkins and AECOM, 2014. *Meta Evaluation of Local Major Schemes. Final Report* [online]. Available from: https://www.gov.uk/

Boardman, A.E., Greenberg, D.H., Vining, A.R. and Weimer, D., 2006. *Cost-Benefit Analysis. Concepts and Practice*. Upper Saddle River, New Jersey: Prentice Hall.

Börjesson, M., Eliasson, J. and Lundberg, M., 2014. Is CBA Ranking of Transport Investments Robust? *Journal of Transport Economics and Policy*, 48 (2), 189-204.

Chih, Y.Y. and Zwikael, O., 2015. Project benefit management: a conceptual framework of target benefit formulation. *International Journal of Project Manag*ement, 33 (2), 352–362.

Chianca, T., 2008. The OECD/DAC Criteria for International Development Evaluations: An Assessment and Ideas for Improvement. *Journal of MultiDisciplinary Evaluation*, 5 (9), 41-51.

Concept, 2017. *Etterevaluering av statlige investeringsprosjekter. Retningslinjer for evaluator* [In English: Ex-post evaluation of government investment projects. Guildelines for evaluator]. Available from: https://www.ntnu.no/web/concept/etterevaluering

Crozet, Y., 2013. *High Speed Rail Performance in France: From Appraisal Methodologies to Ex-post Evaluation*. OECD International Transport Forum, Discussion Paper 2013: 26. . International Transport Forum, OECD, Paris.

Department for Transport, 2015. *Ex-post evaluation at the UK Department for Transport*. Presentation to CGEDD, 24th June 2015.

de Ruz, G., 2012. *Economic Evaluation of the High Speed Rail*. Ministry of Finance, Stockholm.

Eliasson, J., 2014. Introduction to the special issue on appraisal. *Research in Transport Economics*, 47, 1-2.

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

European Investment Bank, 2005. *Evaluation of EIB Financing of Railway Projects in the European Union*. Synthesis Report. Available from: http://www.eib.org/attachments/ev/ev_railways_en.pdf

Flyvbjerg, B., Skamris Holm, M. and Buhl, S., 2002. Underestimation of Costs in Public Works Projects: Error or Lie? *Journal of the American Planning Association*, 68 (3), 279-295.

Fridstrom L. and Elvik R., 1997. The barely revealed preference behind road investment priorities. *Public Choice*, 92 (1/2), 145–168.

Førsund, F.R., 2017. Measuring effectiveness of production in the public sector. *Omega*, 73, 93-103.

Givoni, M. and Dobruszkes, F., 2013. A Review of Ex-Post Evidence for Mode Substitution and Induced Demand Following the Introduction of High-Speed Rail. *Transport Reviews*, 33 (6), 720-742.

Highways England, 2016. Post Opening Project Evaluation (POPE) of Major Schemes. Executive Summary.

Haavaldsen, T., Lædre, O., Volden, G.H. and Lohne, J., 2014, On the concept of sustainability – assessing the sustainability of large public infrastructure projects, *International Journal of Sustainable Engineering*, 7 (1), 2-12.

International Transport Forum, 2017. Ex-Post Assessment of Transport Investments and Policy Interventions. ITF Roundtable Reports. Paris: OECD Publishing.

Jansson, J.-O. and Nilsson, J.-E., 1989. Spelar samhällsøkonomiska kalkyler någon verklig roll i vägvesendet? *Ekonomisk Debatt*, 2, 85–95.

Jones, H., Moura, F., and Domingos, T., 2014. Transport infrastructure project evaluation using cost-benefit analysis. *Procedia – Social and Behavioural Sciences*, 111, 400-409.

Kjerkreit, A. and Odeck, J., 2015. The accuracy of ex-ante cost benefit analysis – evidence from the Norwegian road sector. Paper presented at: *World Road Congress*, Seoul, South Korea, November 3 to November 5, 2015.

Laird, J., Kelly, C., Nellthorp, J., Constantini, S., Carbajo, J., Mas, X., Esteban Leon, J., Richards, P., Macdonald, N., oakes, G., Tindall, J., Arrand, J. and Black, K., 2012. Ex-post Appraisal: What lessons can be learnt from EU Cohesion funded transport projects. Paper presented at: *the European Transport Conference*, Glasgow, October 2012.

Mackie, P., Worsley, T. and Eliasson, J., 2014. Transport appraisal revisited. *Research in Transport Economics*, 47, 3-18.

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

Meunier, D. and Welde, M., 2017. Ex-post evaluations in Norway and France. *Transportation Research Procedia*, 26, 144-155.

Mouter, N., 2017. Dutch politicians' use of cost–benefit analysis. *Transportation*, 44 (5), 1127-1145.

Nellthorp, J., and Mackie, P., 2000. The UK Roads Review—a hedonic model of decision making. *Transport Policy*, 7(2), 127–138.

Nicolaisen, M.S, 2012. *Forecasts: Fact or Fiction? Uncertainty and Inaccuracy in Transport Project Evaluation*. PhD thesis. Department of Develpment and Planning, Aalborg University

Nicolaisen, M.S. and Driscoll, P.A., 2014. Ex-Post Evaluations of Demand Forecast Accuracy: A Literature Review. *Transport Reviews*, 34 (4), 540-557.

Nicolaisen, M.S. and Driscoll, P.A., 2016. An International Review of *Ex-Post* Project Evaluation Schemes in the Transport Sector. *Journal of Environmental Assessment Policy and Management*, 18 (1), 33 p.

Nilsson, J.-E., 1991. Investment Decisions in a Public Bureaucracy: A Case Study of Swedish Road Planning Practices. *Journal of Transport Economics and Policy*, 25(2), 163–175.

Nyborg, K., 1998. Some Norwegian Politicians' Use of Cost-benefit analysis. *Public Choice*, 95, 381-401.

OECD, 2018. *OECD Economic Surveys: Norway 2018*. OECD Publishing, Paris. Available from: http://dx.doi.org/10.1787/eco_surveys-nor-2018-en

OECD Development Assistance Committee, 1991. *Principles for Evaluation of Development Assistance*. OECD, Paris.

Odeck, J., 1996. Ranking of regional road investment in Norway. *Transportation*, 23(2), 123–140.

Odeck, J., 2010. What Determines Decision-Makers' Preferences for Road Investments? Evidence from the Norwegian Road Sector. *Transport Reviews*, 30(4), 473–494.

Odeck, J. 2014. Do reforms reduce the magnitudes of cost overruns in road projects? Statistical evidence from Norway. *Transportation Research Part A: Policy and Practice*, 65, 68-79.

Odeck, J., 2017. Variation in cost overruns of transportation projects: an econometric meta-regression analysis of studies reported in the literature. *Transportation*, DOI https://doi.org/10.1007/s11116-017-9836-5

Odeck, J., Volden, G.H. and Welde, M., 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.

Sager, T. and Ravlum, I.A., 2005. The political relevance of planners' analysis: the case of a parliamentary standing committee. *Planning Theory*, 4 (1), 33-65.

Samset, K., 2003. *Project Evaluation. Making investments succeed*. Trondheim: Tapir Academic Press.

Samset, K. and Christensen, T., 2015. Ex-ante project evaluation and the complexity of early decision-making. *Public Organizational Review*, 17 (1), 1-17.

Trafikverket, 2017. This is Vision Zero. Available from: https://www.trafikverket.se/en/startpage/operations/Operations-road/vision-zero-academy/This-is-Vision-Zero/

Volden, G.H., 2017. Ex-post evaluation of large public infrastructure projects – experiences from Norway. *Nordic Academy of Management Conference*. Bodø: August 23-26 2017.

Volden, G.H. and Samset, K., 2017a. *Statlige investeringstiltak under lupen. Erfaring med evalueringer av de 20 første KS-prosjektene* [in English: A Close-up on Public Investment Cases. Lessons from Ex-post Evaluations of 20 Major Norwegian Projects]. Concept Research Report no. 52. Trondheim: Ex-ante Academic Publisher.

Volden, G.H. and Samset, K., 2017b. Quality Assurance in Megaproject Management. In: Flyvbjerg, B. (ed.). *The Oxford Handbook of Megaproject Management*. Oxford: Oxford University Press, 406-427.

Welde, M., 2017. *Kostnadskontroll i store statlige investeringer underlagt ordningen med ekstern kvalitetssikring* [Cost performance in large government investment projects that have been subjected to external quality assurance]. Concept report no. 51. Trondheim: Ex-ante Academic Publisher.

Welde, M. and Odeck, J., 2017. Cost escalations in the front-end of projects – empirical evidence from Norwegian road projects. *Transport Reviews*, 37 (5), 612-630.

Yu, M.-M., 2008. Assessing the technical efficiency, service effectiveness, and technical effectiveness of the world's railways through NDEA analysis. *Transportation Research Part A*, 42, 1283-1294.