Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 42 (2020) 75–84

Transportation Research Procedia

www.elsevier.com/locate/procedia

46th European Transport Conference 2018, ETC 2018, 10-12 October 2018, Dublin, Ireland

Ex-post evaluation of major infrastructure projects

Gerard de Jong[a,*], Silvia Vignetti[b], Chiara Pancotti[b]

[a]*Significance, The Hague, The Netherlands, and ITS, University of Leeds, UK*
[b]*CSIL Centre for Industrial Studies, Milan, Italy*

Abstract

Ex ante assessment of a project by means of cost-benefit analysis is quite common in transport. On the other hand, ex post evaluations are only done very infrequently, even though such analyses can be very useful. Ex post evaluation can be used to check whether projects really delivered the benefits expected from them at the time, and to learn which projects do better and which do worse than expected, and why. This paper reviews the literature for evaluating projects that have been completed (ex post evaluation). It also presents the methodology that was selected for the ex post evaluation of ten major transport infrastructure projects supported by the European Regional Development Fund (ERDF) and Cohesion Fund (CF) between 2000 and 2013. A conceptual framework on the impacts of a transport project on society has been developed and worked out in terms of an ex post assessment methodology. The conceptual framework includes a typology of effects for investment projects in the transport sector and the timeframe of effects. Ten case studies (major transport projects) have been carried out using this framework, and the outcomes for these case studies are reported.

Keywords: ex post evaluation, infrastructure projects, transport projects, cost-benefit analysis

1. Introduction

In transport, project assessment, e.g. using cost-benefit analysis, is often done ex ante, but systematic ex post evaluations are only done infrequently (e.g. Norway and the UK have systems in place for post-opening evaluation). When present, they cover a time horizon shorter than five years, longer term ex-post evaluations being carried out sporadically. Nevertheless, such analyses can be very useful to check whether projects really delivered the benefits expected from them at the time, and to learn which projects do better and which do worse than expected, and why.

This paper reviews the literature on ex post evaluation which focuses on differences between ex ante and ex post costs and benefits and presents the methodology that was selected for the ex post evaluation of ten major transport infrastructure projects supported by the European Regional Development Fund (ERDF) and Cohesion Fund (CF) between 2000 and 2013. This ex post evaluation was carried out for DG REGIO by a consortium consisting of CSIL, Ramboll, Significance and T-Plan and, rather than merely comparing ex ante and ex post, aimed at gathering an in-depth understanding of the long-term long term contribution of the selected projects to economic development and quality of life.

* Corresponding author:
E-mail address: dejong@significance.nl

10.1016/j.trpro.2019.12.008

A conceptual framework on the impacts of a transport project on society has been developed and worked out in terms of a uniform ex post assessment methodology. A substantial part of the project effects are included in an ex post cost-benefit analysis. Other effects are treated in a qualitative analysis with the help of a uniform scoring system. This paper will summarise the methodology.

Using this common methodology, the study then evaluates projects that are at least five years in operation. The fact that the CBA is carried out during the lifetime of the selected projects leads to a hybrid typology of CBA, sharing the features of both an ex ante CBA and a pure ex post (i.e. retrospective) CBA.

Beyond the assessment of long-term effects, the evaluation tries to elaborate on the mechanisms explaining the observed performance and its key determinants. We develop a classification of stylised patterns for the outcomes of a project (ex ante as well as ex post) and then use this to classify the ten transport project studied.

In section 2 of this paper, we will give a short overview on the literature on ex post evaluation of projects in transport. Section 3 will contain a description of the main features of the common evaluation methodology. In section 4, the stylised patterns for the project outcomes will be discussed and presented for the ten projects that were evaluated. Finally some concluding remarks will be provided in section 5.

2. A brief review of the literature on ex post evaluation

In this section we review papers about differences between ex ante and ex post costs and benefits and especially about the explanations of these differences, which we also interpret as the causes of project success and failure.

Flyvbjerg (with co-authors) is one of the key references on the issue of ex-ante/ex-post comparison and on the challenges affecting the selection and implementation of major infrastructure projects. He has published several papers, books and reports about the comparison of ex ante forecasts of project costs and benefits against the actual outcomes (Flyvbjerg et al. 2003, 2004, 2005; Flyvbjerg, 2007). A database on the construction costs of large-infrastructure projects (projects exceeding 100 mln dollar) was constructed, consisting of 258 projects in 20 countries over a 70-year period. Three types of projects were distinguished: rail, bridges/tunnels and road projects. For all three types, the variation in forecasting errors was large. Nine out of ten projects has a cost overrun. This happens across the 20 countries and has not improved over time. The average cost overrun is 45% for rail, 34% for bridges and tunnels and 20% for roads.

For the comparison of demand forecasts versus observed demand, the database has 208 large-infrastructure projects in 14 countries over a 30-year period. Here the distinction is between rail and road projects. The variation in inaccuracy is large for both types of project. For rail, on average the actual demand is 51% lower than estimated (nine out of ten rail projects overestimate demand), for road actual traffic on average is 10% higher than predicted (about the same number of road projects has overestimated traffic or underestimated traffic).

The situation of cost overruns and benefit shortfalls, bound to lead to failed projects, is therefore not unlikely at all, especially for rail projects.

In Flyvbjerg (2007) the causes for these inaccuracies in forecasts are discussed. Three main types of explanations have been offered:

- Technical explanations, i.e. errors and pitfalls in forecasting techniques (e.g. imperfects models, inadequate data).
- Psychological explanations, such as planning fallacy and optimism bias. According to this explanation (following for instance Kahneman and Tversky, 1979), planners will suffer from cognitive biases focussing too much on the possibly high benefits and largely ignoring costs and risks. A potential cure here would be reference-class forecasting: use the findings from a class of similar projects instead of the details of the project itself.
- Political-economic explanations. Planners and promoters deliberately and strategically overestimate benefits and underestimate costs to get their projects selected.

Flyvbjerg (2007) argues that the technical explanations do not fit the data well. They would lead to both positive and negative errors, but the data mainly show one-sided errors. Also if this were the reason, the inaccuracies should decrease over time as methods and data get better. But this has not happened in the dataset.

Psychological explanations can explain the bias in the errors in the data. But Flyvbjerg (2007) argues that it seems unlikely that "…a whole profession of forecasting experts would continue to make the same mistakes decade after decade", so does not see this as a credible explanation.

The political-economic explanation also explains the observed biases and is supported by outcomes of interviews with planners and consultants in the UK and the US. Flyvbjerg (2007) regards this as the key cause for cost overruns and benefit shortfalls and his suggested cures (especially: improved incentives in the public and private sector to be transparent and honest) are based on this diagnosis.

Cantarelli et al. (2010) focus on the concept on 'lock-in' as a further explanation of cost overruns. Lock-in is created when suboptimal policies are used as a consequence of path dependency, even though a better alternative is present.

Lock-in can occur both at the decision-making level (before the decision to build) and at the project level (after the decision to build) and can influence the extent of overruns in two ways. Due to lock-in, the `real decision to build' is made quite early in the decision-making process and the costs estimated at that stage are often much lower than those that are estimated at a later stage in the decision-making process, thus increasing cost overruns. The second way that lock-in can affect cost overruns is through `practice'. Although decisions about the project (design and implementation) need to be made, lock-in can lead to inefficient

decisions that involve higher costs.

Sunk costs (in terms of both time and money), the need for justification, escalating commitment, and inflexibility and the closure of alternatives are indicators of lock-in.

This suggests that recognition of lock-in as an explanation for cost overruns contributes significantly to the understanding of the inadequate planning process of projects and allows development of more appropriate means.

Eliasson and Fosgerau (2013) presented some numerical evidence that a project selection process on the basis of the benefit-cost ratio on its own is enough to generate bias of typical magnitudes (the cost overruns and benefit shortfalls as found by Flyvbjerg), given plausible parameters. In essence, their argument is a variant of the classic rule of 'regression to the mean' from statistics (just as very short people have a substantial chance of having taller children). This argument does not provide information on causes of project outcomes but treats these as random events. While they can refute the argument of Flyvbjerg, they acknowledge they cannot refute his conclusion. Strategic misrepresentation by project promoters may well exist; but the existence of systematic cost overruns and benefit shortfalls does not prove this. As long as projects compete for selection based on uncertain, formal or informal, predictions of costs and benefits, these phenomena are bound to occur.

Nicolaisen and Driscoll (2016) found that the primary considerations in ex ante transport project appraisal are often construction cost, travel time savings/congestion relief and safety (reduced accidents). Therefore, a number of ex post evaluations has focussed on comparing predicted and actual results on these aspects only. They review the literature on ex post evaluation of transport projects. Little is known about the causal mechanisms explaining the differences between expected and actual cost and demand. Often there is no systematic sampling of projects to assess ex post. The timescale in ex post evaluation is usually 1-5 years after opening, whereas the ex-ante evaluation looked at a much longer timescale (20-25 years ahead). Poor documentation at the different ex ante evaluation stages and lack of monitoring after the opening are recurrent problems. Conceptual problems are defining the reference point for classifying projects as ex ante or ex post (often one uses the decision to fund for this) and defining the counterfactual.

Two of the better ex post evaluation schemes are discussed in detail in Nicolaisen and Driscoll (2016): England's Post-Opening Project Evaluation (POPE) and Norway's post opening evaluation of major projects. POPE looks at: safety, cost, environmental impacts, accessibility and integration with other local, regional and national plans and programs. A severe tendency is found to underestimate investments costs (consistent with a larger body of literature, including the work of Flyvbjerg). Traffic estimates on the other hand are in many cases reasonably accurate, with a balance between overpredictions and underpredictions. Journey time savings are often overestimated. The Norwegian ex post evaluations only look at the monetised costs and benefits (e.g. costs, demand, accidents, local air pollution). For costs, overestimations and underestimations were about equally likely. Most road projects have higher benefits than expected.

3. Our methodology for ex post evaluation

3.1 Mix of CBA and qualitative analysis

Cost-benefit analysis has been selected as a key part of the ex post evaluation methodology for the ten transport projects for the following reasons:

- CBA is the most suitable quantitative method to investigate the details required to isolate the impact of an individual project.
- CBA is a reliable tool to express project benefits and externalities in monetary terms.
- Being founded in welfare economics, CBA measures all impacts in terms of welfare changes. This does not only make it possible to rank project, but also to reach conclusions about social desirability of the project by itself.

Given these considerations, CBA complemented by a qualitative analysis is the methodological option adopted for the present evaluation.

The adopted qualitative techniques are documentary analysis, desk research, and an extensive series of interviews with stakeholders (around 25 for each project) with field missions. The objectives of the qualitative analysis are:

- Describing the project with a critical focus on its identification.
- Analysing the socio-economic context.
- Reconstructing the decision making process.
- Assessing possible alternative options
- Reflecting the views of all the stakeholders involved in the project design and implementation.
- Collecting evidence on non-quantifiable effects and factors influencing project's performance.
- Effects investigated in qualitative terms are then aggregated to measurable effects and a comprehensive assessment is provided through a scoring system from -5 (the highest negative effect) to +5 (the highest positive effect).

The CBA and qualitative scores are based on an analysis of documents for the original ex ante evaluation of the project, data on observed developments since the opening and field missions and interviews (245 interviews in total) with (local) stakeholders.

This approach allows taking into account the variety of effects which determines the long-term contribution of each project. On one hand, the CBA approach allows to measure in the same unit (money) and then to aggregate the different effects produced by the projects (and balancing, for example, negative with positive effects). On the other hand, qualitative analysis is a helpful

complementary tool when a proper monetisation of effects is not suitable. More specifically, in terms of their measurement level, the effects can be distinguished into:

A. Effects that by their nature are already in monetary units (e.g. transport costs savings). These can therefore be easily included in a cost-benefit analysis (CBA).

B. Effects that are quantitative, but not in money units, and that can be converted into money units in a reasonably reliable way (e.g. transport time savings, accidents, air pollution). These effects can also be included in the CBA.

C. Effects that are quantitative, but not in money units, for which there are no reasonably reliable conversion factors to money. We propose not to try to include such effects in the CBA, but to discuss them in a qualitative way together with the overall outcome of the CBA.

D. Effects that are difficult to measure in quantitative (cardinal) terms, but do lend themselves for ordinal measurement (a ranking of the impact of different projects on such a criterion can be provided, such as very good, good, neutral, bad, very bad). We propose to discuss these effects in qualitative terms.

E. Effects that might occur but that are subject to a high degree of uncertainty: these will be treated as part of the risks/scenario analysis that will be included in the CBA.

F. Effects that might occur but that we cannot even express in an ordinal (ranking) manner: they are residual effects that can be mentioned in the qualitative description of the case study report.

All effects in A and B have been included in the CBA and they are the most significant share of long-terms effects. Then, the outcome of CBA is complemented by evidence from C and D, while E and F are used for descriptive purposes.

Table 1 list which variables are included in the CBA (in Annex 1 we explain how the effect is defined and monetised) and which in a qualitative fashion. For some effects there is an entry for both CBA and qualitative analysis. This means that we first try to include the effect in the CBA, but if this is not possible, the effect is treated in a qualitative way.

3.2 Structure and features of the ex post CBA

While the overall methodological reference for CBA is the DG REGIO Guide, the Core Team had to slightly adjust it in order to take into account the mid-term perspective of the assessment.

As said, the selected projects are operating for at least 5 years. This led to two main implications:

1. The ex-post CBA can be more ambitious in terms of effects to be accounted for as the risk of optimism bias is mitigated by the possibility to rely on observed data.
2. As the CBA is carried out during the life time of the project, it has been necessary to adopt a hybrid methodology which shares ex-ante and ex-post perspectives (i.e. backwards and onwards values).

The project did not carry out new runs with transport models, but used existing runs for the ten different projects that were evaluated.

4. Stylized pattern for the outcomes of a project

4.1 Understanding the effects

After having developed an inventory of the outcomes of the different projects, we analyse the elements, both external and internal to the projects, which have determined the observed causal chain of effects and influenced the observed project performance.

On the basis of the literature we have developed a method to classify typical outcomes of transport projects (a stylised pattern of how good or bad a project performs ex ante and ex post). The key variables here are cost (i.e. construction costs) and monetised benefits. For both the ex-ante situation and the ex post situation we score each project in a two-dimensional matrix, using the scores A to D, where A (low costs and benefits) and D (high costs and benefits) stand for intermediate project, B (low costs and high benefits) for good project, C (high costs and low benefits) bad project.

We can now distinguish several typologies of project performance over time, where BD (B before and D after) and AC for instance represent a cost overrun and BA and DC a benefit shortfall. Using this typology, and the causal factors that are most likely to explain the observed patterns, we classify projects with regards to a number of stylised patterns.

Taking inspiration from the literature on the success and failure of projects (see section 2), and particularly on costs overruns and demand shortfalls, and on the basis of the empirical evidence which develops from European Commission (2012) six stylised determinants of projects' outcomes and their development over time have been identified:

- Relation with the context which includes considerations of institutional, social and economic environment into which the project is inserted.
- Selection process which refers to the institutional and legislative framework that regulates how public investment decisions are taken.
- Project design which refers both to designing and the development of the project.
- Forecasting capacity representing the possibility and capacity to predict future trends and forecast the demand level and technical challenges.

- Project governance concerns the number and type of stakeholders involved during the project cycle and how responsibilities are attributed and shared.
- Managerial capacity refers to both the professional ability to react to changes in the context and to unforeseen events and the professional capability to ensure the expected level of services in the operational phase.

Table 1. Effects in the CBA and in the qualitative analysis.

Category	Effect	CBA	Qualitative analysis
Economic growth	Travel time	✓	
	Vehicle operating cost	✓	
	Reliability of journey time	✓	✓
	Income for the service provider	✓	
	Wider economic impacts		✓
	Institutional learning		✓
Quality of life	Safety	✓	
	Crowding	✓	
	Service quality (other than crowding)	✓	✓
	Security		✓
	Noise	✓	
	Aesthetic value	✓	✓
	Urban renewal	✓	✓
Environmental sustainability	Local air pollution	✓	
	Climate change (GHG)	✓	
	Biodiversity		✓
Distributional effects	Allocation over social groups		✓
	Allocation over territorial areas		✓

Table 2. Possible key outcomes in the ex-ante evaluation of transport projects for 'before' and 'after'

		MONETISED BENEFITS	
		Low	High
COST	low	A: intermediate	B: good project
	high	C: bad project	D: intermediate

It is worth noticing that these six stylised determinants are highly interrelated and they may mutually reinforce or dilute: a very unstable context is likely to obstruct the forecasting capacity; bad incentive mechanisms can negatively affect the project design. Moreover, determinants may change over time. Therefore, it is important to make clear the link between identified determinants and the specific effect triggered. In doing so, the research team identified stylised typical "paths" or project behaviours linking the interrelation of different determinants in a dynamic fashion. These patterns represent common stories describing recurring pattern of performances, as well as typical problems that may arise and influence the chronicle of events. The list of stylised patterns is shown in the table below.

These stylised patterns had been designed to be as comprehensive as possible. However, variation on these patterns emerged throughout the case studies analysis as a sort of crossover or adjustment of the six original patterns. This has been done in order to be as accurate as possible when describing project "paths". These new patterns are:

- Blurred star which represents good project(s) falling short to be "bright stars" as one determinant is sub-optimal.
- Little star which represents fairly good projects with some deficiencies in the planning phase.
- Shooting supernova where a propitious context and good ex-ante preconditions have avoided the complete failure of the project from the negative effect of poor forecasting capacity and project management
- Star representing a satisfactory project which is somehow exposed to risks if the underlying assumptions will not materialise in the future.
- Eclipse Sun where – unlike the rising sun case – a poor forecasting capacity led to project under-achievement.

Table 3. Behavioural patterns archetypes

TYPE	DESCRIPTION
Bright star	This pattern is typical of projects where the good predictions made ex-ante (both on the cost side and demand side) turn out to be accurate. Proper incentive systems are in place so that the project actually delivers value for money and success. Even in the event of exogenous negative events, the managerial capacity ensures that proper corrective actions are taken and a positive situation is restored.
Rising sun	This pattern is typical of projects which, soon after their implementation, are affected by under capacity issues because of a combination of low demand forecasting capacity, weak appropriateness to the context, and weak technical capacity to design the infrastructure. However, due to changed circumstances or thanks to responsible management and good governance the project turns around to reap new benefits.
Supernova	This pattern is typical of projects for which the good predictions made ex-ante (both on the cost and demand side) turn out to be accurate. However, due to changed circumstances or because of weak management capacity and/or governance the project eventually turns out to be unsuccessful.
Shooting star	This pattern is typical of projects starting from an intermediate situation and resulting in a failure. This outcome can be explained by a low forecasting capacity affected by optimism bias which yields a cost overrun. Then during project implementation, because of low managerial capacity and/or poor governance (also due to distorted incentives) corrective actions are not implemented, this leading to project failure. The situation is exacerbated if unexpected negative events materialise during the project implementation.
Black-hole	This pattern is typical of projects that since the beginning of their life fail to deliver net benefits. This is a result of a combination of ex-ante bad factors (i.e. low technical capacity for demand forecasting, optimism bias, inappropriateness to the local context and bad incentives affecting both the selection process and the project governance) and careless management during the project implementation or bad project governance (e.g. unclear division of responsibilities, bad incentive schemes).

4.2 Final assessment of the ten projects

Qualitative and quantitative findings are integrated in a narrative way, in order to develop ten project 'histories' and to isolate and depict the main aspects behind the project's long-term performance. A final assessment on each project is then conveyed in the case studies with an assessment structured along a set of project performance criteria:

- Relevance (were the project objectives in line with the existing development needs and the priorities at the programme, national and/or EU level?);
- Coherence (with other national and/or EU interventions in the same sector or region);
- Effectiveness (were the stated objectives achieved, and in time? Did other effects materialise? Were other possible options considered?);
- Efficiency (costs and benefits relative to each other and to their ex-ante values);
- EU added value (was EU support necessary, EU-wide effects, further EU action required?).

5. Case studies and main findings

Table 4 includes the list of which ten projects were included in our ex post evaluation. This is the result of a selection strategy consisting of three phases and ensuring data availability, project varieties in terms of mechanisms and causal chains, willingness to cooperate by potential stakeholders, as well as capacity to produce relevant policy lessons.

The main outcomes for these ten projects, including the scores on the above five performance criteria and on their determinants are given in Table 5 together with the classification of their behavioural patterns.

The objective of putting together the project outcomes along the various line of analysis was clustering (as far as possible) different projects in order to draw some general policy lessons. Arguably, the most straightforward criterion to classify the ten case studies is on the basis of their success or failure (see the table and the graph below).

The ten transport projects that we evaluated did not suffer from the large cost overruns and demand shortfalls that Flyvbjerg (2007) reported earlier for a large set of completed projects. This difference may of course partly be due to the fact that we are evaluating projects 5-10 years after opening, whereas their ex ante appraisals had been carried out for considerably longer time horizons. Except for the Le Havre project, the projects all have a benefit-cost ratio above 1 ex post (monetised benefits exceeding costs), though the benefits for about half of the projects were clearly lower than had been anticipated ex ante.

The most effective projects were those responding to a clear transport need and providing significant transport benefits

(especially in terms of travel time and cost reductions). Underperforming projects were those in which the key objectives of the project included wider effects as regards urban renewal or socio-economic development in the catchment areas.

Table 4. List of selected projects and final classification

PROJECT	COUNTRY	SUB-SECTOR	EC FUND	PERIOD
Autobahn A14	DE	Road	ERDF	2007-2013
Rio Antirio Bridge	GR	Road	ERDF	2000-2006
M43 Motorway	HU	Road	Cohesion Fund	2007-2013
Saulkrasti Bypass	LV	Road	ISPA (Cohesion Fund)	2000-2006
Malaga Bypass	ES	Road	ERDF	2007-2013
Warsaw Line 8 modernisation and airport connection	PL	Rail	Cohesion Fund	2007-2013
Modernistation of railtrack in Žilina	SK	Rail	Cohesion Fund	2007-2013
Tramway in Le Havre	FR	Urban transport	ERDF	2007-2013
Naples Metro Line 1	IT	Urban transport	ERDF	2000-2006
Gdansk Tram	PL	Urban transport	Cohesion Fund	2007-2013

A fully successful project such as the Rio Antirio Bridge is characterized by high quality at entry and solid resilience. Quality at entry refers here to the solidity and quality of the project preparation and selection processes, including the quality of ex ante feasibility and CBA analyses as well as the way the entire selection process is structured and works. Resilience instead is the capacity of the project to recover quickly and effectively from difficulties met during the implementation phase, taking the necessary measures to keep the project on a successful track. It is strictly linked to the managerial capacity and project governance.

The support from organisations like JASPERS, in combination with usually good project management, appears to have contributed to preventing failed projects. JASPERS is a technical assistance partnership between the EIB and the European Commission supporting member states in the preparation of major projects (http://www.eib.org/en/pro/advising/jaspers/index.htm).

Table 5. Project performance and its determinants for the ten selected projects (scores from 1 to 5)

			Final performance					**Determinants**					
Cluster	**Case study**	**Behavioural pattern**	**Relevance**	**Coherence**	**Effectiveness**	**Efficiency**	**EU added value**	**Relation with the context**	**Selection process**	**Project design**	**Forecasting capacity**	**Project governance**	**Managerial capacity**
Successful	Greece – Rio Antirio Bridge *(Bright Star)*	Project in which the good predictions made ex ante turn out to be accurate. **The project delivers value for money and success**. Even in the event of exogenous negative events, the project performance remained positive.	5	5	4	5	5	4	4	5	5	5	4
	Poland - Gdańsk Tram *(Star)*	**The project performance is very positive**. However, due to the fact that the infrastructure and services refer to a small intervention embedded into an existing wider network the positive performance of the project is highly influenced by network effects not fully attributable to the project.	5	5	4	4	4	4	5	3	4	5	4
	Poland – Warsaw Line 8 Modernisation and Airport Connection *(Star)*		5	5	4	5	5	5	3	3	4	3	3
	Latvia – Saulkrasti Bypass *(Rising Sun)*	**Project affected by exogenous unfavourable factors in the initial**	5	5	4	4	4	5	4	4	2	4	5

		phase. However, thanks to the commitment of stakeholders and managerial capacity, **it turns out to be successful**.											
Intermediate success	Spain – Malaga Bypass (*Blurred Star*)	**Project partially successful**. The sub-optimal coordination among level of governments partially clouded the fulfilment of all the expected objectives. However, the most urgent need was successfully addressed.	5	5	3	3	1	5	3	5	4	1	4
	Hungary – M43 motorway (*Little Star*)	**Project performance is positive** but far below the expectations. This is due to some deficiencies in the planning phase.	5	4	3	3	4	3	2	3	-2	4	4
	Slovakia – Žilina Railway Modernisation (*Little Star*)		5	5	3	3	5	-1	-4	3	1	3	4
	Italy - Naples Metro Line 1 (*Shooting Supernova*)	Project in which the propitious context and good ex ante preconditions have avoided the failure of the project from the negative effect of poor forecasting capacity and project management. **The project is underperforming.**	5	5	3	3	3	5	5	4	-3	-5	-5
Least successful	Germany – Autobahn A14 *(Rising Sun)*	**Project affected by a combination of ex ante unfavourable factors (**overoptimistic traffic forecast, inappropriateness to the local context**)**. However, the effective design and a good managerial capacity **prevented the project failure**.	2	5	2	3	1	-2	1	2	-3	4	4
	France - Le Havre tramway *(Eclipsed Sun)*	**Project in which a combination of ex ante unfavourable factors** (optimism bias, inappropriateness to the local context and bad incentives) **prevented the project to reach its expected benefits** and the good managerial capacity is unlikely to save the project from its underachievement.	2	3	2	2	3	-1	-2	5	-3	4	3

6. Concluding remarks

Ten major transport projects carried out in the period 2000-2013 and financed to some degree from the Cohesion Fund or European Regional Development Fund were evaluated ex post using a common methodology. The methodology includes a CBA (looking backward and forward), as well as a more qualitative scoring of project outcomes. The project performance was also related to its determining factors.

The ten transport projects evaluated in this way did not suffer from the large cost overruns and demand shortfalls that were observed in earlier studies by Flyvbjerg (2007). Nine out of the ten projects had a benefit-cost ratio above 1 ex post, though the benefits for about half of the projects were clearly lower than had been anticipated ex ante.

Even the worst performing projects from our list cannot be called failed projects, only underachieving projects. We see projects without optimism bias and projects where negative trends are overcome on the one hand and projects that suffered from great expectations that were only partly met and projects where unfavourable external developments were not properly countered on the other hand. But on the whole, the methods of project selection used in the ex ante stage, with a focus on formal CBA, guidelines on how to do CBA and which values to use and support from organisations like JASPERS in combination with usually good project management appears to work rather well at preventing project failure.

Ex-post evaluation is a useful tool as it incentivises decision makers to ensure good governance and to have a liability towards their decisions. Furthermore, it adds an extra layer of transparency on the outcome of the project. Hence, ex-post evaluation of projects should be done more systematically. In particular, evaluation is viewed as beneficial for managing authorities in order to build internal capacity and to improve the project selection process.

References

Cantarelli, C.C., B. Flyvbjerg, B. Van Wee and E.J.E. Molin (2010) Lock-in and its influence on the project performance of large-scale transportation infrastructure projects: investigating the way in which lock-in can emerge and affect cost overruns, Environment and Planning B, 37 (5): 792-807.

CSIL, Ramboll, Significance and T-Plan (2017) Ex post evaluation of major projects supported by the European Regional Development Fund (ERDF) and Cohesion Fund between 2000 and 2013, First Interim Report, Volume I and II, report for DGREGIO, CSIL, Milan.

CSIL, Ramboll, Significance and T-Plan (2018) Ex post evaluation of major projects supported by the European Regional Development Fund (ERDF) and Cohesion Fund between 2000 and 2013, Final Report, report for DGREGIO, CSIL, Milan.

Eliasson, J. and M. Fosgerau (2013) Cost overruns and benefit shortfalls – deception or selection?, Transportation Research B, 57, 105-113.

European Commission (2014) Guide to Cost Benefit Analysis of Investment projects, European Commission, Brussels.

Flyvbjerg, B. (2007) Policy and planning for large-infrastructure projects: problems, causes, cures, Environment and Planning B, 34 (4): 578-597.

Flyvbjerg, B., M.K. Holm and S.L. Buhl (2003) How Common and How Large Are Cost Overruns in Transport Infrastructure Projects? Transport Reviews, 23 (1): 71-88.

Flyvbjerg, B., M.K.S. Holm and S.L. Buhl (2004) What Causes Cost Overrun in Transport infrastructure Projects?, Transport Reviews, 24 (1): 3-18.

Flyvbjerg, B., M.K.S. Holm and S.L. Buhl (2005) How (In)accurate Are Demand Forecasts in Public Works Projects? The Case of Transportation, Journal of the American Planning Association, 71 (2): 131-146.

HEATCO (2006) Developing Harmonised European Approaches for Transport Costing and Project Assessment, Deliverable 5, Proposal for harmonized guidelines. IER, University of Stuttgart.

JASPERS (2017) JASPERS Appraisal Guidance (Transport), Guidance on appraising the economic impacts of rail freight measures, JASPERS, Bucharest.

Korzhenevych, A., N. Dehnen, J. Bröcker, M. Holtkamp, H. Meier, G. Gibson, A. Varma and V. Cox (2014) Update of the handbook on external costs of transport, European Commission, DG MOVE, Brussels.

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

Wardman, M. and G. Whelan (2011) Twenty years of rail crowding valuation studies: evidence and lessons from British experience, Transport Reviews, 31(3), 379-398

Wardman, M.R., P. Chintakayala and G.C. de Jong (2016) European wide meta-analysis of values of travel time, Transportation Research A, 94, 93-111.

Annex 1: Treatment of the effects in the CBA

IMPACT	Q-SIDE (QUANTIFY EFFECT)	P-SIDE (CONVERSION TO EUROS)
Travel time savings	Observed time reductions and predicted time reductions from regional transport models both for passengers and for freight.	National values, or unitary values of time (VOT) for passengers and freight: - Passengers VOTs are from the meta-analysis by Wardman et al. (2016, Table 9). They are the most recent available data estimated at country level. Data were converted in 2017 PPP euro. - Freight VOT are from HEATCO (2006; Table 4.11), converted in 2017 PPP euro.
Reliability	Changes in the observed and/or predicted time distribution.	- For passenger reliability ratio is from Wardman et al., 2016. - For freight reliability ratio are from JASPERS (2017). Reliability ratios should be meant as factors to adjust the value of time.
Vehicle operating cost savings (VOCs).	Observed costs reductions and predicted reductions from regional transport models (distance reductions only).	VOCs are country or project specific and they are already expressed in monetary values from regional transport models or feasibility studies. If such values are not available, a simplified formula to calculate VOCs is provided based on the approach suggested in the EC Guide.
Income for the service provider	Level of income from revenues and/or avoided operating cost savings	Financial value corrected to avoid double counting with possible other already considered economic benefits (such as VoC or VoT) that may be reflected in the income value.
Safety	Observed accidents and predicted from regional transport models accidents variations.	Social accidents costs by country and type of injury from Table 10 in Korzhenevych et al. (2014), converted in 2017 PPP euro. The latter is a report prepared for DG MOVE, which provides updated estimates of HEATCO (2006) data.
Crowding	Compare development of capacity and demand in public transport.	Wardman and Whelan (2011)
Service quality (other than crowding)	Observed number of passengers declaring a higher service quality.	Project-specific willingness to pay values when available.
Climate change (Greenhouse gases)	Observed vehicle km reduction and predicted reduction from regional transport models.	Unitary monetary values of CO2 emission (in 2017 €/ton CO_2 equivalent) from EIB (2013) as suggested in the EC Guide and by JASPERS.
Local air pollution	Observed vehicle km reduction and predicted reduction from regional transport models.	Unitary monetary values of main local pollutants expressed in CO2 equivalent. Original data (€/ton CO_2 equivalent) from Korzhenevych et al. (2014; Table 15) have been expressed in 2017 prices for the purpose of this study. Korzhenevych et al. (2014), which provides updated estimates of NEEDS data.
Noise	Observed vehicle km reduction and predicted reduction from regional transport models.	Values from Korzhenevych et al. (2014; Table 28), converted in € 2017 per 1000 vehicle km.
Aesthetic value	Observed number of residents reporting a positive/negative experience for aesthetic reasons	Project-specific willingness to pay values when available.
Urban renewal	Number of observed and estimated m2 of surface of real estates affected by a variation in the economic value due to the proximity to the transport infrastructure	Project-specific observed or expected increase in market prices as reported by real estate sector experts or available from official local sources
Allocation over social groups	Number of different social groups	ENPV accruing to different social groups