International Transport Forum

Ex-Post Assessment of Transport Investments and Policy Interventions

Roundtable Report

OECD

Ex-Post Assessment of Transport Investments and Policy Interventions

Roundtable Report

Please cite this publication as:
ITF (2017), *Ex-Post Assessment of Transport Investments and Policy Interventions*, ITF Roundtable Reports, OECD Publishing, Paris.
http://dx.doi.org/10.1787/9789282108154-en

ISBN 978-92-82-10814-7 (print)
ISBN 978-92-82-10815-4 (PDF)

Series: ITF Roundtable Reports
ISSN 2519-8777 (print)
ISSN 2519-8785 (online)

Photo credits: Cover © JM Lecollier/LGV SEA

The International Transport Forum

The International Transport Forum is an intergovernmental organisation with 57 member countries. It acts as a think tank for transport policy and organises the Annual Summit of transport ministers. ITF is the only global body that covers all transport modes. The ITF is politically autonomous and administratively integrated with the OECD.

The ITF works for transport policies that improve peoples' lives. Our mission is to foster a deeper understanding of the role of transport in economic growth, environmental sustainability and social inclusion and to raise the public profile of transport policy.

The ITF organises global dialogue for better transport. We act as a platform for discussion and pre-negotiation of policy issues across all transport modes. We analyse trends, share knowledge and promote exchange among transport decision-makers and civil society. The ITF's Annual Summit is the world's largest gathering of transport ministers and the leading global platform for dialogue on transport policy.

The Members of the ITF are: Albania, Armenia, Argentina, Australia, Austria, Azerbaijan, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Canada, Chile, China (People's Republic of), Croatia, Czech Republic, Denmark, Estonia, Finland, France, Former Yugoslav Republic of Macedonia, Georgia, Germany, Greece, Hungary, Iceland, India, Ireland, Israel, Italy, Japan, Korea, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Mexico, Republic of Moldova, Montenegro, Morocco, the Netherlands, New Zealand, Norway, Poland, Portugal, Romania, Russian Federation, Serbia, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, the United Kingdom and the United States.

International Transport Forum
2, rue André Pascal
F-75775 Paris Cedex 16
contact@itf-oecd.org
www.itf-oecd.org

ITF Roundtable Reports

ITF Roundtable Reports present the proceedings of ITF roundtable meetings, dedicated to specific topics notably on economic and regulatory aspects of transport policies in ITF member countries. Roundtable Reports contain the reviewed versions of the discussion papers presented by international experts at the meeting and a summary of discussions with the main findings of the roundtable. This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of International Transport Forum member countries. This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Table of contents

Figures

Tables

Executive summary

What we did

This publication examines the role that ex-post project evaluation can play in enhancing decision-making for future transport infrastructure investments. It presents the results of a Roundtable held in Paris in 2014, which brought together more than 20 experts from 12 countries to discuss the issues surrounding ex-post project assessment and review the experience with different approaches in a number of countries, with case studies from France, the United Kingdom, the United States and Spain at the core.

What we found

Ex-post analysis makes investment appraisal more accurate by validating the inputs to evaluation. Investigating past successes and failures not only improves the quality of appraisal for upcoming projects but enhances accountability. Ex-post assessment may also be critical to maintaining confidence in the value of project appraisal. It is now accepted that errors in traffic forecasts and cost estimates can be systematic rather than incidental. If left unaddressed, such errors can seriously undermine the value of ex-ante analysis.

Collecting detailed information on the impacts of interventions may also help us understand mechanisms beyond those captured in current appraisal methods and improve investment appraisal. Ex-post assessment can also contribute to understanding the impact of infrastructure investment on the economy, although it can be a methodological challenge to separate the net contribution of a transport project from a multitude of other relevant factors. At the same time advances in other areas of research such as the new economic geography, used increasingly to investigate potential economic impacts overlooked by standard transport sector cost-benefit assessment techniques, suggest an expansion of ex-ante evaluation methods may be necessary.

Many countries require ex-post evaluation of transport projects, however, few actually enforce this requirement. This failure is often related to a lack of dedicated funding for ex-post evaluation. Ex-post project evaluations can also be hamstrung by limited availability of relevant data.

For the ex-post analysis to have a good chance of success, the fleeting nature of data needs to be addressed. Data collection needs to be planned at the outset of a project, because post-hoc reconstruction of relevant data sets is often impossible. France has addressed this through the creation of transport observatories, established by law, which collect data, set benchmarks and publish audits of transport projects. Transport observatories have greatly improved data quality for major transport schemes. For the largest investments, a project-specific observatory is now set up at the same time as the project is approved.

In the United States, an online database with information on 100 projects allows decision makers to select a sample of similar projects to the one they are considering and access relevant information on the performance of those projects.

Ex-post evaluation can be applied even before a project is completed. If systematic data collection is put in place at the outset, an evaluation can be carried out at any point in time. In effect such evaluations can also help steer projects while they are still underway. In the United Kingdom, the National Audit

Office is tasked with reviewing the progress of major infrastructure projects and aims to counter project deficiencies as projects progress, reporting to relevant parliamentary committees.

What we recommend

Data collection for evaluation needs to be planned for from the outset

Much of the data needed for accurate assessment will not be available ex-post unless it is collected during planning, procurement, construction and early operating phases. Data collection can be facilitated by the technology used for pricing infrastructure and public transport use. The French permanent observatories have set a new standard for monitoring changes over the long periods for which major projects influence travel demand and territorial development.

Audit transport projects throughout the project stages

By monitoring planning, procurement and operational stages of projects, evaluation can provide those ultimately responsible for the project, such as a parliament, with up to date accounts of the management of the project by the responsible authority. Evaluation can highlight risks to successful project delivery and recommend changes when needed to reduce the risk of project delays or cost overruns.

Use independent organisations to carry out audits of transport projects

There is a strong case for ensuring that audits are carried out by an organisation independent of the project sponsor. An independent audit at an early stage in project planning and delivery reduces the risk of complications only becoming apparent when it is too late to take remedial action.

Recognise the variety of economic goals targeted by transport investments

Transport investment can be motivated by both short-term and long-term economic goals and by multiple combinations of land-use, employment and income growth goals. Evaluations should use methodologies that address the different goals and benefit streams sought. Case-study approaches to evaluation can help decision-makers understand which complementary policies (land-use planning, investment in skills and training, institutional changes, etc.) increase the likelihood of transport investments delivering their objectives.

Involve local partners in providing evidence on performance

Collecting the short-lived data needed to produce pertinent indicators of performance will often be facilitated by involving multiple local partners. This also strengthens public debate on the value of infrastructure investments and supports the primary goal of the ex-post evaluation that will always be to help prepare better ex-ante evaluations.

Chapter 1
Summary of discussions

Tom Worsley - Institute for Transport Studies, University of Leeds, United Kingdom.

This summary of discussions begins by outlining the meaning of ex-post vs. ex-ante evaluation and examining the accuracy of the conventional ex-ante methodology. It then describes two basic systems/approaches to data collection and evaluation. Next it presents the challenges in expanding the conventional investment appraisal to broader development effects and the statistical tools available to evaluate the broader development effects in ex-post evaluation. Lastly, the need to expand the concept of ex-post evaluation to monitoring is discussed and concluding remarks are given.

Introduction

Ex-post evaluation can be used to serve multiple purposes, at the core of which is the improvement of ex-ante analysis:

- It can help policy makers better identify the kinds of projects that work best in certain situations.
- It can identify the effect of concurrent non-transportation investments and their interaction with transportation investments.
- It can contribute to establishing the time frames in which we expect to see economic impacts materialise, thereby helping to set realistic expectations for the effect of investments and economic development.
- We can make use of findings of ex-post assessments to support communication with the public, improve the information provided and support consensus-building efforts.

But ex-post evaluation is often also perceived as a weak link in the assessment process for transport infrastructure and policy initiatives. The assessment methods have tended to rely on ex-ante appraisal, making predictions of how a scheme or policy might perform rather than being based directly on the outcomes of past decisions. Substantial resources have been devoted to improving and updating appraisal methods by those countries that rely on economic appraisal as a means of informing decision makers. Many of these developments, for example, research on the valuation of crowding on public transport or on the valuation of agglomeration and other wider economic benefits, have been the subject of ITF discussion papers and Roundtables (ITF 2014a, 2014b, 2011) and most countries have access to extensive national guidelines on transport investment appraisal.

Many countries also have legislation and guidelines in place that require ex-post evaluation but few actually enforce it. In the United Kingdom, for example, the National Audit Office (NAO) in its reports regularly “reminds” the Department for Transport to conduct ex-post evaluation. The non-execution of ex-post evaluation is not unrelated to the lack of dedicated funds.

Although one might think that ex-post evaluation is a far more straightforward exercise than investment appraisal and the forecasting of what might happen in the future, the experience of evaluators and an increasing volume of research literature would beg to differ. There are a number of reasons for this. When funds are available, there may be limitations in terms of available data. Data collection protocols are seldom introduced before a project starts in order to enable ex-post analysis. Our investment/intervention is very likely not the only change that happened to a certain location or region. Multiple other changes probably occurred in the years since our project became operational, which can make it very difficult to ascertain what the net contribution of our investment/intervention was. This can be a substantial methodological challenge, even when data are available.

These challenges involve methodological and practical issues for the execution of ex-post evaluation. A more fundamental question is: are there other roles for ex-post evaluation? The primary role is informing and improving the ex-ante appraisal process, but can an ex-post analysis also serve accountability in decision-making?

The following four chapters (representing the four discussion papers prepared for the Roundtable) address these questions. Each chapter addresses either a role ex-post evaluation can play in enhancing decision-making or an approach to improving the execution of ex-post assessment.

This summary of discussions begins by outlining the meaning of ex-post vs. ex-ante evaluation and examining the accuracy of the conventional ex-ante methodology. It then describes two basic systems/approaches to data collection and evaluation. Next it presents the challenges in expanding the conventional investment appraisal to broader development effects and the statistical tools available to evaluate the broader development effects in ex-post evaluation. Lastly, the need to expand the concept of ex-post evaluation to monitoring is discussed and concluding remarks are given.

The roles of ex-ante appraisal and of ex-post evaluation in the assessment of transport investments and initiatives

This focus on ex-ante appraisal in transport sector planning provides a contrast to the methods adopted by politicians and their advisors in many other fields of evidence-based public sector decision-making. Here the approach is to conduct well-specified trials or experiments and to monitor the effects of the intervention. Such trials often require the establishment of two or more otherwise similar groups of people or places, and administering the treatment to one segment of this population and not the other. Comparison between the two groups provides evidence of the outcome. If successful, the treatment or policy can then be rolled out to other similarly affected groups. Conversely, if the test cannot be shown to deliver the desired outcomes, it can be discontinued and the experiment either abandoned or alternative treatments can be tested.

Examples of such an approach to decision making are found in the health sector, where drug trials are administered by testing two otherwise similar populations, with one given the treatment to be tested and the other group a placebo. Education provides another example of public policy decisions being based on experiments to test different teaching methods and monitoring the outcomes.

Governments may provide guidance to ensure consistent and cost-effective evaluation methods, such as the UK Treasury's Magenta Book (HM Treasury, 2011). Guidance in these fields is concerned with setting up the trials, selecting a control group, correcting for the placebo effect and avoiding or eliminating confounding effects.

Transport investments and policy initiatives are generally different from the examples in health and education quoted above. Location is an important consideration. Places differ from each other in terms of their populations, their transport networks, and the quantity and location of household and economic activity. The solution to any problem, its cost and its impacts on transport users and others tends to be specific to the location. What works in one place may not work so well in another.

Investment in transport is lumpy and the assets generally cannot be moved to another location if the initial outcome is not a success. So the evaluation "model" of trialling an initiative, careful monitoring and rolling it out if successful, perhaps modified if the evaluation suggests such changes, is not usually a practical option for transport. It is not an option to build a short section of a high-speed rail line, evaluate this trial length and deduce from this the economic and other impacts of a full-length city-to-city scheme. So in the transport field there is a need for good-quality appraisal methods to be used in advance of any decision so as to inform the decision maker of the likely outcomes of that decision.

There is another sound reason for this emphasis on ex-ante appraisal in the transport sector. There is good evidence about how and why people travel and use the transport networks, which are often congested, overcrowded and slow and which decision makers, responding to their voters, want to improve. Based on this evidence, the well-established theory of consumer behaviour can be used to estimate the benefits – their willingness to pay and the surplus in excess of their willingness to pay – that transport users and others derive from an intervention.

Most countries collect transport use and traffic data, which include information on the use that business travellers and industry make of their transport networks. From these data, estimates can be made of the transport cost savings and hence economic benefits gained by the productive sector of the economy, that result from providing more capacity, reducing congestion and providing for quicker, more reliable journeys. Offsetting these benefits are the costs that agents responsible for the provision and operation of the infrastructure and associated services incur, estimated using the judgement of the engineering professionals who are responsible for the design, construction and operation of the scheme.

The evidence about how and why people travel is complemented by information gained from research - Stated Preference and Revealed Preference based studies – about how people value those impacts of a scheme which do not result in direct cost savings, including savings in non-working time, reductions in crowding on trains and improvements in reliability. Incorporating these values into the utility-maximising model of transport users' behaviour shows how transport users will respond, in terms of changing their choice of mode, destination or route, when presented with the option of a new scheme. Research continues to update and improve the values that make up this model and to extend it to cover the external impacts of transport investments, including the impacts on the environment and on accessibility, economic mass and hence on the productivity of urban agglomerations.

The existence of a widely accepted theoretical basis and a rich set of behavioural-based evidence has had an interesting consequence which singles transport out from many other areas of public spending. Most of the resources devoted to research in transport economics have been aimed at methods of improving the ex-ante appraisal process: fewer resources have been targeted at carrying out ex-post evaluation studies. If the appraisal methods are judged to be fit for purpose, because of the plausibility of the theory on which the methods are based and because of the evidence which supports that theory, there is, arguably, no need for evaluation. Moreover, on those occasions when an evaluation is carried out, the results of the evaluation become available long after the scheme has opened. In the intervening years, many improvements have been made to the modelling, forecasting and appraisal methods and the findings of the evaluation may have little relevance to the choices that currently face decision makers

Shortcomings of the appraisal process – errors in forecasts

The ex-ante cost benefit appraisal methods have, quite rightly, been questioned. There have been many cases of errors in forecasts of patronage of a public transport scheme or in forecasts of road traffic volumes. Project costs are often underestimated (Flyvbjerg, 2004 and Flyvbjerg, Holm and Buhl, 2005). Changes are made to the scope of a project, sometimes after decision makers have approved it, without any assessment of the incremental costs and benefits of this change. Questions have been raised about some of the key values in the cost benefit appraisal. For example, criticism has been made of the United Kingdom's approach to valuing rail business passenger time-savings in the context of the economic case for high-speed rail, on the grounds that business passengers often work productively while on a journey.

Some errors are due to exogenous changes – changes in the reference case, for example, on account of the failure to anticipate at the time of the appraisal future changes in fuel prices, GDP or in local or regional economic or territorial development assumptions. Many of the transport models used to appraise schemes assume that land-use changes are not induced by the scheme, since the relaxation of this assumption complicates the attribution of the benefits of the scheme. Differences between forecast and actual values also occur because of errors in the traffic or transport model's prediction of mode, destination or route choice. More recent studies (Eliasson and Fosgerau, 2013, Rose and Hensher, 2013 and Makovšek, 2014) also suggest that the systematic errors in forecasting of traffic and cost estimation may at least in part be the inherent result of the characteristics of the methods applied.

A further source of error arises because the economic appraisal generally omits any assessment of the responses of transport providers to the investment or policy intervention. There has been a systematic failure to anticipate the growth of low-cost air carriers and their response to high-speed rail, or to consider the scenario in which an open access operator might provide on-rail competition to the operator of a new high-speed line either by using the freed-up capacity of the existing line or on the new one. Nor is any account generally taken in the appraisal of transport schemes of technological change – for example, the implications of self-guided cars on the economic case for additional highway capacity.

Systems for evaluation

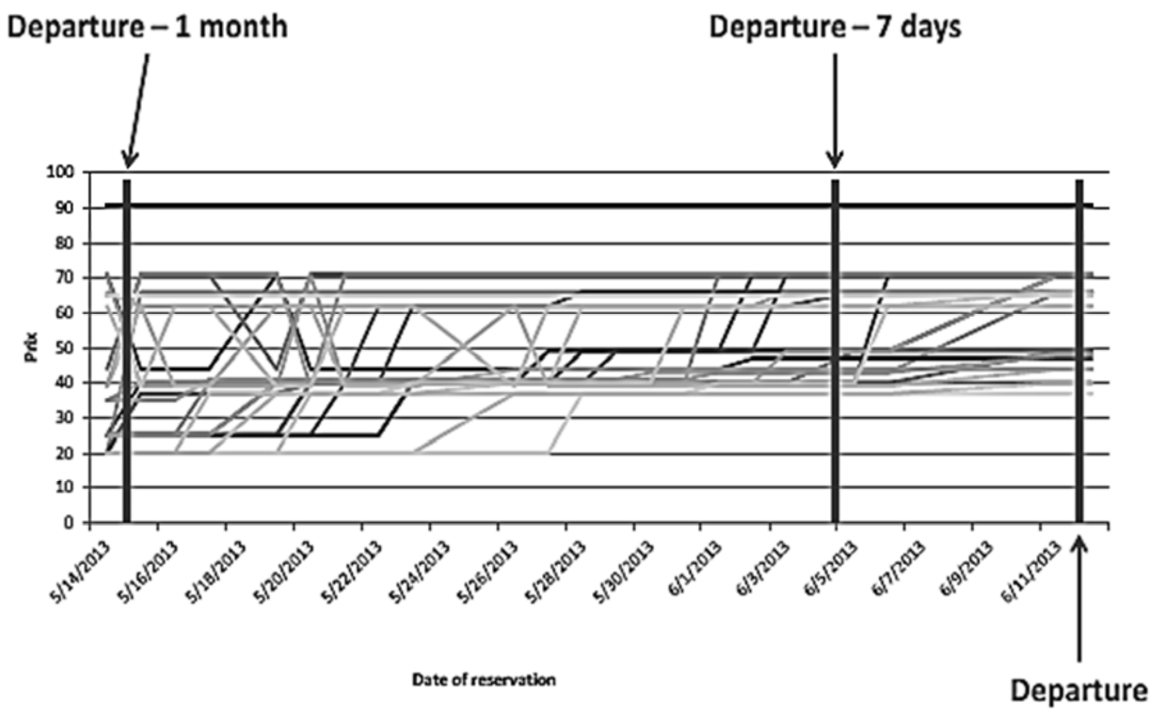
These deficiencies in the appraisal process provide a clear motivation for carrying out evaluation studies. They can help decision makers and their analysts understand some of the causes of error in the appraisal and learn from these. However, evaluation studies need planning well in advance. Transport schemes take many years to plan and construct. The impacts expected of a scheme evolve over a long period and take time to develop. Time-series data are essential for any evaluation. Rent-seeking agents respond well before a scheme has opened, while other effects materialise some years after its completion. Data on networks, services and rail fares, while publicly available at the time of travel, are multi-dimensional, evanescent and complex. As a result they are often impossible to reconstruct post hoc.

In France, there is a long tradition of setting up observatories, established by law (*Loi d'orientation des transports intérieurs* (LOTI) of 1982), for collecting data to facilitate a detailed evaluation of all major transport schemes. Information is published, setting out key indicators of ex-post performance, including data on actual and forecast costs, patronage or, in the case of highway schemes, road traffic volumes. The LOTI audit provides an estimate of the predicted and actual economic internal rate of return for the investment. In addition, the LOTI audits include a high-level assessment of the causes of any observed differences between forecast and out-turn values as a means of informing analysts about potential improvements to appraisal methods. The audit is published so that decision makers and the public have some understanding of the impacts of the project and any errors in the forecasts. LOTI imposed a legal obligation on the infrastructure companies to fund an audit of each scheme, to be carried out by a different department in the organisation to the one that prepared the ex-ante analysis. The report is then reviewed by another body, although within the same line ministry, which may limit the independence of the reviewers.

Despite the discipline imposed by the LOTI-based audit process, certain limitations were identified, in particular on account of the difficulty of attributing local and regional development impacts to the scheme and comparing these ex-post estimates with the ex-ante aspirations. In addition, the time period over which data were collected and the coverage of the data were both judged to be inadequate for a full evaluation of projects.

The establishment of permanent observatories has greatly improved the quality and coverage of data on major transport schemes in France. For the largest projects, an observatory is now set up at the same time as the project is approved, so as to ensure that the "before" conditions are adequately recorded. Data which are available for only a short period of time – for example the fare charged by a rail or air operator using yield management methods to maximise revenue – can only be collected through a systematic and timely monitoring process. Information on attributes of demand, such as the price actually paid by the passenger and the restrictions on the availability of the ticket, improves the basis for analysing and understanding trends in patronage and can explain some causes of errors in the forecasts of demand. Figure 1.1 illustrates the volume of information – price shifts in yield management on a train line Bordeaux-Paris – which becomes unavailable after a few months as nobody stores the information.

Figure 1.1. **An example of the fleeting nature of information: Minimum fares for all Bordeaux-Paris trains, based on length of time elapsed since reservation to departure date**



Source: Paul Joho, in Bonnafous, 2014

A different approach to evaluation has been followed in the United States. A Transport Project Impact Case Studies (T-PICS) database (http://www.tpics.us/) has been set up, based on 100 projects that vary by size, location and motivation. The online database contains details of the characteristics of each scheme and the objectives it was intended to achieve. Information is provided on the indicators of change used to measure its success when compared with otherwise comparable locations, the regional and local area context and data from local interviews about the extent of complementary policies and their role in delivering the outcomes in terms of economic development and local employment.

The T-PICS database (Fitzroy, Weisbrod and Stein, 2014) is intended to provide decision makers and the public with information about the range of outcomes delivered by a sample of projects that are broadly comparable to one that these decision makers might be considering for their locality. The database provides information on the type of scheme and the circumstances in the area served, so as to facilitate the matching of the project under consideration with a number of broadly comparable projects in the database. Information is provided on the factors that are judged to have contributed to the success or failure of the schemes in the sample. Of the 100 projects, 15 failed to deliver any direct employment effects, although these projects did have other impacts on property prices and productivity.

The database is intended to be used well in advance of any detailed design or appraisal of a project, to be considered at the early stage of the decision process as a means of sifting potential options. It also helps decision makers and the public understand whether, in circumstances similar to theirs, complementary policies have been a necessary condition for delivering the objectives of the project, or whether the transport scheme alone has in these other examples been sufficient.

Challenges in extending the forecasting methodology development impacts

Conventional appraisal methods estimate, at best, the first-round impacts of a scheme on transport users – the travel time/generalised cost savings. These time savings make it possible for users of the improved transport network to change where they live, work and carry out the whole range of activities that every household partakes in. The savings in travel time estimated and measured in the ex-ante appraisal are converted into changes in land use, with second-round effects on house prices, labour markets and wages. Firms use the cost savings to reorganise the location of their production, distribution and retail networks; passing cost savings on to consumers. These changes result in a change in the spatial distribution of economic activity and, in some cases, its overall national level.

Policymakers' concerns – in this age of austerity – are with the economic development impacts of transport. The impact of a transport scheme on economic development is far more important to decision makers than the present value of the time-savings – and of more interest to voters. Transport appraisal methods are largely silent about these second-round effects. Transport appraisal experts regard these impacts as difficult to estimate and in any case largely accounted for in terms of the overall size of economic benefits by measuring the direct impacts of the project on users. Some progress has been made through the incorporation in the appraisal of wider economic benefits, covering agglomeration, labour market and imperfect competition effects within the cost benefit framework. But these developments of the cost benefit model provide only limited information on possible changes in the location of economic activity. They do not provide decision makers with a comprehensive assessment of the local or regional impacts of the project on economic development. While there have been advances in integrating models of land-use change into the conventional transport model, these land-use transport interaction models are costly to set up, complex and the predictions from these models have not been subject to extensive testing.

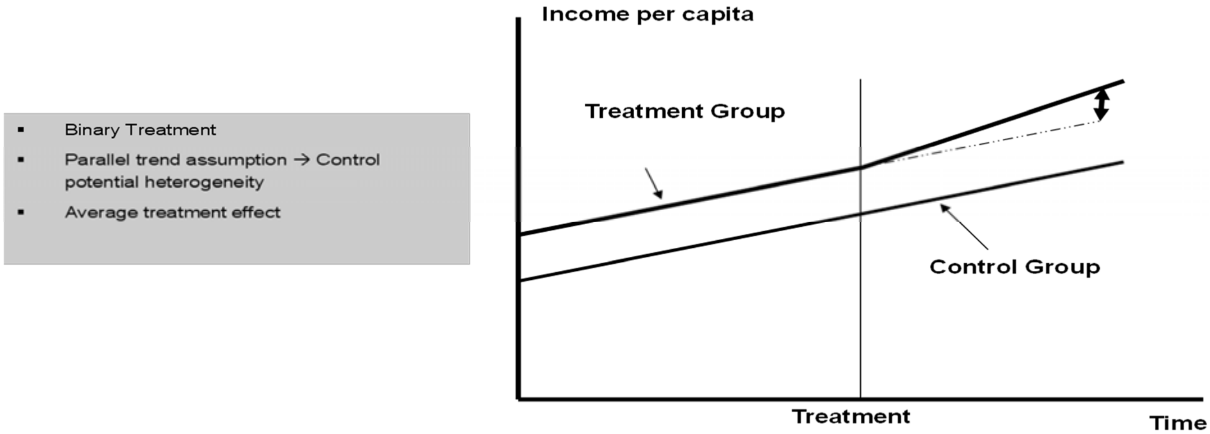
Evaluation, as is shown in the United States and French chapters, has the potential to provide decision makers with information about the economic development effects of a scheme. It can tell policymakers what they want to know about the decisions on transport investment that they or their predecessors have taken. The United States chapter describes the process of collecting an adequate sample of completed schemes and data on them. It describes a process of identifying the extent of any economic development effects in the area affected by the scheme and the circumstances that might have facilitated those effects. While this does not demonstrate causation, it can help decision makers and the public understand which schemes and associated policies might have a chance of delivering the desired economic development effects. The evaluations carried out by the French observatories provide similar evidence about the effects of the investment on land-use change and economic development.

Evaluation-causation

Decision makers need to be confident that a transport investment or initiative they approve will be the cause of the outcomes that they are elected to deliver. Yet establishing that any observed correlation can be interpreted as causation is a challenge. Decisions about where to intervene are not made across a random set of choices. Investment is targeted on those locations where congestion or overcrowding is at its worst, where the accident record is highest, or where there is a particular need to stimulate economic activity. Methods that assume interventions are made at random will produce biased results if applied to transport. In addition, the specific nature of each transport scheme – each place is different in terms of its transport networks on the supply side and of the population and economic circumstances determining demand for that network – means that data on the counterfactual, the hypothesis that the investment had not taken place, is unobtainable.

The discussion paper by Daniel Graham (Graham, 2014) in Chapter 5 describes a means of avoiding these potential sources of bias through the use of statistical inference methods. The aim of the approach is to identify the differences between the areas that benefit from transport interventions and other areas where no such intervention takes place. The effect of the scheme, or of a sample of schemes, is identified by netting off the influence of these "confounding effects" when comparing the outcome in locations benefiting from transport investment with what has happened elsewhere. The use of causal inference models to identify the impacts of transport schemes avoids imposing on the evaluation the assumptions about transport user behaviour that underpin the conventional cost-benefit-based economic appraisal model.

Figure 1.2. **An example of a causal inference method: Differences in differences**



Source: Graham, Brage-Ardao and Melo, 2014.

The paper provided examples of applying these techniques United States to a sample of highway schemes in the United States and to high-speed rail in Spain. In the case of the United States study, the evidence did not support the hypothesis that providing more road capacity in urban areas reduced road congestion or increased productivity in the areas where the investment took place. These findings should not be interpreted as implying that there were no benefits from the investment. The trips induced by the schemes delivered benefits dependent on increasing mobility. But if the objective of the programme was to reduce congestion, it failed to deliver. The same method was used to assess the impact on regional economic performance of the Madrid-Barcelona high-speed rail line. The results showed that, once confounding factors had been accounted for, there was no significant difference between economic growth rates in the target corridor and in other unaffected provinces in Spain. While these examples show the application of the technique, they do not provide any more general lessons about the role of transport in facilitating economic development as they depend on the nature of the projects assessed.

Evaluation to inform the process of procurement, implementation and financing a transport project

The term ex-post immediately suggests that something will take place after an action/event is completed. With regard to evaluation of transport investments, ex-post usually refers to evaluation after the completion of the construction phase or later. But the term is generic. An evaluation can be executed at any point during the project development and operation.

Indeed, one part of the work of the United Kingdom's National Audit Office is concerned with monitoring the progress of major infrastructure schemes or policy interventions prior to their completion

(Barker, G. et al., 2014) note that the NAO provides the United Kingdom Parliament's Public Accounts Committee with the evidence that they need to understand whether government departments and their ministers are obtaining value for money in their spending and investment programmes. The PAC has the right to investigate spending decisions and to ask questions of departmental officials and ministers about whether they are delivering, or are likely to deliver, value for money.

Major infrastructure projects take many years to deliver and the NAO will often review progress during the implementation and construction phases if, in its view, such an intervention might be beneficial. The NAO seeks to ensure that the discipline imposed on project sponsors by the process of cost-benefit appraisal is not relaxed once the project has been approved. For example, changes in the scope of a major project frequently arise after it has been approved and, while such variations might not be subject to economic appraisal, the NAO's examination can act as an incentive on project sponsors and funders to manage the costs of such alterations to the initial contract.

The NAO's role is not restricted to reviews during the phase of project implementation. It also carries out evaluations of a sample of major projects after they have opened, to understand whether they have delivered the expected benefits in terms of traffic volumes, time savings and regeneration impacts. These post-opening audits provide the NAO with the opportunity to make recommendations to the Department for Transport about improvements to its appraisal methods.

In general, it can be noted that both the NAO and T-PICS's approaches to ex-post assessment rely heavily on interview and interaction with people and organisations directly affected by the investment(s) being assessed.

Conclusions

While economic appraisal tends to follow a standard approach based on the theory of consumer behaviour, there are several different approaches to conducting an ex-post evaluation. The model chosen depends on the purpose of the evaluation, and the linkage between the purpose of the evaluation and the method for the evaluation is examined here.

Purpose of evaluation – audit during implementation

By monitoring these stages of project development and delivery, evaluation can provide those responsible for approving the funding of the project (Parliament, in the case of the United Kingdom's NAO) with an account of the management of the project by the transport department or other responsible authority. The evaluation can highlight risks to successful project delivery and recommend changes to tighten overall management so as to reduce the risk of project delays and of cost overruns. This also provides lessons for management of future projects.

There is a strong case for ensuring that such an audit is carried out by an organisation that is independent of the project sponsor. The United Kingdom NAO has noted that independent audit at an early stage in the delivery of a project reduces the risk of any over-run only becoming apparent when it is too late to take remedial action.

Experience in the United Kingdom suggests that the threat of external audit imposes a stronger discipline on the Department for Transport and supports the Department's process of internal audit.

Many countries have within government an audit department that is independent of the department responsible for transport and other spending departments. Not all are as active as the United States GAO or the United Kingdom's NAO, or have the support of the United Kingdom Parliament's PAC to ensure that the legislature is aware of their findings. There is a case for strengthening the role of independent government audit departments in many jurisdictions and encouraging them to contribute to the objective of better management of the delivery of transport projects, thereby reducing cost overruns and encouraging those responsible to improve cost and demand forecasts.

Purposes of appraisal – did the project deliver its objectives?

Policymakers need to know whether a project has delivered its objectives. Reassurance that the outcomes delivered by the investment are in line with impacts that were predicted at the time the decision was made provides policymakers with the confidence to continue with or extend the transport investment programme. It can also provide evidence to support the ex-ante appraisal methods used to justify the initial decision to approve the investment.

Evaluation can serve a number of objectives. These include:

- Identifying the differences between forecast and actual values for project costs, road traffic or public transport patronage volumes and other key impacts such as carbon dioxide emissions.
  - One approach is to provide a high-level assessment of the possible causes of any differences, such as the impact of the economic downturn, not anticipated when the scheme was approved. While such an approach might not provide the level of detail needed for a full comparison of the counterfactual with observed demand, it can help to improve the allowance to be made for risk and uncertainty in future decisions.
- Validation of the transport model used to estimate mode, route and destination choice.
  - Evaluation is only used infrequently for the purpose of validating the transport model adopted to appraise a scheme. In many circumstances models become obsolete by the time the scheme has opened and become established, or the data and model runs are no longer available. Decision makers are unlikely to be willing to prioritise funding for a new best-practice model to represent a counterfactual that, because the scheme has been built, could not exist.
- Understanding some of the key relationships between transport investment and its outcomes through econometric analysis of a sample of schemes.
  - Several studies have been carried out in the United Kingdom to examine the link between changes in transport costs and changes in employment and productivity.
- Determining the extent to which the decision-makers' objectives for the scheme or programme were delivered.
  - These objectives might include economic development, defined in terms of additional jobs in the area, increases in productivity or take-up of industrial sites, population growth, encouraging housing developments, shift from car to public transport or to active modes, and reductions in carbon intensity or other environmental goals. This forms the main focus of most evaluations of transport schemes.

– A point mentioned in the Roundtable discussions is also the need to properly define the project's objectives in the first place. For example, one can only speak about having a problem if there is at least one solution available. If this is not the case, then our problem is only a fact. An example would be a road connection, which is congested and winding and the objective was to reduce congestion. An investment into a straight road would reduce travel time, but it would not reduce congestion (due to induced traffic, for example). Thus the success of an intervention is also dependent on the proper definition of the objectives.

– It should be noted that the objectives of some projects do not necessarily coincide with the value-for-money concept in economic theory or with the purpose of maximizing socio-economic welfare. For example, it is entirely possible that it is in the public interest to improve a transport connection to a dislocated community even if this intervention does not increase the social welfare of the nation (i.e. the project is assessed to have a negative net present value). In this context, the challenge for the ex-post evaluator is how to devise a clear measure of success, which would differentiate between abuse of power and social necessity.

Evaluation to determine whether an investment or a programme of projects has delivered on policy objectives does not necessarily establish any causal link between the project and the outcomes. Even when the evaluation takes the form of validating the transport model and the responses to the changes in transport costs as a result of the investment, there is an implicit assumption in such an evaluation about the structural form and behavioural relationships within the model. The use of causal inference methods to attribute causation is particularly suited to transport investments because the decision about where and when to invest is not a statistically random one but one based on the greater perceived need in the location where the investment is made. In addition, the approach does not require any hypothesis about the counterfactual or about transport users' responses to changes in transport costs. While its application to transport schemes is still at the research stage, it has the potential to provide a valuable method for better evaluation.

Data are essential in any evaluation method. The French permanent observatories have set a new standard for monitoring changes over the long period during which a major new scheme influences travel demand and territorial development. The process of data collection needs to encompass the technological developments that have made possible changes in the way public transport is priced and accessed and how travel time is used.

Each approach has strengths and weaknesses – in terms of costs of collecting the data, the target audience, and how it helps to improve decision making and provide for more informed decision makers. Case-study approaches in particular help transport ministers understand which complementary policies (land-use planning, skills and training, institutional changes, etc.) might increase the likelihood of transport programmes or schemes delivering their objectives.

The way forward?

The discussion papers presented at this Roundtable each introduce a unique element:

- NAO's assessments, already in the project preparation and execution stages, enable monitoring on the spot, while the project is still running, corrective action still possible and accountability still enforceable.

- The T-PICS case-study approach, among other things, involves field research and interviews. Potentially, it provides a source of information that can explain the relations between cause and effect.
- Transport observatories represent a robust approach to ensuring that short-lived data can be preserved and used in the ex-post evaluation.
- Causal inference methods provide a tool for the objective assessment of effects (but do not explain the functional relations which lead to them).

The latter three approaches also enable an improvement or calibration of structural models used in ex-ante assessments in terms of functional relations applied, as well as the accuracy of predicted outcomes.

Combining these elements together may represent a holistic approach to ex-post evaluation, creating a high-resolution image of the outcomes we might expect from transport investments and other interventions. What perhaps remains to be determined in the future is a standard or best practice of ex-post evaluation that would combine these methods and their unique elements in the most efficient and cost-effective way.

References

Barker, G., G. Beardsley and A. Parsons (2014), "The National Audit Office's value-for-money assessment of transport investments", *International Transport Forum Discussion Papers*, No. 2014/12, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/5jrw2z0kvgg0-en

Bonnafous, A. (2014), "Permanent observatories as tools for ex-post assessment: the French case study", *International Transport Forum Discussion Papers*, No. 2014/10, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/5jrw2z3ttwf7-en

Economic Development Research Group (EDR Group), Transportation Project Impact Case Studies http://www.tpics.us/.

Eliasson, J. and M. Fosgerau (2013), "Cost-overruns and demand shortfalls – deception or selection?" *Transportation Research Part B: Methodological*, vol. 57, pp. 105-113.

Fitzroy, S., G. Weisbrod, N. Stein (2014), "TPICS, TIGER and US Experience: A Focus on Case-Based Ex Post Economic Impact Assessment", *International Transport Forum Discussion Papers*, No. 2014/11, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/5jrw2z2fqgf1-en

Flyvbjerg, B., M.K. Holm and S.L. Buhl (2005), "How (In)accurate Are Demand Forecasts in Public Works Projects? The case of transportation", *Journal of American Planning Association*, Spring, Vol. 71, No. 2.

Flyvbjerg, B. (2004), Procedures for Dealing with Optimism Bias in Transport Planning, Department for Transport, London.

Graham, D.J. (2014), "Causal inference for ex post evaluation of transport interventions", *International Transport Forum Discussion Papers*, No. 2014/13, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/5jrw2z02frjk-en

Graham, D.J., R.C. Brage-Ardao and P. Melo (2014), "Quantifying the Economic Development Impacts of Major Transport Infrastructure Projects: A Case Study of High-Speed Rail in Spain", Presentation at the Roundtable on Ex-post Assessment of Transport Investments and Policy Interventions, 15-16 September 2014, OECD, Paris.

HM Treasury (2011), The Magenta Book: Guidance for evaluation, UK Government, https://www.gov.uk/government/publications/the-magenta-book.

ITF (2014a), *"Appraising Transformational Projects: The Case of the Grand Paris Express"*, International Transport Forum Policy Papers, No. 3, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/5jlwvzhlj935-en

ITF (2014b), *Major Transport Infrastructure Investments and Economic Development*, ITF Round Tables, No. 154, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/9789282107720-en

ITF (2011), *Improving the Practice of Transport Project Appraisal*, ITF Round Tables, No. 149, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/9789282103081-en

LOTI (1982), Loi d'orientation des transports intérieurs, France, http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006068730.

Makovšek, D. (2014), "Systematic construction risk, cost estimation mechanism and unit price movements", *Transport Policy*, vol. 35, pp. 135-145.

Rose, J.M., and D.A. Hensher (2014), "Toll roads are only part of the overall trip: the error of our ways in past willingness-to-pay studies", Transportation, vol. 41 no. 4, pp. 819-837.

Chapter 2
Permanent observatories as tools for ex-post assessment: The French case study

Alain Bonnafous - Emeritus Professor, Lyon University/Laboratoire d'Économie des Transports (LET), Lyon, France.

In France, the obligation to conduct an ex-post assessment is relatively recent (1982) but it has, in fact, revealed some methodological problems that flow from the retrospective nature of the exercise.

This chapter shows that some of those difficulties can be overcome through the establishment of permanent observatories. It presents examples of such observatories, with a particular focus on motorway ("autoroute", or "freeway" in North American usage) investments. A particular case is then investigated, concerning the socioeconomic observatory on the effects of the Sud-Europe-Atlantique (South-Europe-Atlantic) high-speed train line that is now under construction.

Besides an overall presentation of the mechanism, the report deals in particular with the metrological precautions that must be taken in order to monitor the multimodal offer of transport for the areas concerned.

Introduction

The decision to make public funds available for studies or research subsequent to the coming on-stream of new infrastructure in Europe was first made in the 1970s. At that time, the primary concern was not to conduct a socioeconomic assessment in the strict sense, but rather to identify the effects that had not been taken into account in the standard evaluation procedure then in force, and that were referred to as "indirect" or "structuring" effects of infrastructure investments. The European Conference of Ministers of Transport (ECMT) anticipated (ECMT, 1969) and then supported these initiatives (ECMT, 1975).

Clearly, then, the first question was how to measure the impacts not caught in the socioeconomic calculation, i.e. those that might reveal further advantages from the investment. This concern not to underestimate advantages was an important consideration at that time. At the end of the 1960s, Europe had a series of uncompleted motorway networks that were built as an emergency response to relieve major traffic bottlenecks, and the resulting reduction in congestion implied very high returns in socioeconomic terms. Because the highway segments needed to complete that priority network presented more modest rates of return, it was legitimate for the sponsoring administrations to seek out advantages for the projects that might have been overlooked, if only to be able to select the "best" projects.

Attention therefore turned to drawing lessons from concrete cases and studying regional trends in areas where transport conditions could be significantly improved through large-scale infrastructure investments. A number of ex-post studies were thus conducted on a number of major projects (ECMT, 1969 and 1975). These include the Autostrada del Sole motorway between Milan and Naples, inaugurated in 1964, the M62 motorway in the United Kingdom (Lancashire-Yorkshire Motorway), the Severn Bridge linking Wales and England, and the A7 Autoroute along the Rhône Valley, south of Lyon.

A scientific comparison of the results of the studies yielded some lessons on which there was fairly broad consensus. First, each case study presented so many specific features that this in itself constituted a problem. Thus, a motorway linking Lombardy, one of Europe's richest regions, with the Mezzogiorno, which contained some of the poorest areas, raised the question of regional rebalancing; the A7 autoroute (*Autoroute du Soleil*) in France or the M62 in the United Kingdom, which served areas of roughly equivalent development, posed the problem of "transit" effects and, in the second case especially, the question of effects on the urban hierarchy: the Severn Bridge linked two regions so closely as to raise the issue of the merger of market areas.

Secondly, all the studies pointed to the lack of systematic impacts from these major projects and the difficulty, when looking at the effects ex-post, of specifying the role of the new infrastructure in relation to other factors. For example, although statistics might reveal a certain concentration of activities near motorway interchanges, could this be interpreted as an effect of infrastructure or as representing the attractiveness of urban centres located, as a rule, near interchanges?

The studies from that time ultimately concluded that if regional development apparently induced by these investments was observed it could reasonably be assumed that this represented a displacement of activity towards a better-served region, rather than any net creation of value. Consequently, what was needed was not so much to evaluate some regional enrichment as to learn to use infrastructure investments as territorial development tools in the sense of regional rebalancing (Bonnafous, 1979).

This, of course, tended to shift the initial problem from poorly estimated value creation to another, no less difficult problem. That was to understand the mechanisms of regional development and the specific role that the availability of transportation could play. It was quite natural, then, to see some initiatives made to apply a methodological approach to this problem. In France, these efforts led to the establishment of permanent observatories of socioeconomic developments in regions served by new infrastructure.

These initiatives were launched at the end of the 1970s. Shortly afterwards, in the 1980s, legislation was adopted requiring an ex-post evaluation of major projects (or of significant technological innovations), an evaluation known as the *Bilan* LOTI, the "LOTI audit" or balance sheet (LOTI stands for *Loi d'orientation des transports intérieurs*, or "Domestic transport planning law" in English), the principals and selected results of which are presented in the next section. As we shall see through some examples of such audits, a number of lessons can be drawn from these experiments.

At the present time, the usefulness of these LOTI audit reports is generally admitted, although their authors have themselves identified the limitations of such retrospective investigations. In the section "From ex-post evaluations to the notion of a permanent observatory", we shall identify what may have been the first initiatives to create permanent observatories in France, which related to a problem complementary to that of the LOTI audits and which contained information of a different nature. These initiatives were initially conducted by academic circles, but were then supported by the autoroute operators. A presentation of some of them allows us to appreciate their methodological contribution. Recognition of these contributions naturally sparked the idea of taking a more ambitious and long-term approach that would combine the advantages of permanent observation and of ex-post evaluation. The case study of the socioeconomic observatory of the LGV Sud-Europe-Atlantique, which is unique as it is supposed to function until 2027, will be presented in the subsection "The socioeconomic observatory for the LGV LISEA", with some details on the particular methodological difficulties revealed by its implementation.

French experience with "LOTI audits"

For a long time it seemed to be the practice in France that the file on a major investment was effectively closed once the infrastructure entered into service. Officially, the notion of ex-post evaluation was introduced only in the 1980s.

The regulatory mechanism

The term "LOTI audit" reflects the fact that these ex-post evaluations of major transportation projects became compulsory with the passage of the "Domestic transport planning law" of December 1982. In 2010, that obligation was reiterated and spelled out in a new "transport code", which updated all the legislative provisions organising the transportation sector in France.

First, there is an obligation for the sponsoring authority or project owner (*maître d'ouvrage*) to prepare an ex-ante evaluation and an ex-post assessment of major infrastructure projects that are publicly financed, if the project cost exceeds EUR 83 million. The project owner is responsible for conducting the assessment and has from three to five years after entry into service to produce it. This ex-post balance sheet is published, notably on the project owner's website; for example, the site of the RFF (Réseau Ferré de France) for railway projects. It gives rise to an official *avis* (opinion) from the reviewing authority, which includes the inspectors general of the ministry responsible for transport. This opinion is also published on the website and in two newspapers of national circulation.

The purposes of these ex-post reports are explicit. They are to:

- Inform the public about the status of the operation, and in particular about the discrepancies between forecast and actual outcomes, and to reveal the causes of those discrepancies.
- Render accounts for the use of public funds by evaluating ex-post the economic, social and environmental efficiency of the investments made.
- Provide feedback for improving the methods of ex-ante project evaluation.

This is, then, a very conventional approach that consists essentially of producing a critical analysis of the forecasts and assessments made prior to the decision to proceed with the project. This can cause a problem within the sponsoring institution, which will have been responsible for these forecasts and achievements and will then be asked to render a critical assessment of them. To prevent the exercise from being self-serving, different agencies are involved. Thus, within the RFF, the LOTI audit for railway projects is entrusted to the "audit and risks division", which is functionally independent of the units responsible for ex-ante evaluations. Lastly, this report is enriched by the opinions of the independent body identified earlier.

The lessons to be drawn from these exercises must be distinguished according to the nature of the investment. The complete list of official audits available is shown in Annex 2.1. Given the purpose of this report, we shall focus only on the two main investments involving toll motorways (autoroutes under concession) and high-speed rail lines.

LOTI audits for concessioned autoroutes

The authors of the LOTI audit reports and the official opinions place particular stress on the quality of traffic forecasts. Thus, the official site presenting these opinions specifies that the purpose is to "compare the forecasts on which the previous choices were based and the actual outcomes of those choices". To summarise the main outcomes, we shall look at the discrepancies between the forecast and actual economic internal rates of return (EIRR). Table 2.1 shows the most important autoroutes covered by LOTI audits and indicates the apparent reasons for the main discrepancies between planned and actual outcomes.

Table 2.1. **Ex-ante/ex-post comparisons of LOTI balance sheets: Economic returns for the main concessioned motorways**

Motorway Segment	Ex-ante/ex-post discrepancies	Principal explanation
A49 **Grenoble Valence** **(Opened in 1992)**	Initial forecast EIRR: 14 % Initial observed EIRR: 19 %	Costs closely controlled and traffic flows higher than forecast.
A57 **Cuers-Le Cannet des Maures** **(Opened in 1992)**	Forecast EIRR: 20 % Ex-post EIRR: 14.8 %	Traffic flows higher than forecast, but very great cost overruns.
A54 **St Martin de Crau-Salon de Provence** **(Opened in 1996)**	Initial forecast EIRR: 30 % Initial observed EIRR: 15.4 %	Costs controlled but traffic flows far below forecasts.
A837 **Saintes-Rochefort** **(Opened in 1997)**	Initial forecast EIRR: 13 % Initial observed EIRR: 5 %	Traffic flows far below forecasts.
A83 **Nantes-Niort** **(Opened in 2001)**	Ex-post EIRR: 15 % Greater than forecast EIRR not specified in the audit report	Cost overruns more than offset by higher-than-expected traffic flow.

Motorway Segment	Ex-ante/ex-post discrepancies	Principal explanation
A20 Brive-Montauban (Opened in 2003)	Forecast EIRR: 8 % Ex-post EIRR: 8 %	Cost overruns offset by higher-than-expected traffic flow
A28 Alençon-Tours (Opened in 2005)	Forecast EIRR: 15.5 % Ex-post EIRR: 10 %	Cost overruns

Source: Bilans LOTI, cf. Annex 2.1.

Generally speaking, traffic forecasts tend to underestimate rather than overestimate future traffic flows. This reflects the fact that the uncertainties inherent in forecasts have to do primarily with the allocation of traffic across the highway network, and more particularly between toll and toll-free routes. As the intermodal dimension is of little importance, we may conclude that traffic allocation represents the most consistent modelling technique.

The disappointing EIRRs are generally due to the fact that costs are underestimated to such an extent they cannot be offset by higher-than-expected traffic volumes. This is a complex problem. If we examine the cost estimates in detail, the discrepancies can indeed be glaring for some projects. In the case of the A57 autoroute, for example, right-of-way acquisition costs exceeded forecasts by 80% and safety equipment costs by 108%. On the other hand, civil works costs (a very significant item for investments in a hilly region) came in 30% below estimate. It must be recalled, though, that this project was prepared more than 25 years ago, and the LOTI reports noted that, with the current system of competitive tendering in force in France since 2001, the discrepancies are more reasonable.

It will be noted, however, that these ex-post audits say little about the impact on the local area, a reticence that contrasts with the importance of these indirect effects in the considerations preceding the decision. The key point relates to the local developments that these new investments are supposed to bring with them. The audits are generally confined to comparing forecast and actual employment effects generated by the construction phase and, after entry into service, by operation of the autoroute and its service plazas.

LOTI audits of the new high-speed rail lines (lignes à grande vitesse, LGV)

It is not necessary here to select the most important projects in order to draw some lessons from these LOTI audits, as there have been only six, all of them large-scale and costly. The main discrepancies between forecasts and outcomes are summarised in Table 2.2 below for the comparative EIRRs, as well as for the financial internal rates or return (FIRR), which by law must be calculated ex-ante for LGV projects.

The differentials between forecast and actual internal rates of return are the result in part of poorly controlled costs, as well as of traffic volumes that have fallen short of expectations. The most frequent cost overruns relate to infrastructure, generally stemming from supplementary investments or tightness in the public works market. They also relate to rolling stock purchase and operating costs, which seem inexplicably unpredictable.

With respect to traffic, it is clearly important to identify the causes of error, especially those relating to traffic increases beyond the benchmark. For LGVs, traffic forecasting is more difficult than for autoroutes, where the problem is primarily one of allocating traffic across the highway network. In the case of high-speed rail lines, the task is to project the traffic induced by a significant movement in service, as well as a modal shift that is harder to control. While the forecasts were quite accurate for the

LGV Atlantique, they were overestimated for the LGV Nord-Europe. There, the LOTI audit report offers no lessons as to the mistake that was made in calculating the modal shift from highway to rail for short runs such as Paris-Lille. As to air competition, it reflects the steep but unanticipated fare cuts on Paris-London and Paris-Brussels flights. In the first case, there was a mistake in model specification, while in the second case there was an erroneous hypothesis concerning an exogenous parameter.

Table 2.2. **Ex-ante/ex-post comparisons of LOTI audits: economic rates of return for high-speed rail lines**

LGV Project	Ex-ante/ex-post differentials		Principal explanation
	EIRR	FIRR	
LGV Atlantique (Opened in 1992)	Expected: 23.6 % Ex post: 14 %	Expected: 12.9 % Ex post: 8.5 %	Traffic and revenues higher than forecast, but heavy cost overruns (more than 20%).
LGV Nord-Europe (Opened in 1993) (extended to Belgium in 1996)	Expected: 20.3 % Ex post: 5 %	Expected: 12.9 % Ex post: 2.9 %	Traffic below forecasts; revenues close to forecast thanks to increased fares, but 20% infrastructure cost overrun.
Interconnexion Ile-de-France (Opened in 1994)	Expected: 14.1 % Ex post: 6.9 %	Expected: 22.3 % Ex post: 15 %	Traffic increases below forecast and overruns on rolling stock and operating costs.
LGV Rhône-Alpes (Opened in 1994)	Expected: 14 % Ex post: 10.6 %	Expected: 9 % Ex post: 6.1 %	Benchmark traffic below forecast and overruns on rolling stock and operating costs.
LGV Méditerranée (Opened in 2001)	Expected: 11 % Ex post: 8.1 %	Expected: 8 % Ex post: 4.1 %	Benchmark traffic close to forecast but lower traffic increases and overruns on rolling stock and operating costs.
LGV Est (Opened in 2007)	Expected: 8.5 % Ex post: 4.2 %	Expected: 7.2 % Ex post: 5.9 %	Cost overruns (+20.2 %) partially offset by higher-than-expected traffic

Source: Bilans LOTI, cf. Annex 1.1.

Some methodological lessons

The lessons that can be drawn from these LOTI audits are of general application and can be summarised in two points. First, there is a real problem with cost control, resulting in glaringly excessive overruns, particularly for LGV rolling stock and operating costs. The methodological response is to strengthen the risk assessments by taking cost uncertainties more thoroughly into account. This has been done for the ambitious Tours-Bordeaux LGV project (EUR 7.8 billion), discussed below.

Second, the analysis of traffic forecasting errors generally involves choosing between a few common explanations: a mistaken benchmark estimate, most often due to faulty macroeconomic assumptions; errors in modal distribution, often due to mistaken assumptions about the competitive context; and errors in traffic allocation or modal distribution resulting from faulty modelling.

Erroneous assumptions lead to the same recommendation as costing errors, and call for development of a risk analysis methodology. Model specification errors are a much more intractable problem, as the ex-post evaluation is done several years after entry into service and does not always have the statistical data needed for a rigorous distinction between exogenous assumption errors and specification errors.

Taking intermodal competition as an example, it is nearly impossible to reconstruct an airfare history, given the multitude of different fares resulting from the airlines' "yield management" approach.

In this case, only the operators will have a chronological series for the different airfare levels and their weighting; and commercial secrecy prevents them from disclosing these data (a point that applies as well to rail operators). The only methodological response to this problem is to collect real-time data, which brings us to the need for a permanent observatory.

Similarly, beyond traffic considerations and the elements of economic and financial return, the LOTI audits have attempted to provide some information on the economic impacts of projects. The idea is to compare what could be observed following entry into service with what had been promised in the project's initial plan. Two extensive chapters are typically devoted to these issues in the documentation that is used for the public hearing, prior to the official "declaration of public utility". One of them addresses the economic activities surrounding the construction phase itself, and then the operation of the facility, and they are assessed in terms of their direct job-creation effect. The other chapter looks at the anticipated indirect economic effects of all kinds, which are generally highlighted in the public hearing documents. In both cases, the idea is to ensure positive fallout for the regions traversed by the infrastructure in order to counter frequently vocal opposition, including the NIMBY ("Not In My Back-Yard") phenomenon.

With respect to the promised "direct employment creation", this is not generally a focus of observation. For example, for the six LGV projects listed in Table 2.2, the public hearing document mentioned a specific number of anticipated new jobs in each case. Only one of the six LOTI audits contains an ex-post estimate. For the LGV Atlantique, a specific study cited in the LOTI audit estimated actual employment at the work site at 21 600 job-years, compared to the forecast of 33 400, and 11 000 job-years for the rolling stock, and the projected 20 000. In defence of the LOTI report authors, it should be noted that retrospective reconstruction of such effects is dauntingly complex in the absence of real-time data.

As to the indirect economic effects (which may be highly variable, as we shall see in the following sections), the public hearing documents are often very optimistic about the expected fallout for regional development. The LOTI audits are more discreet on what is observed. The LGV North was supposed to attract new development to Lille, but the balance between activity relocating to Lille from Paris and movements in the opposite direction is very difficult to establish, according to the official audit report. The LGV Rhône-Alpes was supposed to encourage air-rail trade-offs at Saint-Exupéry Airport, but in fact only 0.5% of air traffic has been affected. The LGV Méditerranée was supposed to promote specific developments around the new train stations at Valence, Avignon and Aix-en-Provence but the LOTI audit revealed nothing significant. The LGV Est was supposed to reinforce the role of Strasbourg as a European capital, but the LOTI audit contents itself with quoting the Chamber of Commerce of Strasbourg to the effect that the city now ranks in the class of big cities for the TGV.

In fact, for direct and indirect effects alike, these retrospective assessments have the greatest difficulty in establishing the statistics or the facts needed to identify them. Thus, each analysis is based on data that at best require a great deal of reconstitution and, at worst, do not exist at all. The LOTI audits are thus reduced to hunting around for elements in specific studies that may have been undertaken at the initiative of local or national agencies. The methodological response to this problem lies in the implementation of permanent observatories.

From ex-post evaluations to the notion of a permanent observatory

It is not by accident that some operators, in particular autoroute concessionaires, have gradually been drawn toward permanent observation arrangements. These initiatives have generally recognised that they were following the example of a series of academic studies.

Early experiments

The first French experiment with permanent observatories was purely academic. It involved the implementation of a system called SPOT (*Système Permanent d'Observation sur le Triangle Lyon-Chambéry-Grenoble).*

From the methodological viewpoint, this operation was largely inspired by the studies that will be mentioned here and that were conducted by the same research team looking at the indirect effects of the A7 autoroute in the Rhône Valley. That highway was put into service between 1962 and 1968. The retrospective study thus benefited from surveys that happened to be conducted in those same years. Moreover, an ex-post survey in 1975 provided valuable information on the demographic and economic situation seven years after the last autoroute segments were brought into service. Many other reports were also assembled and analysed for the periods before and after 1968.

The vast statistical bases thus constituted were used to establish an ex-ante typology of municipal dynamics between 1962 and 1968 and an ex-post typology of those dynamics between 1968 and 1975. The methodology was based on comparing these two typologies in order to identify significant shifts that might reveal the effects of the autoroute (Plassard, 1977). Thus, of the more than 400 communes involved, some saw changes in their typical dynamics, probably influenced by their relative proximity or distance from an autoroute interchange. A field survey was then undertaken to validate or discard one or other interpretation of the role of the autoroute in bringing about these changes.

The field survey, though, was limited to some 30 communes designated by these statistical treatments. In-depth interviews served to validate and flesh out some of the results. For example, as the autoroute ran along the left bank of the Rhone, the right bank fell into something of an economic depression, as borne out by local residents, who felt that the other bank was too attractive. Another example is the revival of economic activity along the old highway (the famous *Nationale 7*), which was relieved of much of its congestion with the opening of the A7, and where managers of newly established businesses confirmed that this was a key factor in their choice of location. However, the field studies, which were conducted five or six years after the A7 opened, sometimes faced difficulties in gaining access to information; some players could no longer be reached, others reconstituted economic performances that were in total contradiction with the statistical record, some significant information was not conserved, etc.

These difficulties led the researchers to undertake a comparative investigation that could be conducted in real time so as to catch information before it disappeared. In support of an ambitious programme of the National Scientific Research Centre (CNRS) on the observation of social change, the LET *(Laboratoire d'économie des transports*, the "Transportation Economics Laboratory") was selected to institute, on an experimental basis, a permanent observatory for the geographic area served by the new autoroute linkages in the Rhône-Alpes region. As it happened, this research programme was launched at the same time as the opening of the A43 autoroute Lyon-Bourgoin-Chambéry (completed in 1974) and the section of the A48 linking Bourgoin and Grenoble (completed in 1975).

This SPOT mechanism enabled close monitoring (Gérardin et al., 1981) of a sufficient number of indicators to reveal demographic or economic changes in the 245 communes that make up the triangle under observation. For example, they allowed the following themes to be addressed:

- Industrial activity and its transformations were monitored by a monthly survey of high- or medium-voltage electricity consumption in each commune. This indicator was particularly useful for revealing local trends in industrial output, either up or down. In the case of business closures or relocations, it was generally possible to identify important trends several months before they actually occurred, and in this way to conduct in-depth interviews with the players concerned.

- Municipal budgets were systematically recorded, making it possible to identify those communes that were purchasing new equipment or that were experiencing particular growth, in terms of housing or business activities.
- Building permits and property transfers were also recorded, as a supplement to this coverage.
- A set of 21 communes comprising the *Ville Nouvelle* de l'Isle d'Abeau was subjected to especially close monitoring, entailing a number of field studies. This revealed the relative failure to shorten home/work-commuting distances, which the design of the new city was supposed to promote.

From a methodological viewpoint, the SPOT mechanism proved to be a powerful explanatory tool. For example, this area had historically been characterised by a large number of textile industries scattered across small towns and villages. It so happens that the opening of the autoroute coincided with a profound crisis in the French textile industry, which was already facing stiff competition from emerging countries. If a retrospective study had been conducted at the end of that time, the effects of the autoroute would have been found to be offset by employment losses in the hinterland and barely sustained activity near the interchanges. Field studies conducted immediately after abrupt changes appeared in certain types of industrial output made it possible to sort out the local effects of the textile crisis from the effects on the attractiveness of activity zones located near the autoroute.

More generally, this experience showed that permanent observation could be more useful than ex-post studies, and that it could be an effective response to fleeting phenomena, the disappearance of data, and the loss of stakeholder memory.

The involvement of the autoroute companies

Based on its initial results, the Rhône-Alpes autoroute company (AREA) was induced to co-operate with this first experiment, and subsequently to provide financial assistance. Its interest in such an observatory went beyond the elements determining the project's economic profitability, and other autoroute companies came to share this view.

In France during the 1980s and 90s, the decision-making process regarding major capital projects changed significantly, to make greater room for debate and consensus building. As a result, there was a demand for specific knowledge. One of the main themes systematically addressed in the context of these participatory procedures had to do, of course, with territorial development, which became an essential element for a project's social acceptability. As a result, autoroute observatories were launched by the Ministry of Equipment, by its research unit on highway and autoroute techniques (SETRA), and by three autoroute concession companies: *Autoroutes Paris-Rhin-Rhône* (APRR), *Autoroutes du sud de la France* (ASF), and the *Société française du tunnel routier du Fréjus* (SFTRF).

The ten or so observatories that were launched met with varying success, as these evaluation mechanisms were not covered by any regulatory requirements. We shall mention here only the two main experiments, which benefited from the solid support of APRR. The observatory for the A71 autoroute (Bourges/Clermont-Ferrand) was in operation from 1986 to 1997, in partnership with the CERAMAC laboratory of the University of Clermont-Ferrand. The observatory for the A39 (Dole/Bourg-en-Bresse), operated between 1993 and 2004, in partnership with the ThéMA laboratory of the University of Franche-Comté.

The report from the A71 observatory identified three series of major facts, relating to territorial interaction (Varlet and Jamot, 2002):

- The A71 and the more southerly route via the A75 and the southeast, make the Massif Central a "potential transit space". The more recent arrival of the A9 has resulted in a major autoroute crossroads at Clermont-Ferrand, which has reinforced this transit function.
- The autoroute link has sparked a "temporary imbalance in local economic systems through the emergence of exogenous as well as endogenous dynamics". The external influences can be seen in the establishment of a hotel chain beside the A71, by the expansion of shopping centres, and by a tendency for businesses installed in the hinterland without links to transport networks to move closer to autoroute interchanges. There has also been a boom in the establishment of secondary residences.
- Growth in nodes located near to the autoroute has encouraged a degree of urban sprawl as businesses located in denser urban centres made the short move to areas closer to the interchanges, which were deemed more functional.

The observatory for the A39 started work at the outset of the construction phase in 1992, and continued operations until 2004, or six years after the last segment of the autoroute entered service. This was, then, the first opportunity for an up-close examination of the direct effects on the local economy and employment, and it produced some very specific results for the construction phase (procurement from local producers, employee spending and consumption habits, local taxes paid, etc.), on operation of the autoroute (jobs created by the concessionaire and subcontractors, local taxes, etc.), and on the reorganisation of traffic (the shift of long-distance traffic from the national route to the autoroute, increased traffic near the interchanges, etc.).

Two series of observations were conducted to pinpoint the indirect effects as well:

- One set of effects related to the autoroute itself, which gave rise to three types of developments: adaptations of the spaces physically impacted by the autoroute (regrouping of farmlands and modernisation of feeder roads from the interchanges into the surrounding territory), the economic and local development effects of accessibility and proximity to the autoroute (activity zones, preparation of development projects and inter-municipal co-operation), and promotion of the territory (installation of high-quality service areas).
- The other set concerned the use and usefulness of the infrastructure for the broader territory. As far as mobility practices and business logistics are concerned, the adaptations observed were modest.

The most significant results related to the concerns of the public authorities (and the autoroute companies), described earlier. In a public debate still marked by the NIMBY phenomenon, it is important to be able to identify the typical effects cited by those interviewed, especially in the case of the last observatory mentioned. Finally, we should note that the A39 observatory provided considerable input to the LOTI audit, supplying information that an ordinary ex-post balance sheet would never have been able to reconstruct.

We may note that, for autoroute investments, the permanent observatory methodology has been well tested, with accumulated experience that allows us to delimit thoroughly the object of the observation, whether this relates to the direct effects that the observatories have found to be important, or the indirect effects that are still highly dependent on the potential of the areas affected and on exogenous variables. The situation is clearly different in the case of investments in high-speed rail links, where the "gravitational pull" of the interchanges is not an issue.

Permanent observatories for the LGV

There is no inherent reason why the methods of ex-ante evaluation or ex-post audit cannot be applied to these investments as well. However, expanding the approach to include the socioeconomic effects of an LGV raises a problem that is quite different from that of the autoroute case.

Positioning the problem

The first difference relates to the mode of transport itself. When it comes to traffic and advantages for users, autoroute projects typically involve the redistribution of traffic from the old highway and, in contrast to the time savings offered by a new project, there is generally very little in the way of a modal shift. In the case of high-speed rail lines, modal transfers are naturally much more important, whether this involves a shift away from highway travel for short and medium distances or a shift away from air travel for longer distances.

In each case, the relative difficulty can be readily appreciated. In the first case, it is enough to monitor autoroute traffic (something that is easy to do as most autoroutes operate under the "closed toll" system) and then to supplement this with regular vehicle counts on alternative routes. This provides a sufficient time series for comparing forecast and actual traffic and for explaining the discrepancies. In the case of a new high-speed rail line, we must make a distinction between traffic diverted from highways and airlines and rail traffic induced by the improvement. This presupposes a degree of consistency among different statistical sources, and usually specific passenger surveys as well. It also assumes that the characteristics of the different modes available are captured in detail, including speed of travel, schedules and fares, in order to analyse forecasting errors.

The infrastructure itself and its related services have characteristics quite different from those of the autoroute, starting with the fundamental difference between a train station and a highway interchange. Although some suburban or rural stations have been built on the new lines, TGVs generally serve central stations, and a considerable portion of traffic is city-centre to city-centre. Autoroute interchanges, on the contrary, are generally located far from city centres.

As a result, the spatial diffusion is quite different. The autoroute interchange serves exclusively to ensure highway accessibility to city centres and rural areas, while the train station distributes traffic into urban areas via complementary modes: private automobile, foot traffic, public transit, taxi, or transfer by train to secondary stations.

We may conclude, then, that the spatial effects of an LGV will not be the same as those of an autoroute. At the risk of oversimplifying, we may summarise the differences between their respective “structuring effects” by noting that autoroutes produced primarily "transit effects" (*effets de traversé*: i.e. the impact on the areas through which they pass) while an LGV is likely to induce “pole” or "hub” effects (*effets de pôle*”: i.e. the impact on areas at each end of the line), which may lead to a rebalancing of urban hierarchies (Bonnafous, 1980).

All of these differences can be appreciated by presenting the case of the socioeconomic observatory for the LGV Sud-Europe-Atlantique.

The socioeconomic observatory for the LGV LISEA (Sud-Europe-Atlantic) - 2012-2027

The LGV from Tours to Bordeaux is now under construction and will come into service in 2017. Its 300 km of new line will complete the high-speed link from Paris to Bordeaux, cutting the travel time from three hours to two hours and five minutes. The new line has been let under concession by the Réseau Ferré de France to the LISEA company, a subsidiary of VINCI. The concession contract calls for serving the stations on the existing rail network via a system of connections between the LGV and the existing line: ten connections are planned, representing 40 km of supplementary rail segments. The main cities served by these connections will be Poitiers and Angoulême (see maps in Annex 2.2 and travel times table in Annex 2.3).

This is an innovative project in terms of financing, as it uses the concession system to limit public borrowing. The overall cost of EUR 7.8 billion will be financed as follows:

- EUR 3.8 billion from LISEA, of which 20% is in equity and the remainder borrowed
- EUR 1 billion from RFF, based on a loan backed by the expected additional revenues from expanding traffic on the existing network, and hence fare revenues
- EUR 3 billion in subsidies, half of which will come from the State and half from the local governments concerned.

It may be noted that this last item of financing involved negotiations between 5 regional councils, 19 departments and 33 *communautés de communes* or metropolitan areas. Each of these partners had an interest in the layout of the new line, either because of the TGV (high-speed train) services it would offer them or because of subsequent extensions of the line. This aspect is obviously important for the content and the geographic scope of the observatory, especially as the financial contributions were justified by the project's expected economic effects.

Given the scope of the project (presented as the biggest concession project in Europe), the prospect of a LOTI audit by 2020 or so, and the expectations of the local governments co-financing the project, the responsible authority (Réseau Ferré de France) introduced a clause in the concession contract obliging the concessionaire to establish and finance a socioeconomic observatory for the effects of the new line, known as OSE-LISEA. This contractual provision thereby ensures permanent financing for the observatory, which is to function for ten years after the line comes into service, i.e. until 2027 at least. This is a considerable methodological advance, for the socioeconomic effects of a structural nature will only appear with the passage of time. In nearly all the studies cited above, the authors have stressed that the observation time after entry into service has been too short.

Moreover, the fact that the observatory was put in place when the works had barely begun helped to prevent any loss of information on the "construction phase effects", as confirmed by the first results on this topic (Fouqueray, 2013). It also served to identify, early on, the potential questions and expectations of some of the key players (Manceau, 2012).

The objectives of the OSE-LISEA are spelled out in the concession contract. Thus, it is to "evaluate the direct and indirect effects of the LGV SEA on mobility, the local economy, and territorial development. The output of the Observatory must:

- be useful to local stakeholders for integrating the LGV into their territory to best advantage
- provide input to national stock-taking and debate on the effects of high-speed rail lines and be useful for forward planning."

The resulting organisation of the observatory involves units of the State as well as the local partners. It is managed by a technical and administrative team within LISEA, assisted by a scientific committee of academics to help in preparing and evaluating progress with the work programme. The State units participate in a monitoring committee that validates this programme, while the local partners (elected officials, local government units, chambers of commerce) are consulted by two regional commissions (North and South) set up to address the particular expectations of the regional geographic hubs, namely Aquitaine/Midi-Pyrénées and Poitou-Charentes/Centre.

Primary themes

The scientific committee proposed an initial work programme (Manceau, 2013) calling for observations and analyses based on six themes.

- **"Construction phase" effects.** Experience with autoroute projects (Bérion et al., 2007) has shown the usefulness of putting a mechanism in place promptly to monitor the local effects of the construction phase and to identify the geographic distribution of the effects on the production system. This is obviously an issue of pressing local political interest, as revealed at the public presentations of the initial results.
- **Transportation availability and traffic volumes.** This is obviously the most important theme for the detailed design and usefulness of future LOTI audits. The required information can be broken down into three categories:
  - **User-friendly time schedules**. Given the heavy competition from airlines on the longer routes, the schedules must be surveyed for both modes and for the train stations and airports involved. For rail service, the schedule also constitutes a demand for “track time” slots from RFF and LISEA, an important consideration for the financial profitability of the project.
  - **Fares**. While information on time schedules can be stored or even reconstituted, the situation is quite different with fares, as they are governed by "yield management" both for air and for TGV services. As we shall see in the next section, information on this point is typically short-lived.
  - **Passenger traffic**. Paradoxically, statistics on autoroute or air traffic are more accessible than those for rail traffic, where the historic operator insists on commercial confidentiality. Fortunately, a recent regulation will make most of these data available, although with a one-year lag. It will likely be necessary to conduct passenger surveys before and after entry into service.
- **"Station effects".** Numerous observations (Bazin et al., 2009, Richer et al., 2009) have found very significant urban developments in the immediate vicinity of the TGV stations. Urban developments should in fact be monitored closely, starting several years before entry into service, given the expectations of stakeholders.

- **Metropolitan and territorial dynamics.** This theme arises whenever an LGV makes significant changes to travel distances and times between urban areas. An ex-ante/ex-post study of the very first TGV project in France produced important results on this theme (Buisson et al., 1986). Spatial relationships are necessarily part of an urban hierarchy that involves a regional capital, Bordeaux, of recognised dynamism, but also Paris, the economic clout of which is based on centuries of centralisation. It is important, then, to observe as closely as possible the transformations in this hierarchy, which obviously involves the medium-sized cities.
- **Effects on tourism.** There is a wealth of statistical information on this theme in France, relating both to accommodation capacities and visitor numbers. The OSE-LISEA will however be limited to areas likely to be strongly affected by differences in supply. The work programme will therefore focus in particular on urban tourism.
- **Strategies of stakeholders and organisations.** The first contacts with local stakeholders who spontaneously declared their interest in the observatory's work programme revealed a very optimistic outlook. This sometimes seems to be based on an overestimate of the new availability of transportation, but it may also lead to real initiatives that should be inventoried and analysed, for the "branding effect" of TGV service has been highlighted in nearly all the studies.

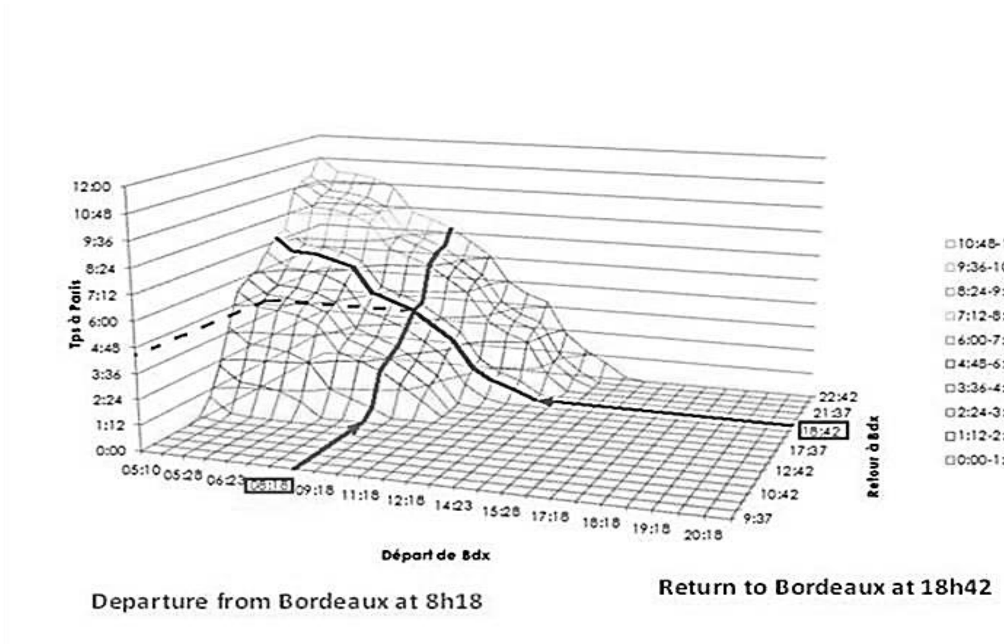
The main methodological challenges

This work programme has been introduced progressively and has already revealed some methodological problems, some of which are inherent to the very notion of a permanent observatory. The most important issues relate to two principal difficulties. One is the need to synthesise complex information into a limited number of indicators, the relevance of which can pose a problem. The other has to do with the fleeting nature of information, which can disappear if it is not compiled promptly. We shall confine ourselves here to two illustrations of these problems, both of which relate to supply monitoring.

With regards to the relevance of the indicators, the example selected concerns the problem of delivering a synthetic description of time schedule changes. As things stand now, the local political authorities seem to attach great importance to the number of daily train runs. However, the real question is not whether there will be at least six TGVs per day in each direction, as one mayor demanded during the negotiations on public financing, but rather whether the schedules and their related services will make it easier to get to certain destinations, starting with Paris.

Work was therefore undertaken to compile a composite body of information, by its nature complex, representing hundreds of departure and arrival times and to make sense of that information. To illustrate these initial efforts (Joho, 2013) we take the notion of "available time at destination" (ATD), which is determined on the basis of the proposed schedules, but under the constraint of a departure time ("not before *X* o'clock") and return time ("not after *X* o'clock"). Figure 2.1 below illustrates this concept using the example of a one-day round trip between Bordeaux and Paris.

Figure 2.1. **Available time at destination (Paris), with a departure constraint (after 8 a.m.) and a return constraint (before 7 p.m.)**



Source: Paul Joho, 2013.

On this three-dimensional graph, the outward-bound train's departure time is shown on axis 1 and the arrival time of the return train is shown on axis 2. The length of time available at destination can then be determined from the vertical axis (axis 3). This value corresponds to the co-ordinate of a point that is at the intersection of the curve representing the consumption of time on the outbound train and the curve representing the consumption of time on the return train.

The advantage of these representations has to do with the synthetic nature of the indicator obtained. In a single number we can represent a duration of useful time, which is obviously fundamental for organising the programme of activities at destination and which certainly constitutes a pertinent indicator of the need to change the schedule. For example, the value of this indicator, as shown in Figure 2.1, is 4 hours and 48 minutes with the current Bordeaux-Paris schedules for a traveller who does not wish to leave before 8 a.m. and who wants to return before 7 p.m. This useful time would increase to 6 hours and 38 minutes at the running speeds authorised for the LGV, and with the same departure and arrival times.

This type of indicator offers a good explanation of certain observations concerning traveller demand following the introduction of a high-speed train service. In the first French case, that of the Paris-Lyon LGV introduced in the 1980s, traffic exceeded forecast demand, and the principal explanation had to do with the fact that the travel time, which was reduced from nearly four hours to two hours, and the high frequency of service allowed for many more same-day roundtrips, and even half-day trips, in place of longer stays and overnights at destination. This confirmed that there was a strong social demand, in particular from business travellers (Buisson et al., 2006) for this ATD.

Given the interest already declared by local stakeholders in this observatory and in monitoring the offer of transportation, it remains for them to express the types of ATD that should be promoted. It is clear that for each city the maximum ATD in Paris is of interest (as can be seen from Figure 2.1, it is ten hours and 46 minutes with the current train schedules from Bordeaux, if one takes the first train in the morning and returns on the last at night). An ATD with more convenient schedules, such as those we have imagined above, should certainly be posted as well as an ATD for a half-day round trip.

Of course, the observatory will have to establish these indicators for all LGV stations and for destinations other than Paris, which are yet to be determined, as well as for stations which are not on the high-speed line but which benefit from the TGV services using that line (for example La Rochelle, Toulouse). Where airline service exists, the schedules should be the subject of comparable treatment and posting. This assumes that it is possible to evaluate transportation time to the terminals, as well as the advance check-in or arrival times that passengers must observe for each mode.

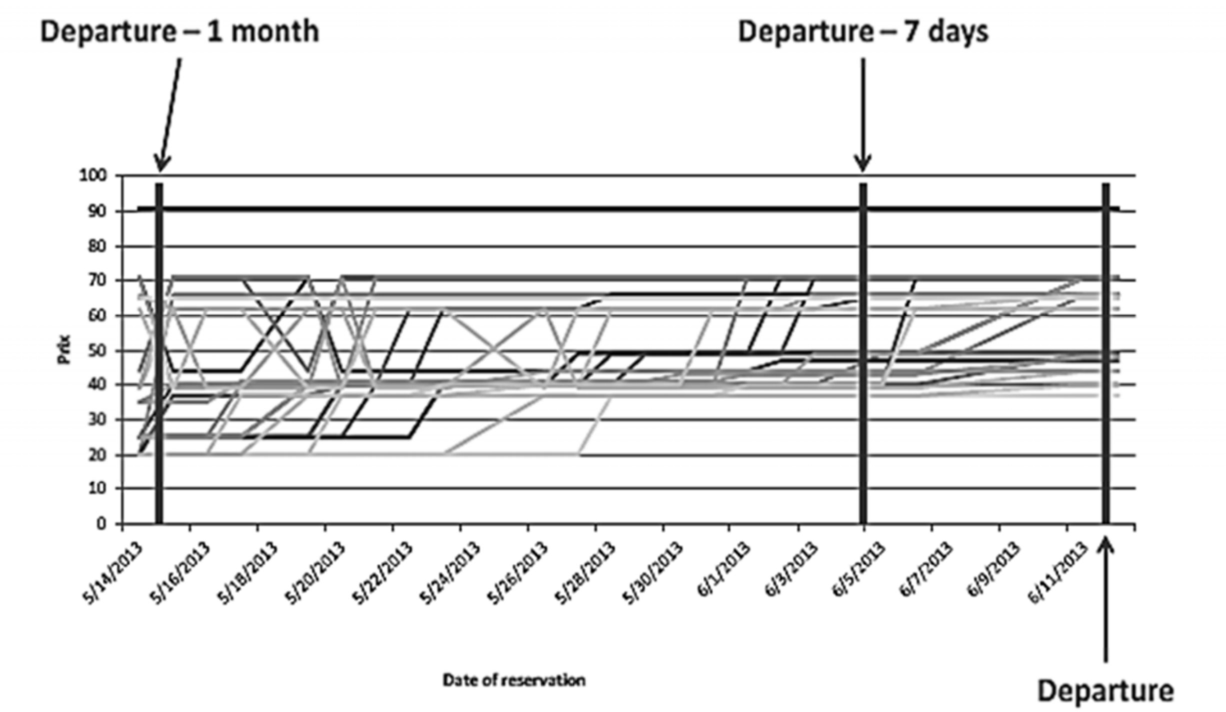
With respect to the short life of information – the other main methodological challenge – the best illustration is no doubt that concerning fares. The fact that airline and TGV operators practice "yield management" makes the problem especially complex. Information is available on the minimum and maximum values of these fares for a second-class TGV ticket, because SNCF is legally obligated in its operating specifications to publish it. A recent version of the official text added the following sentences: "SNCF shall publish and communicate clearly and fully the lowest fare, apart from promotional fares, and the highest fare for a second-class ticket applicable for each route. This fare information shall be made readily accessible to users of the route." Airlines are not under the same constraint.

However, a close examination of actual rail fares suggests that this obligation to publish is a weak constraint and not very helpful for a real understanding of fares. Figure 2.2 below, taken from studies by Paul Joho (2013), provides persuasive evidence on this point. It shows that, for a given departure date on the Bordeaux-Paris trains, the minimum price rises almost daily according to the time elapsed since the reservation was made.

We can distinguish the particular case of a single train for which the minimum fare is constant and fixed at EUR 90. This is a train leaving Bordeaux at 6:23 a.m. (which can also be identified in Figure 2.1) for which "yield management" does not apply, probably because it must run as nearly full as possible. For other trades, there is both a sharp chronological series variance (depending on time since reservation) and a sharp cross-series variance (depending on the train).

The problem of short-lived information is a glaring one: for a departure on 11 June, as shown in Figure 2.2, the fares posted at the ticket reservation sites on 22 May are no longer available on 23 May, and there is no chance that the carrier will be willing to offer them, unless forced to do so by competition from airlines or, within a few years, from other rail operators. The only solution, then, is to conduct a systematic and comparative survey of reservation sites.

Figure 2.2. **Minimum fare for all Bordeaux-Paris trains, based on length of time elapsed since reservation to departure date**



Source: Paul Joho, 2013.

A systematic survey will require selecting trains in such a way as to replicate the diversity of origins-destinations as well as of travel days (weekdays, weekends, and major vacation periods). It also means distinguishing the profiles of buyers and the type of ticket they choose (second-class ticket with a senior card, exchangeable second-class ticket, first-class ticket etc.). Lastly, for a given departure date, different advance reservation times must be set (for example, three months, one month, seven days and previous day). A similar exercise will have to be undertaken for airline connections.

A thorough survey of this kind will require compiling and working with thousands of fares, and consequently thought should be given to using synthetic indicators. When the LOTI audits mention rail fare increases or airfare cuts, they give no indication of the orders of magnitude. The audits should be improved in this regard, not only to provide useful information to stakeholders but also to be able to explain eventual errors in traffic or revenue forecasts. The observatory is now examining the question of synthetic indicators for fares.

Conclusion: Addressing the challenges of complex and short-lived data

Between the launch of the SPOT discussed in the section "From ex-post evaluations to the notion of a permanent observatory", the first permanent observatory of a new infrastructure project introduced at the end of the 1970s, and the OSE-LISEA, established just two years ago, more than three decades have elapsed and yet there has been an obvious methodological continuity. In fact, for all these French experiments, the challenge has been to find a dual response: on one hand, to address the volume of information gathered (remembering that information overload is lethal), and on the other hand to address the well-known erosion and even disappearance of information. One may be tempted to gather ever more information, recognising that it is likely to disappear soon, and this makes for a massive and unwieldy database, as illustrated in the last example cited here.

However, besides its longevity, the socioeconomic observatory for the LGV South Europe-Atlantic has the particular advantage of offering answers to these problems. The involvement of multiple local partners can help to appreciate the pertinence of synthetic indicators and thus assist in choosing those that will be useful for enlightening public debate. This choice must be made with knowledge of the uses to which the data will be put, particularly for the most technical aspects – for example, analysing the quality of forecasts on which the investment decision is based.

The primary goal of the ex-post evaluation will always be to help prepare better ex-ante evaluations

.

Annex 2.1

List of "Bilans LOTI" [LOTI audits] – Accessible in PDF format at: http://www.cgedd.developpement-durable.gouv.fr/les-bilans-loti-r245.html

21 August 2013

- **LGV Est européenne (phase 1)**
  Avis [Opinion] no_009145-01– July 2013
- **Construction of Autoroute A 28 Alençon-Tours**
  Avis n° 008517-01 – February 2013
- **Aeronautical infrastructure upgrade at Paris-Charles de Gaulle Airport**
  Avis n° 008039-01 – July 2012
- **Electrification of the Rennes–Saint-Malo line**
  Avis n° 007847-01 – July 2011
- **Autoroute A83 Nantes–Niort**
  Avis n° 007538-01 – June 2011
- **Autoroute A 20 (Section Brive Montauban)**
  Avis n°007284-01 – April 2011
- **Operation of tramway T2**
  Avis n°007715-01 – May 2011
- **Autoroute A75 Engayresque-La Cavalerie south section, including the Millau viaduct**
  Avis n° 007191-01– February 2011
- **Autoroute A66 (bifurcation A61 – Pamiers)**
  Avis n° 007283-01 – November 2010
- **Re-routing of the RN 12 to Jouars–Pontchartrain**
  Avis n° 007447-01 – November 2010
- **A16: concessioned section: L'Isle-Adam–Amiens–Boulogne and A16: non-concessioned section: Boulogne–Belgian border**
  Avis n° 007087-01 – July 2010
- **A75 Clermont-Ferrand–Sévérac-le-Château**
  Avis n° 006652-01 – April 2010
- **Electrification of the Paris–Clermont-Ferrand rail line**
  Rapport n° 007156-01 – January 2010

- **Autoroute A26 Châlons-en-Champagne–Troyes**
  Avis n° 003820-02 – December 2009

- **Autoroute A51 Sisteron-La Saulce**
  Avis n° 006654-01– November 2009

- **Autoroute A29, Pont de Normandie-A13 and Le Havre-Saint-Saens sections and the Pont de Normandie**
  Avis n° 006600-01 – October 2009

- **Electrification of the Paris–Caen–Cherbourg rail line and upgrade of the Paris-Granville rail line**
  Avis CGEDD – n° 006948-01 et n° 006949-01 – September 2009

- **Autoroute A5 La Francilienne – Troyes**
  Avis CGEDD n°005616-01– January 2009

- **Electrification of rail lines in Bretagne**
  Avis CGEDD n° 005921-01– July 2008

- **Autoroute A 39 – Sections Dijon–Dôle and Dôle–Bourg-en-Bresse**
  Avis CGPC n° 005322-01 – July 2008

- **Autoroute A 19 – Section Sens–Courtenay**
  Avis CGPC n° 005138-01– July 2008

- **LGV Rhône-Alpes et Méditerranée**
  Rapport [Report] CGEDD n° 005448-01
  Avis délibéré du CGEDD n°005448-01 – July 2008

- **CGPC opinion on the bilan LOTI of Autoroute A 77 – Section Dordives–Cosne-sur-Loire**
  Avis CGPC n°005781-01 – June 2008

- **CGPC opinion on the *bilan* LOTI of train speed control by tracking antenna (KVB)**
  Avis CGPC n° 005721-01 – March 2008

- **Autoroute A54 (Saint-Martin-de-Crau–Salon-de-Provence)**
  n° 005295-01 – December 2007

- **Puymorens Tunnel**
  Avis CGPC n°005172-01 – December 2007

- **RER stations at Stade de France**
  Bilan LOTI RER Stade de France – RFF – November 2006
  Avis CGPC n° 005438-01– August 2007

- **Autoroute A837 (Saintes-Rochefort)**
  Concessionnaire ASF
  n° 5173-01 / Avis du CGPC : A837 – July 2007

- **RER D (Gare de Lyon-Châtelet connection), RER E (Eole) and metro line 14 (Météor)**
  Bilan LOTI interconnection du RER D Châtelet-gare de Lyon – RFF March 2006
  Bilan LOTI RER E – RFF – mars 2006
  Avis CGPC n°004956-01 – June 2007

- **Non-concessioned Autoroute A28 (Rouen-Abbeville)**
  DRE Haute-Normandie et Picardie
  Le rapport relatif à l'A28
  n° 004891-01 / L'avis du CGPC : A28 – February 2007

- **Non-concessioned Autoroute A20 (Vierzon-Brive)**
  Direction Régionale de l'Équipement du Limousin
  Le rapport et ses annexes (1,2,3)
  n°004812-01 / L'avis du CGPC : A20 – February 2007

- **RN 24**
  Direction Régionale de l'Équipement de la Bretagne
  n° 004813-01 / L'avis du CGPC : RN24 – January 2007

- **LGV Nord Europe et Ile-de-France Interconnection**
  Bilan LOTI LGV Nord – RFF – May 2005
  Bilan LOTI Interconnexion Ile-de-France – RFF – September 2005
  Avis CGPC n°004624-01 – July 2006

- **Autoroute A14 (Orgeval-Nanterre)**
  Concessionnaire SAPN
  L'avis du CGPC : A14 – November 2005
  La brochure de présentation de SAPN

- **Autoroute A57 (Cuers/Le-Cannet-des-Maures)**
  Concessionnaire ESCOTA
  n° 2004-0263-01 / L'avis du CGPC: A57 – December 2004

- **Tramway line between Saint-Denis and Bobigny (93)**
  Avis CGPC n°2002-0140-01 – November 2003

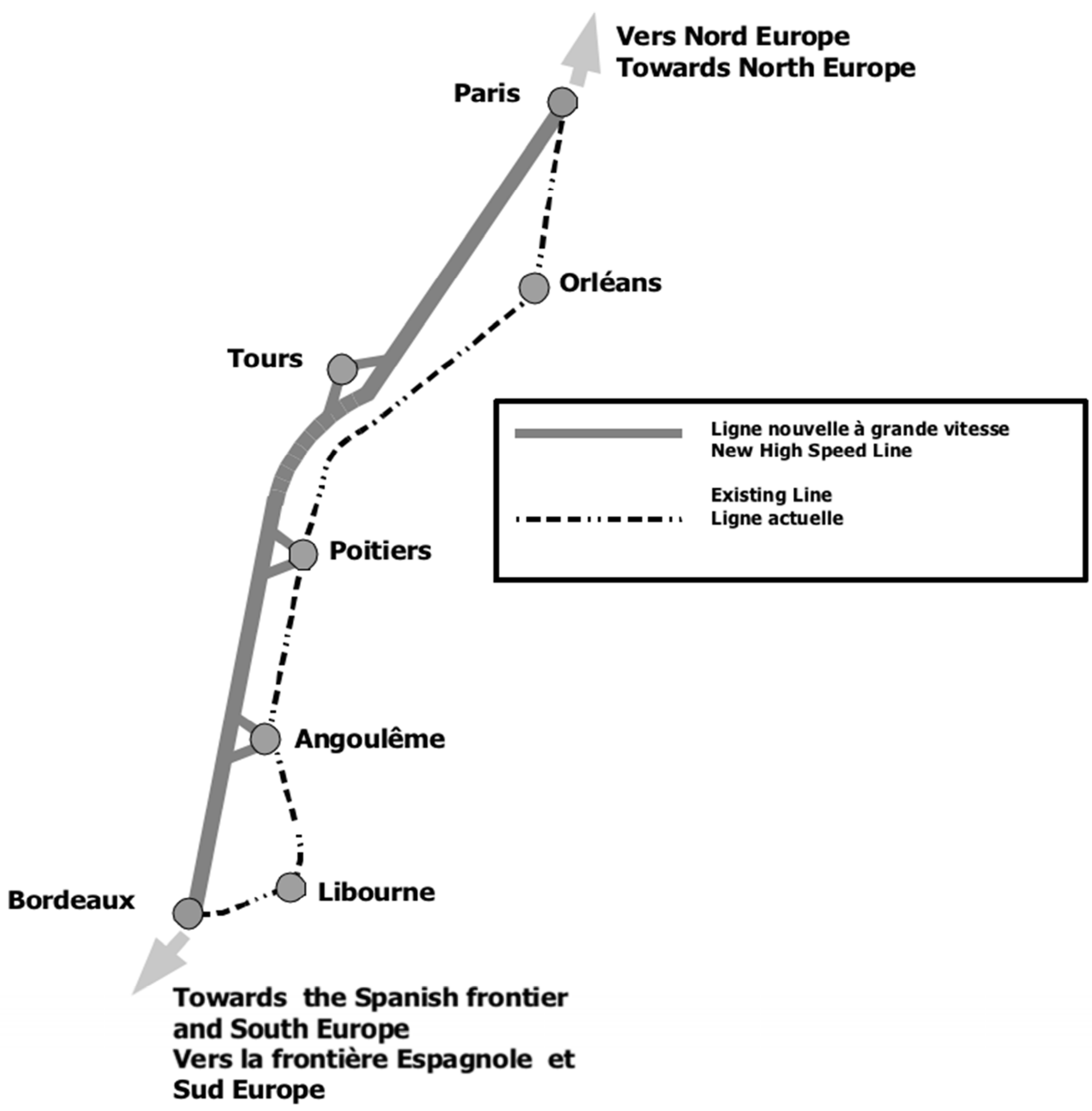
- **TGV Atlantique**
  Avis CGPC n°1999-0163-01 – July 2001

- **Upgrade of the Montmélian-Albertville-Moûtiers link**
  Concessionnaire AREA
  L'avis du CGPC : A43 – November 1999

- **Autoroute A49 (Grenoble-Valence)**
  Concessionnaire AREA
  L'avis du CGPC : A49 – November 1999

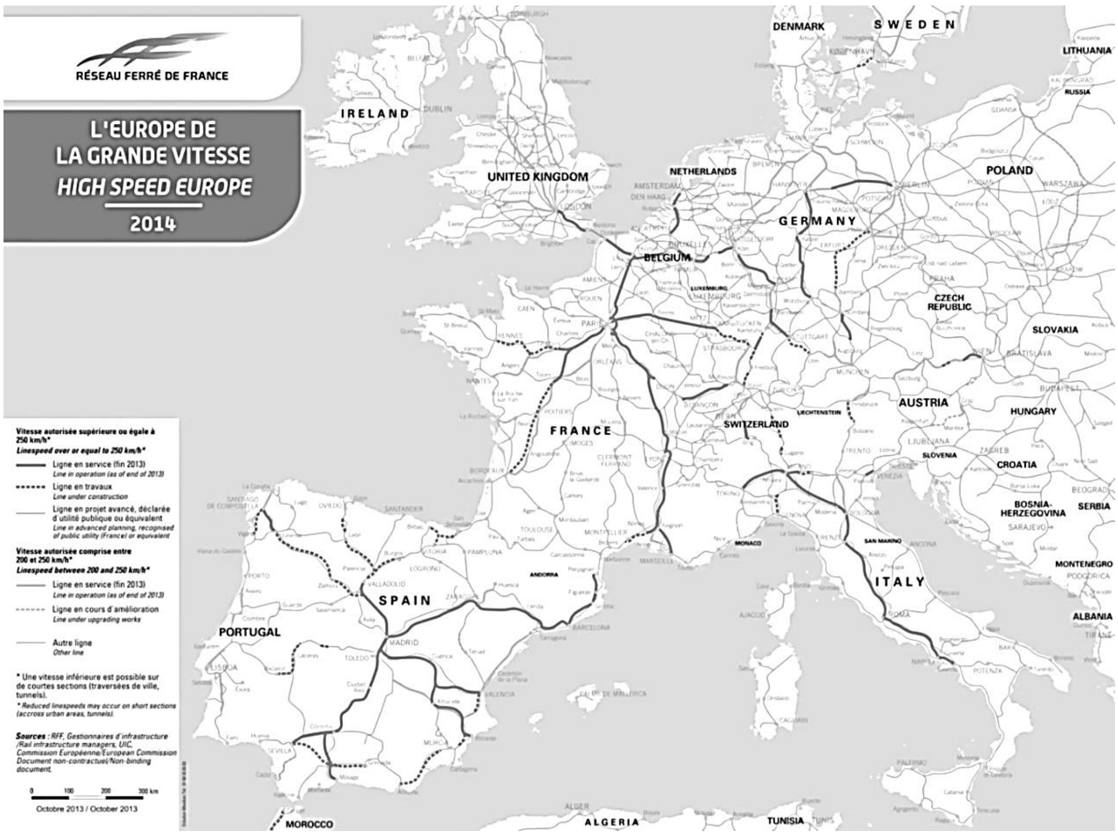
Annex 2.2

Figure A2.2.1. **Outline of high-speed rail project, South-Europe-Atlantic (SEA)**



Source: Abeneal-commonswiki, 2006.

Figure A2.2.2. **High-speed Europe in 2014**



Source: Réseau Ferré de France, 2014.

Annex 2.3

Table A2.3.1. **Travel times: Table comparing current line / LGV SEA line**

Routes	2007	2017
Paris-Poitiers	1h26	1h17
Paris-Angoulême	2h05	1h40
Paris-Bordeaux	3h00	2h05
Paris-La Rochelle	2h50	2h27
Bordeaux-Tours	2h30	1h30
Bordeaux-Angoulême	0h52	0h35
Poitiers-Bordeaux	1h32	0h55
Poitiers-Angoulême	0h44	0h37
Poitiers-Tours	0h47	0h30
Tours-Angoulême	1h32	1h07
Paris-Toulouse	4h56	4h03

Note: Average indicative travel times

References

Bazin, S., C. Beckerich and M. Delaplace (2009), "Desserte TGV et localisation des entreprises dans les quartiers d'affaire: nouvelle accessibilité ou nouvelle offre de bureau? Le cas de la gare centre de Reims", *Les Cahiers scientifiques du transport,* No. 56.

Bérion, P., G. Joignaux and J.F. Langumier (2007), "L'évaluation socio-économique des infrastructures de transport: Enrichir les approches du développement territorial", *Revue d'Économie Régionale et Urbaine*, No. 4, pp. 651-676. (http://dx.doi.org/10.3917/reru.074.0651)

Bérion, P. (1995), "L'Observatoire de l'autoroute A39: démarche utilisée et présentation de l'état initial", *Autoroutes, économie et territoires*, CERAMAC, Clermont-Ferrand, 1998, pp. 123-137.

Bonnafous, A. (1980), "Rhône-Alpes, capitale Paris: les effets prévisibles du TGV", *Revue de géographie de Lyon*, No. 3. (http://www.youscribe.com/catalogue/presse-et-revues/savoirs/sciences-humaines-et-sociales/rhone-alpes-capitale-paris-les-effets-previsibles-du-t-g-v-1082747)

Bonnafous, A. (1979), "Underdeveloped regions and structural aspects of transport infrastructure", in: W.A. Blonk (Ed.), *Transport and Regional Development*, Saxon House, Farnborough, pp. 45-62.

Buisson, M.A., A. Bonnafous, J.L. Bernadet, C. Jafflin and D. Mignot (1986), "Effets indirects du TGV et transformations du tertiaire supérieur en Rhône-Alpes", *Études et recherches*, Laboratoire d'Économie des Transports. http://temis.documentation.developpement-durable.gouv.fr/docs/Temis/0002/Temis-0002152/5928.pdf

ECMT (1975), *Impact of Infrastructural Investment on Industrial Development: Report of the Twenty-Fifth Round Table on Transport Economics Held in Paris on 14th and 15 March 1974*, ECMT Round Tables, No. 25, OECD Publishing, Paris.
DOI: http://dx.doi.org/10.1787/9789282104057-en

ECMT (1969), The Impact of Infrastructural Investment on Economic Development: Report of the Fourth Round Table on Transport Economics Held in Paris on 23-24 October 1969, ECMT Round Tables, No. 4, OECD Publishing, Paris.
DOI: http://dx.doi.org/10.1787/9789282104798-en

Fouqueray, E. (2013), "Sous-traitance et fourniture du chantier de la LGV SEA Tours-Bordeaux : quelles retombées pour les territoires?", Observatoire Socio-économique LISEA. (http://www.lgv-sea-tours-bordeaux.fr/engagements-durables/observatoire-socio-economique-lisea)

Gérardin, B., J.-M. Cusset, A. Bonnafous and J. Valère (1981), *Système permanent d'observation du triangle Lyon-Chambéry-Grenoble* (SPOT), pp. 237 (http://www.let.fr/).

Joho, P. (2013), "Suivi de l'évolution de l'offre de transport liée à l'arrivée de la LGV SEA Tours-Bordeaux dans le cadre de l'Observatoire socio-économique mis en place par LISEA, Observatoire Socio-économique", LISEA. (http://www.lgv-sea-tours-bordeaux.fr/engagements-durables/observatoire-socio-economique-lisea)

Manceau, E. (2013), "Programme de travail de l'Observatoire socio-économique 2013-2020", Observatoire Socio-économique LISEA. (http://www.lgv-sea-tours-bordeaux.fr/engagements-durables/observatoire-socio-economique-lisea)

Manceau, E. (2012), "La ligne grande vitesse Tours-Bordeaux et ses effets vus par les acteurs du territoire", Observatoire Socio-économique LISEA. (http://www.lgv-sea-tours-bordeaux.fr/engagements-durables/observatoire-socio-economique-lisea)

Ministère de l'écologie du développement durable et de l'énergie (2014), Bilans LOTI. http://www.developpement-durable.gouv.fr/Bilans-LOTI.html

Plassard, F. (1977), "Les autoroutes et le développement régional", *Economica*, Paris, pp. 342

Richer, C., P. Bérion and V. Facchinetti-Mannone (2009), "L'observatoire des effets territoriaux des gares du TGV Rhin-Rhône : contexte, enjeux et perspectives", *Images de Franche-Comté*, (40), pp. 2-5.

UIC (2011), *High Speed Rail as a Tool for Regional Development*, In-depth Study, (http://www.shop-etf.com/fr/high-speed-rail-as-a-tool-for-regional-development.html)

Varlet, J. and C. Jamot (2002), "Autoroute A71, acteurs et territoires. Bilan d'une décennie d'observations", *Géocarrefour*, Vol. 77, No. 1. (http://www.persee.fr/web/revues/home/prescript/article/geoca_1627-4873_2002_num_77_1_6259).

Chapter 3
The National Audit Office's value-for-money assessment of transport investments

Geraldine Barker, Grace Beardsley and Annie Parsons - The National Audit Office, London, United Kingdom.

The United Kingdom National Audit Office (NAO) scrutinises public spending on behalf of Parliament, helping it to hold government departments to account and helping public bodies improve performance and delivery. We publish around 60 value-for-money studies each year which look at how government projects, programmes and initiatives have been implemented and make recommendations on how services can be improved.

This chapter will focus on the work of the Transport value-for-money team in the NAO, examining our role; how we carry out assessments of the Department for Transport, its major projects and areas of expenditure, including examples of techniques used to assess value for money; and exploring the difficulties in concluding whether a project or programme is value for money.

Summary

The United Kingdom National Audit Office (NAO) scrutinises public spending on behalf of Parliament, helping it to hold government departments to account and helping public bodies improve performance and delivery. We publish around sixty value-for-money studies each year across a range of government activities, of which around three of these usually cover transport topics. Our reports look at how government projects, programmes and initiatives have been implemented and make recommendations on how they can be improved.

Our value-for-money work is not strictly ex-post assessment in the usual sense of assessing a programme once it has been in operation for some time. Due to the length of time needed to complete major transport investments and our remit to focus on accountability, we often carry out an assessment of a project before its completion. In some cases, particularly for significant infrastructure investments, a series of value-for-money reports is appropriate as the programme will develop over time. These tend to focus on how the programme is being delivered, in terms of the planning, procurement or construction phases of infrastructure projects.

We have had cause to comment on the Department for Transport's (the Department) failure to carry out full ex-post assessment of the United Kingdom transport projects that we have examined. Ex-post assessments have a valuable role to play in capturing learning and feeding into decision making about current and future projects. Given the long lead times such projects and programmes often have before they deliver their intended benefits, ex-post assessments should be one part of a wider programme of reviews throughout a project or programme's life.

In carrying out our work our three main aims are to:

- Obtain robust evidence and analysis in order to draw sound conclusions about whether the investment of public money provided value for money.
- Address the issues which taxpayers and their elected representatives believe are pertinent in making that assessment.
- Draw out lessons for future programmes both within the transport sector and elsewhere in government.

The first and third of these – the need for robust evidence and the desire to learn for the future – clearly have strong parallels in "classic" ex-post assessment.

We base our work on a standard approach that we apply across the range of central government's activities and services. All our value-for-money work refers to an analytical framework shown in Annex 3.2 of this chapter. Audit teams use this framework as a starting point and the basis on which to develop their detailed methodologies. They flex the application of the framework according to the particular topic, applying their professional judgment and experience. The analytical framework examines the economy, efficiency and effectiveness with which resources are used. We also consider factors such as whether the use of resources was optimal and key decisions were reasonable at the time they were taken.

This chapter focuses on four of the most significant and most recurring challenges we encounter when carrying out value-for-money work in the transport sector:

- **Choosing the appropriate point or points at which to assess a programme.** We need to balance a number of considerations, such as allowing the programme

sufficient time to become established, the expectations of our stakeholders for timely evaluation and the scope for our recommendations to influence the programme going forward.

- **Evaluating the wider economic impacts of transport investment.** Where wider economic impacts are cited as the reason why a transport project is required, decision makers and evaluators have found it hard to quantify these impacts. However, we have seen increasing interest recently in carrying out full ex-post assessments, including effects such as economic growth and regeneration, and using the learning from these to inform approaches to planning and delivering future programmes. For example, the Department for Transport has also published an evaluation strategy and an evaluation and monitoring programme.
- **Obtaining and interrogating data and information.** While our statutory rights give us unique access to data and documentary evidence, we too face issues of having to make judgments on a programme's success where there is insufficient data or information. We provide an example of how we have used available data to test and validate decisions made.
- **Assessing the impact of government interventions where there are many other factors at play.** For example, transport regulatory bodies have some influence on the rate of road accidents through their work to enforce vehicle safety standards but road safety is also influenced by, for example, the weather and drivers' health.

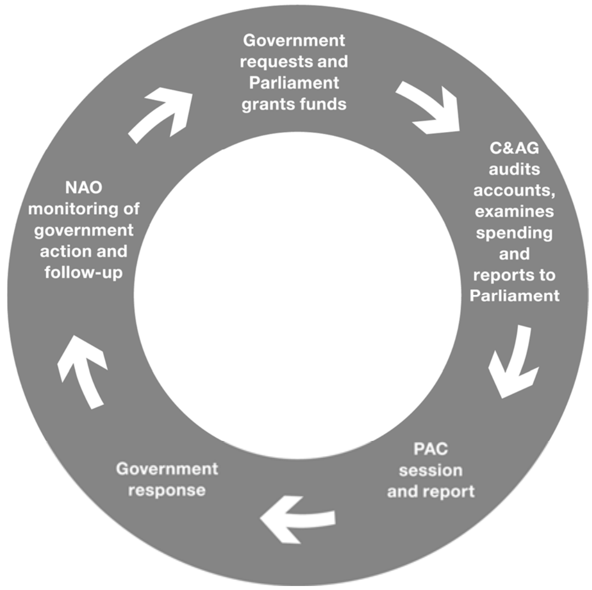
While most of our discussion focuses on the examination of major infrastructure programmes, we also refer to examples of other interventions by government agencies in the transport sector. In assessing such activities we have used techniques such as benefit-cost calculations to conclude on whether agency interventions are value for money.

Our value-for-money reports contain recommendations for the audited body; however, the extent to which they have been implemented is variable. Certain recommendations often recur, such as the robustness of underlying data and quality of risk management. We seek to address this by taking a systematic approach to following up our recommendations, with the aim that the Department and its agencies will increase their focus on these areas and improve performance over time.

The role of the National Audit Office

The work of the National Audit Office (NAO) is part of the United Kingdom accountability process for funds granted by Parliament (Figure 3.1).

Figure 3.1. **The accountability process**



Source: National Audit Office.

The NAO is headed by the Comptroller and Auditor General (C&AG) who has the powers to undertake financial audits of government accounts and value-for-money audits. We produce around sixty value-for-money reports each year, across the range of government's activities. Each report will:

- Consider the economy, efficiency and effectiveness with which government departments and other public sector bodies have used their resources.
- Form a judgment on whether value for money has been achieved.
- Make recommendations on how it could be improved.

Most value-for-money reports are presented to the Committee of Public Accounts (the Committee) in Parliament. The Committee uses our report as the basis of a hearing at which the Committee will question the government officials responsible about our findings. It subsequently publishes its own reports, including recommendations, to which the Government must respond.

Each year, around three NAO reports are published on transport topics (see Annex 3.1). Where transport projects or activities are directly funded by the Department for Transport (the Department), we have unique access to conduct value-for-money audits. We also examine projects which the Department delivers via local government or other partners in the private and third sectors. We select projects and areas of expenditure for audit on the basis of various criteria, including their financial value, strategic importance, level of risk and topicality. Our work in recent years has included:

- **Major investments in rail infrastructure**, including the construction and sale of the high-speed rail line linking London and the Channel Tunnel ("High Speed 1"), and the setting up of the programme to build a high-speed rail line linking London and northern England ("High Speed 2").
- **Major procurements** such as the purchase of new trains for intercity services and the "Thameslink" service across London.

- **Operational activities** of the Department and its agencies, for example, the inspection of heavy goods vehicles by the Vehicle and Operator Services Agency.
- **Interventions by the Department**, for example, its decision in 2009 to terminate the franchise for the East Coast Mainline rail service, held by a private company and subsequently transferred to be run by a public operator.
- **Delivering transport improvements at the local level**, for example, the maintenance of roads by both the Highways Agency and local authorities.

Our reports are not strictly ex-post assessments of the Department's programmes, in the sense of assessing a programme once it is fully in operation, as defined by Campbell and Rozsnyai (2002). We tend to examine major investments before they have been completed and their benefits fully realised, focusing on the planning, procurement and construction phases. This is due in part to the interests of our audience – Parliament and the taxpayer – in holding departments to account for the way they use public money, at a sufficiently early stage to influence the remainder of the programme, and to our aim to help improve performance and service delivery. It is also due to the scale and duration of the programmes that we examine, which take many years to come into operation. Of our recent work, our third report on the Channel Tunnel Rail Link, now known as "High Speed 1", published in March 2012, most resembles a classic ex-post assessment of transport investment. It involves a reworking of the original business case using available data. We also commented specifically on whether the project had been completed to time and cost and on whether forecasts of passenger demand had been achieved.

Within our reports, when relevant, we have commented on the extent to which the Department has evaluated its transport programmes. In general, we have found a lack of ex-post assessment. For example, in our review of the Department's funding of local authority major capital schemes, we found that the Department did not enforce requirements for local authorities to evaluate their schemes, and it had received evaluation reports for only two of seven projects which required them. More recently, the Department has taken positive steps. In March 2013, the Department produced a monitoring and evaluation strategy, followed by a document in October 2013 setting out its monitoring and evaluation program, which it will update annually.

Our approach to assessing value for money

Key to our approach is the requirement to conclude on value for money. Our analytical framework (described below and reproduced in Annex 3.2) helps us to define value for money and how we will assess it, at the beginning of a piece of work. This will lead ultimately to the value-for-money conclusion in the published report.

The analytical framework has three key components:

- Establishing exactly what we will examine[1]
- Identifying what good performance would look like, taking into consideration the environment at the time, which could constrain what is achievable.
- Assessing actual performance against "what good looks like" to draw conclusions and identify recommendations.

The framework draws on traditional approaches to assessing value for money, examining the:

- **Economy** with which resources are used.

- **Efficiency** of the relationship between the output of goods, services or other results and the resources used to produce them.
- **Effectiveness**, i.e. the relationship between the intended results and the actual results of the projects, programs and services.

The value-for-money conclusion, included in all of our value-for-money reports, is the C&AG's view on how well resources have been used to achieve particular objectives. It is usually founded on some type of comparative assessment such as cost-benefit analysis, where we evaluate the net benefits of a programme against a counterfactual, performance benchmarking against appropriate comparator programmes, or comparison to a model of good practice. There are times when there is insufficient information to conclude on value for money. In these cases, we may decide how far the audit body is at fault, or whether data limitations are a feature of the environment in which the body operates.

The challenges

There are a number of specific issues related to transport investments and projects that present us with challenges when conducting our value-for-money assessments and could also present problems for ex-post assessments. We discuss these further below, with reference to case examples taken from our reports. The issues and case examples we will cover include:

- Determining when is the right time to assess the value for money of a programme:
  - Department for Transport: The completion and sale of High Speed 1
  - Department for Transport: High Speed 2
- Evaluating whether wider economic impacts have been achieved:
  - Department for Transport: The completion and sale of High Speed 1
- Making an assessment when there is insufficient data or documentary evidence:
  - Department for Transport: The InterCity East Coast Passenger Rail Franchise
- Assessing the impact of an intervention when there are other factors and agencies at play:
  - vehicle and Operator Services Agency: Enforcement of regulations on commercial vehicles

It takes time to construct major transport projects and bring them into service. For example, the Crossrail service, which is currently being constructed east-west across London, will open fully in 2019, 10 years after construction began in 2009. Phase 1 of High Speed 2, a new rail link from London to northern England, is expected to be operational in 2026, some 14 years after the department's decision to go ahead with the project. When considerable sums of public money are spent there is a natural desire to establish whether that expenditure is worthwhile. Crossrail will cost GBP 14.8 billion, while High Speed 2 is expected to cost around GBP 42.6 billion[1]. Our key stakeholders are eager for information to inform other projects and are not inclined to wait for a full ex-post assessment (which might be many years hence).

Our challenge is to establish at what point an NAO assessment will be most valuable. We are now taking the approach of examining the Department's most significant programmes at key stages during their initiation, development and delivery. This enables the Department to address the risks and issues we identify at an earlier stage. We talk more about this in the final section of this chapter. For example, we first examined High Speed 1 in 2001, focusing on the financing of the project. We returned to the topic in 2005, reviewing progress in construction and revised expectations for public funding and in 2012 to report on the completion of the programme and the sale of the operating company, High Speed One Limited.

We are adopting a similar approach to High Speed 2, a programme to develop a new high-speed rail network between London, the West Midlands, Manchester and Leeds. The Department for Transport took the decision to develop Phase 1 of the scheme in January 2012 and the line is expected to be operational in 2026. We published our first report in May 2013, on the Department for Transport's progress in putting in place the foundations for successful programme delivery, examining specifically:

- the Department's case for building a high-speed railway
- the Department's cost estimate for Phase 1 of the programme, between London and the West Midlands and its assessment of the programme's affordability
- how the Department has set up the programme.

We evaluated the Department's performance against the key elements of success that we had identified in our *Guide to Initiating Successful Projects*, which is based on our experience of 40 major government projects. We examined specifically whether:

- the programme had a clear rationale and objective: we concluded that the Department had poorly articulated the strategic case for high-speed rail
- the programme was well-priced and affordable: we concluded that cost estimates were at an early stage and there were risks to affordability
- plans for programme delivery were realistic and feasible: we concluded that the timetable for planning phase one, from London to the West Midlands, was challenging
- roles and governance arrangements were clear: we concluded that there were weaknesses in programme management which the Department was taking steps to improve.

During the construction phase of a project we would typically examine a major programme once a key stage had been completed, looking in more depth at issues such as whether the programme is being delivered on time and within budget, and whether risks to delivery are being effectively managed. Our recent reports on Thameslink and Crossrail are examples of such examinations.

Even when the project is completed and is operational, an issue remains for us as to when our final and, in effect, ex-post assessment should be carried out, as it can take years for the full benefits to be realised. However, experience has shown that some of the information that we need for our assessment, particularly on costing, is most likely to be available shortly after the completion of the project.

Our March 2012 report on the completion and sale of High Speed 1 included a cost benefit analysis, which involved reassessing the costs and journey time saving benefits of the project. The analysis enabled us to determine the likely benefits and costs to the taxpayer. We made the following conclusion, which illustrates the difficulties in drawing a value-for-money conclusion at this stage:

In assessing whether a project will deliver value for money, the Department considers a wide range of impacts that a project might have, some of which it can quantify and others on which it has to make more qualitative judgements. The original business case in 1998 was based on benefits to transport users, from faster journey times and increased rail capacity, and regeneration benefits. The data available only allows us to estimate that the value of journey time savings benefits, over a 60-year appraisal period to 2070, would be GBP 7 000 million. We estimate that the net costs to the taxpayer to 2070 would be GBP 10 200 million. On these measures we would conclude that the project is not value for money. When including other impacts from the project, some of which are unmeasurable, we accept that such a clear conclusion is not possible. The Department, however, would need to demonstrate that these benefits are going to be at least GBP 8 300 million, giving a higher contribution than originally expected, to achieve the benefit-cost ratio of 1.5 to 1 estimated in 1998.

At the time of publishing the report in March 2012, the Department was developing a plan to evaluate the High Speed 1 project. The Department's view was that a robust assessment of transport benefits from the high-speed line could only be made after December 2012, three years after the introduction of the domestic high-speed services. This was based on standard industry demand forecasting guidance, which states that the expected change in demand would be complete three years after a major service change, with the majority of change occurring by the end of the first year. However, we believed and stated that the Department should already have had an evaluation plan in place that identified the data it needed to collect and monitor to measure project benefits. We acknowledged that the Department had started work to identify the method it would use to evaluate wider economic impacts and regeneration benefits and how it would establish a counterfactual. We felt, however, that there was a risk that the Department would not be able to measure robustly the impact of the project because it was not able to demonstrate that it had collected the information it would need. The Department is currently carrying out its evaluation of the project, and a report is expected to be published in 2014.

Evaluating whether wider economic impacts have been achieved

One of the challenges for both ex-ante and ex-post assessments of transport projects is the measurement of wider economic impacts. This is illustrated by the High Speed 1 programme. The main project benefits the Department identified in 1998 were benefits to transport users: faster journey times and increased passenger rail capacity, and regeneration benefits. The Department chose to route the line through east London to stimulate regeneration. Including a monetary value for regeneration was unconventional for a public transport project at the time because the Department did not have an agreed method for calculating such benefits. The Department's approach for other projects such as Thameslink, Crossrail and High Speed 2 has been to base the initial benefit cost ratio on the transport benefits and to then produce an additional benefit cost ratio which includes an assessment of wider economic impacts. In the case of High Speed 1, it valued the expected regeneration benefits based on the 50 000 jobs that it originally estimated the line would create at sites around the three international stations, and the amount that the Government would be prepared to pay to create these jobs through other interventions.

When we examined the project in 2012, we found that London and Continental Railways (who delivered the project) had commissioned analysis of the approved developments at all three locations and estimated they would support at least 70 000 jobs. The Department had yet to review the regeneration benefits and told us that they intended to do so after the 2012 Olympic Games, when the legacy plan for the Olympic Park had been implemented. This was because the high-speed line has a station at Stratford, where the Olympics took place. Under its transport analysis guidance the Department would need to identify the impact, for example, on unemployment in areas served by the high-speed line to quantify regeneration benefits. The Department told us that these impacts are not as easy to isolate as the impact on transport benefits, where the Department already collects data, and a specific study would be required to assess the wider economic and regeneration benefits.

As can be seen above, we did not attempt to quantify the wider economic impacts ourselves, but used our reworking of the cost benefit analysis to show the scale of benefit needed if the project was to deliver value for money.

It appears to us that in the United Kingdom there is an increasing interest in ex-post assessments and in making investment decisions using benefit-cost ratios, which capture the full benefits of the project. For the latter, this involves having a better understanding of the wider impacts that infrastructure projects such as those in transport generate. For example, the July 2014 Committee of Public Accounts report on Crossrail recommended that the Department should improve its understanding of the wider economic benefits of transport projects and include this in its investment decisions. The Department told the Committee that it agreed it needed to do more work on understanding wider economic benefits, such as changing land use, since these could not currently be quantified in the benefit-cost ratio.

Making an assessment when there is insufficient data and documentary evidence

The availability of data will usually determine what methods we use and the focus of the report. While we can carry out or commission primary research, in the transport sector we generally analyse existing datasets held by the Department, regulatory bodies or other stakeholders such as the rail infrastructure manager, Network Rail. These datasets vary in their complexity, completeness and comparability, and we plan our work taking into account their limitations.

In 2011, we published a report that examined the Department's decision to terminate the InterCity East Coast franchise, in response to the National Express Group stating that it would no longer financially support the franchisee, National Express East Coast. We wanted to determine whether the Department's decision to terminate the franchise offered better value for money than renegotiating the terms of the contract with National Express, or negotiating a consensual exit from the contract. The Department had not carried out this analysis and believed that relaxing contract terms would have encouraged the operators of other rail franchises to seek similar deals, exposing the taxpayer to increased costs. The Department's data was not held in a way that would easily allow analysis of the available options, which were: to renegotiate the terms of the franchise with the operator; negotiate a consensual exit; or terminate for contract default. We therefore drew together the Department's data on the actual and forecast financial performance of train operators facing financial difficulties, and adjusted the data so that it was comparable. We used this evidence to develop a financial model that allowed us to calculate the potential costs to the taxpayer of the three options available to the Department. The analysis helped to support our overall value-for-money conclusion that the Department's decision to terminate the franchise was the best means of protecting the taxpayer, when compared to other potential options.

Assessing the impact of an intervention when there are other factors and agencies at play

A challenge in assessing the operations of government transport agencies is attributing ultimate outcomes to their activity. For example, a number of factors contribute to road safety. Although the Vehicle and Operator Services Agency's work (described below) had an impact on road safety, there are many other factors and agents other than the Agency which affect the rate of road accidents, including weather patterns, economic growth rates, drivers' health and drivers' behaviour patterns.

In January 2010, we published a report that examined the work of the Vehicle and Operator Services Agency (VOSA). Until its replacement by the Driver and Vehicle Standards Agency in April 2014, VOSA was the executive agency of the Department for Transport, responsible for ensuring that

Heavy Goods Vehicles (HGVs) and Public Service Vehicles (PSVs) complied with a wide variety of roadworthiness and traffic regulations covering physical maintenance, weight limits and drivers' hours. We wanted to determine whether the benefits from the Agency's enforcement activities met its costs. One aspect of our methodology was to conduct a benefit-cost calculation of its enforcement work. This enabled us to conclude that the benefits are likely to exceed the Agency's expenditure, but in our opinion it could deliver significantly better value for money. To carry out our assessment, we:

- estimated the proportion of vehicles with defects for which the Agency checks, that go on to cause accidents
- used this to estimate the number of accidents prevented by VOSA's inspections
- estimated the average benefit of preventing an accident involving an HGV
- applied this to the number of accidents prevented by VOSA's activities to estimate their value to the economy.

We calculated that if the Agency's roadside checks prevented 283 accidents and the average value of each of these was GBP 143 529, the roadside checks would have delivered GBP 40.7 million of benefits. This compares to the Agency's expenditure of GBP 32.9 million on HGV enforcement in 2008-09. We also performed sensitivity analysis by varying the number of accidents prevented by the Agency and the average benefit of a prevented accident by up to +/-50% and then looking at the effect different combinations of these changes had on the benefits delivered by roadside checks. This suggested that roadside checks could deliver a minimum of GBP 10 million and a maximum of GBP 91 million in benefits. The benefit exceeded the Agency's expenditure in 2008-09 in 44% of the combinations analysed.

This assessment enabled us to conclude that the benefits are likely to exceed the Department's expenditure and the result was included in our value-for-money conclusion on VOSA's enforcement activities. The conclusion, which found that the Agency had achieved "satisfactory results" also highlighted other issues, such as the potential to improve value for money through system improvements, for example, staff deployment, better location of the sites for checking vehicles and closer working with the Department. This example also serves to illustrate how our value-for-money conclusions take account of multiple aspects to reach an overall judgment. Benefit-cost or other quantitative analysis plays an important part but is not the whole picture.

Our recommendations leading to improvement

Although our assessments are not strictly ex-post evaluations, our value-for-money reports include a number of recommendations for the audited body. As explained earlier, the Committee of Public Accounts will use our reports as the basis of a hearing to question the government officials responsible and subsequently publish its own report. Its report will also contain recommendations which the Department must respond to, and a summary of its response is made public in HM Treasury's Treasury Minutes.

In our reports on major transport projects, the recommendations generally address the same areas identified by De Jong et al. (2013):

- Improvement of cost and benefit estimation approaches
- Risk management measures
- Increasing accountability
- Clarifying project scope and objectives.

Some examples of how we have addressed these themes include:

- **Improvement of cost and benefit estimation approaches:** Our recommendations in this area have focused on the Department's work in ensuring and making more transparent the robustness of its cost benefit analysis.

  Our March 2012 report on High Speed 1 recommended that the Department ensures its demand forecasts, which feed into the benefit cost estimation, are subject to rigorous scrutiny and scepticism. We also recommended that it should assess the benefits under a range of different scenarios, perform a sensitivity analysis of key assumptions and a sense check to understand the reality of meeting forecast demand. In the November 2012 Treasury Minutes, the Department stated that it now takes greater account of downside risks, and typically undertakes extensive sensitivity analysis to test the robustness of the business case to varying input assumptions.

  Our May 2013 report on High Speed 2 highlighted a number of issues with the calculation of the benefit cost ratio in the economic case, including errors in earlier calculations, the need to update the data underpinning key assumptions and the lack of analysis of the effect of premium pricing on forecast passenger demand. The revised business case in October 2013 did seek to address some of these concerns by updating some of the data and revising some assumptions, for example, around journey time savings, and by quoting the benefit cost ratio as a range rather than point estimates to recognise explicitly the uncertainty of the economic case.

- **Risk management measures:** The importance of risk management was particularly evident in our examination of the failure of Metronet in 2009. Metronet was a private infrastructure company responsible for the maintenance and upgrade of sections of the London Underground. It went into administration in July 2007. Although Transport for London (TfL) had guaranteed 95% of Metronet's borrowing, the Department had also informally given assurances to investors that it would guarantee the borrowing[2]. When Metronet failed, the Department had to make a grant payment of GBP 1.7 billion to help London Underground purchase Metronet's debt obligations, a sum that would otherwise have been repaid over the 30-year life time of the contracts. The Department was exposed to this risk but lacked direct ways of gaining assurance over the management of the risk[3]. We advised that the Department should: collect and analyse a range of financial and performance data held by parties to the contract or available independently; request regular risk reports from London Underground and TfL as the contracted clients; and review the devolved body's understanding of the key risks to the project to allow it to identify and investigate any issues relevant to the management of its own risk.

  The quality of the Department's oversight of large programmes is an issue that we continue to examine. For example, our recent report on Crossrail commented favourably on the Department and Transport for London's oversight of that programme, highlighting the use of a probability-based approach to forecasting the delivery date and final cost, and to monitor and manage risks which allows the sponsors and Crossrail Limited to identify when there are risks to delivery and to take action to mitigate those risks.

- **Increasing accountability:** In our December 2012 report on the cancellation of the InterCity West Coast franchise competition, we found that staff in the project team reported to different parts of the organisation, which meant no one person oversaw the whole process, or could see patterns of emerging problems. We recommended the Department appoint someone with sufficient seniority to oversee each significant commercial transaction and major project, with the knowledge, skills and authority within the Department to take action if things are going wrong. The Department has since taken action to review its existing Senior Responsible Owners (SROs) to ensure they have the right seniority, experience and expertise for the projects for which they are currently responsible. Additionally, it has provided training for SROs to ensure they fully understand their responsibilities.
- **Clarifying project scope and objectives:** As can be seen above, one of our criticisms of the early preparations for High Speed 2 was the Department's poor articulation of its strategic case for the route. The Committee of Public Accounts subsequently called for the Department to publish detailed evidence, which clearly showed why it considered High Speed 2 to be the best option for increasing rail capacity into London, improving connectivity between regional cities and rebalancing the economy. The Department sought to address the Committee's concerns in its revised strategic case for the project in October 2013.

The extent to which recommendations are implemented has varied and in recent years we have sought to address this by following up implementation more systematically. As can be seen above, some issues such as the robustness of underlying data for business cases and the quality of risk management are recurring themes in our work. By returning to these themes, we hope to increase focus on them within audited bodies and raise standards over time. Moreover, where we have conducted early examinations of programmes, we examine specifically in subsequent reports whether earlier recommendations have been implemented. One example of this is our final report on High Speed 1 in May 2012, in which we commented that the Department had not yet reassessed the project costs and benefits since 2001, despite making a commitment to the Committee of Public Accounts to do so.

Notes

1. We may choose to examine the objectives and rationale for the programme; progress in delivering the programme to time and budget; and, looking forward, whether risks to delivering the next stage of the programme are being managed effectively. This includes whether the Department is collecting the information and establishing the baselines to enable it to conduct ex-post assessments in the future.

2. The Department provides a grant to TfL. London Underground is a subsidiary of TfL.

3. Under the Greater London Authority Act 1999, strategic and investment responsibility for London Underground was devolved to TfL and the Mayor of London. The Secretary of State of the Department could only direct the Mayor to make changes to transport strategy where it would be inconsistent with national policy and have an adverse effect outside London. DfT was not a party to the contracts and had no direct influence over performance.

Annex 3.1

Recent NAO value-for-money studies on transport

All reports available at www.nao.org.uk:

- Procuring new trains (July 2014)
- Maintaining strategic infrastructure: roads (June 2014)
- Crossrail (January 2014)
- Progress in delivering the Thameslink programme (June 2013)
- High Speed 2: a review of early programme preparation (May 2013)
- Lessons from cancelling the InterCity West Coast franchise competition (December 2012)
- Funding for local transport: an overview (October 2012)
- The completion and sale of High Speed 1 (March 2012)
- Reducing costs in the Department for Transport (December 2011)
- Local Authority Major Capital Schemes (May 2011)
- Regulating Network Rail's efficiency (April 2011)
- The Intercity East Coast passenger rail franchise (March 2011)
- Procurement of the M25 private finance contract (November 2010)
- Increasing passenger rail capacity (June 2010)
- Highways Agency: Contracting for Highways Maintenance (October 2009)
- The Department for Transport: The failure of Metronet (June 2009).

Annex 3.2
Analytical framework for assessing value for money

The framework on the following pages is a key reference source for NAO auditors scoping and planning value-for-money work. It provides a guide to the types of question we will need to answer to draw clear value-for-money conclusions on topics across the range of government activities. The framework is necessarily broad and acts as a foundation for the development of audit teams' work. Its application to a particular topic will be informed by auditors' professional judgment and experience. Neither this nor any other tool can provide a simple, "mechanical" conclusion on value for money.

NAO National Audit Office

Analytical framework for assessing Value for Money

"Good value for money is the optimal use of resources to achieve the intended outcomes."

The objective of our VFM work is to form a clear judgement on whether value for money has been secured in the area under examination. This analytical framework is designed to help teams do this in a consistent manner.

At **planning stage** teams should follow Steps A and B as set out on page 2. Having outlined the focus of the report, it is essential to define at this stage what the comparators of good performance will be for each of the study elements. This is crucial not only for a consistent logic of the report, but also to determine how we will measure performance. It also forces us to critically assess what we think possible. Our definition of VFM refers to 'optimal' use of resources – in practice, defining and securing 'optimal VFM' is inherently difficult and often impossible. Therefore, teams are encouraged to establish for each of their evaluative study elements what 'good' would look like in the particular circumstances in 'place' (assisted by the comparator matrix, page 4).

During **fieldwork** teams will capture the totality of resources relevant to the subject being examined (Step C), as well as all the relevant processes and outputs/outcomes achieved (Step D). This is the stage in the process where we identify data gaps, test causation as well as learn about wider aspects of the environment in which the client operates which might constrain what is achievable (Step E). All those issues have a considerable impact on the type of conclusion we can draw in due course.

When **analysing** the evidence, teams should in Step F refer to their previously identified comparators of good performance (Step B) and contrast observed performance with what – under the given circumstances – could reasonably have been expected. This always involves assessing and then concluding on a given level of performance in the context of the resources used to achieve this (analysing, as appropriate, economy, efficiency or effectiveness). During this step teams should refer to the *Guide on Drawing and Drafting VFM Conclusions*.

Having identified where the weaknesses in performance and arrangements lie, teams can then make robust **recommendations** designed to improve value for money (Step G)

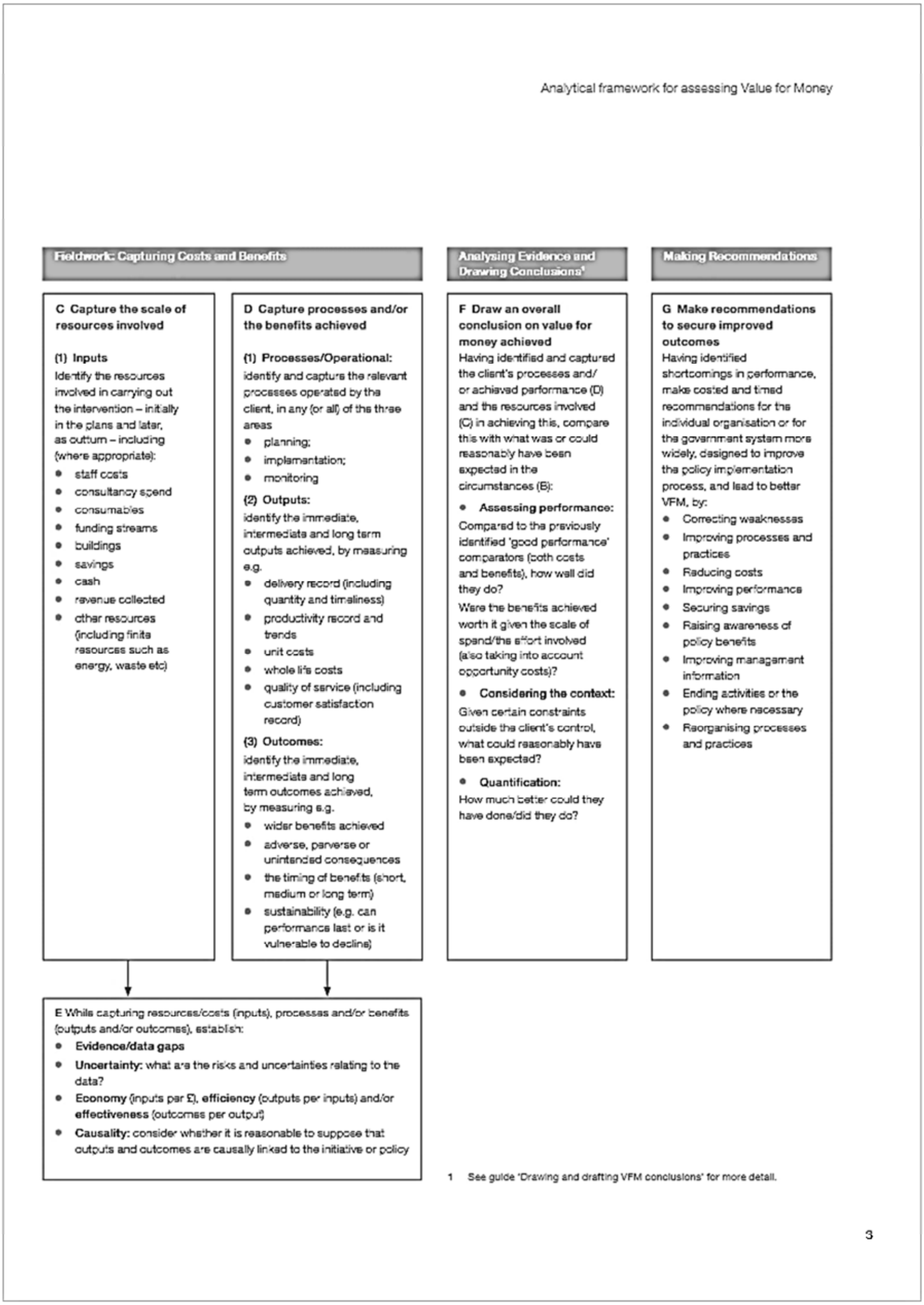
1

Analytical framework for assessing Value for Money

Planning: Defining the evaluative framework of the study

A Establish the focus of the report	B Establish the comparator for what good performance would be (See Comparator Matrix on page 4)
(1) Inputs: → The resources that contribute to the production and delivery of an output	**(1) Previous performance** – over time/trend analysis **(2) Programme targets** **(3) External or internal performance benchmarks**
(2) Operational Processes: → A series of linked activities leading towards outputs	**(1) Operational good practice, e.g.** **Optimal planning:** Being clear about what is wanted. For example: • Well evidence based initiative with clear objectives, based on appropriate best practice/industry benchmarks, agreed targets etc • Realistic plans taking account of constraints and developed in the light of an understanding of risks • Stakeholder support for plans • Required resources in place to deliver plan • Clear understanding of what relevant costs should be **Optimal implementation:** Delivering or procuring well. For example: • Having the right technical and managerial skills • Selecting the best delivery or procurement route • Avoiding false economies • Managing supply markets • Obtaining economies of scale through programmes of investment or contracts • Minimum sustainable impact of procurement • Establishing clear and good VFM risk transfer **Optimal monitoring:** Being able to assess performance. For example: • Clear performance measures set and monitoring arrangements in place • Sound performance management systems • Sound policy evaluation process in place, with timely feedback • Ability to make operational changes well in light of feedback • Ongoing value testing • VFM assessment methodology in place **(2) Previous Performance** **(3) Programme targets**
(3) Outputs: → The immediate results of Government activities **(4) Outcomes:** → The ultimate impacts, or consequences, of the activities	**(1) Previous performance** **(2) Programme targets** **(3) External or internal performance benchmarks** **(4) Counterfactual** – what would outputs be if alternative action had been taken?

2



Analytical framework for assessing Value for Money

Fieldwork: Capturing Costs and Benefits

C Capture the scale of resources involved

(1) Inputs
Identify the resources involved in carrying out the intervention – initially in the plans and later, as outturn – including (where appropriate):

- staff costs
- consultancy spend
- consumables
- funding streams
- buildings
- savings
- cash
- revenue collected
- other resources (including finite resources such as energy, waste etc)

D Capture processes and/or the benefits achieved

(1) Processes/Operational:
identify and capture the relevant processes operated by the client, in any (or all) of the three areas

- planning;
- implementation;
- monitoring

(2) Outputs:
identify the immediate, intermediate and long term outputs achieved, by measuring e.g.

- delivery record (including quantity and timeliness)
- productivity record and trends
- unit costs
- whole life costs
- quality of service (including customer satisfaction record)

(3) Outcomes:
identify the immediate, intermediate and long term outcomes achieved, by measuring e.g.

- wider benefits achieved
- adverse, perverse or unintended consequences
- the timing of benefits (short, medium or long term)
- sustainability (e.g. can performance last or is it vulnerable to decline)

E While capturing resources/costs (inputs), processes and/or benefits (outputs and/or outcomes), establish:

- **Evidence/data gaps**
- **Uncertainty:** what are the risks and uncertainties relating to the data?
- **Economy** (inputs per £), **efficiency** (outputs per inputs) and/or **effectiveness** (outcomes per output)
- **Causality:** consider whether it is reasonable to suppose that outputs and outcomes are causally linked to the initiative or policy

Analysing Evidence and Drawing Conclusions[1]

F Draw an overall conclusion on value for money achieved
Having identified and captured the client's processes and/or achieved performance (D) and the resources involved (C) in achieving this, compare this with what was or could reasonably have been expected in the circumstances (B):

- **Assessing performance:**
Compared to the previously identified 'good performance' comparators (both costs and benefits), how well did they do?
Were the benefits achieved worth it given the scale of spend/the effort involved (also taking into account opportunity costs)?
- **Considering the context:**
Given certain constraints outside the client's control, what could reasonably have been expected?
- **Quantification:**
How much better could they have done/did they do?

Making Recommendations

G Make recommendations to secure improved outcomes
Having identified shortcomings in performance, make costed and timed recommendations for the individual organisation or for the government system more widely, designed to improve the policy implementation process, and lead to better VFM, by:

- Correcting weaknesses
- Improving processes and practices
- Reducing costs
- Improving performance
- Securing savings
- Raising awareness of policy benefits
- Improving management information
- Ending activities or the policy where necessary
- Reorganising processes and practices

1 See guide 'Drawing and drafting VFM conclusions' for more detail.

3

Analytical framework for assessing Value for Money

Comparator Matrix

The matrix below illustrates what types of comparator are generally suitable for assessing the performance of particular aspects of a policy/programme.

Teams should use this matrix during study planning to think through systematically how the VFM of the different aspects of their study are going to be assessed. The types of comparator used will determine what type of conclusion we can draw (see Guide on Drawing and Drafting VFM Conclusions, for more detail about different types of conclusions).

		Element(s) of performance analysed						Type of Conclusion
	Relative to	Inputs [Economy]	Processes/Operational: Planning	Implementation	Monitoring performance	Outputs [Efficiency]	Outcomes [Effectiveness]	
Comparator applied (specify how this will be assessed/ measured)	'alternative action' counterfactual	x	–	–	–	x	x	Type A
	'do nothing' counterfactual	x	–	–	–	–	x	Type A
	external benchmarks	x	–	–	–	x	x	Type B
	internal benchmarks	x	–	–	–	x	x	Type B
	programme targets	x	–	–	–	x	x	Type B
	previous performance	x	x	x	x	x	x	Type B
	operational good practice	x	x	x	x	–	–	Type C
	insufficient data to allow for assessment	x	x	x	x	x	x	Type D

For example, if we are examining the procurement processes of a Department we can either assess them by comparison with an established set of good practice criteria, or (if the processes have previously been assessed) we can comment on whether their monitoring regime has improved over time (or both). If we assessed the outputs of those procurement processes, we could compare them with the Department's output targets, contrast outputs over time or compare the performance of different units of the Department. For a more objective, 'stronger' assessment of performance, we could contrast performance with that of comparable external organisations.

4

References

Campbell, C. and C. Rozsnyai (2002), "Quality Assurance and the Development of Course Programmes", Papers on Higher Education Regional University Network on Governance and Management of Higher Education in South East Europe, Bucharest, UNESCO.

Comptroller and Auditor General (July 2014), "Department for Transport: Procuring new trains", Parliamentary Session 2014-15, HC 531.

Comptroller and Auditor General (June 2014), "Department for Transport and Highways Agency: Maintaining strategic infrastructure: roads", Session 2014-15, HC 169.

Comptroller and Auditor General (May 2013), "Department for Transport: High Speed 2: A review of early programme preparation", Session 2013-14, HC 124.

Comptroller and Auditor General (December 2012), "Department for Transport: Lessons from cancelling the InterCity West Coast franchise competition", Session 2012-13, HC 796.

Comptroller and Auditor General (October 2012), "Department for Transport: Funding for local transport: an overview", Session 2012-13, HC 629.

Comptroller and Auditor General (March 2012), "Department for Transport: The completion and sale of High Speed 1", Session 2010-2012, HC 1834.

Comptroller and Auditor General (December 2011), "Guide: Initiating successful projects".

Comptroller and Auditor General (March 2011), "Department for Transport: The Intercity East Coast passenger rail franchise", Session 2010-2011, HC 824.

Comptroller and Auditor General (June 2010), "Department for Transport: Increasing passenger rail capacity", Session 2010-11, HC 33.

Comptroller and Auditor General (January 2010), "Vehicle and Operator Services Agency: Enforcement of regulations on commercial vehicles," Session 2009-2010, HC210.

Comptroller and Auditor General (June 2009), "Department for Transport: The failure of Metronet", Session 2008-2009, HC 512.

Comptroller and Auditor General (July 2005), "Department for Transport: Progress on the Channel Tunnel Rail Link", Session 2005-2006, HC 77.

Comptroller and Auditor General (March 2001), "Department of the Environment, Transport and the Regions: The Channel Tunnel Rail Link", Session 2000-2001, HC 302.

De Jong et al. (2013), "How to Build Major Transport Infrastructure Projects within Budget, in Time and with the Expected Output; A Literature Review", Vol. 33, Issue 2, pp. 195-218.

Department for Transport (October 2013), "Monitoring and Evaluation Programme"

Department for Transport (March 2013), "Monitoring and evaluation strategy", HM Treasury (all reports available at: www.gov.uk/government/collections/treasury-minutes)

Department for Transport (May 2013), Treasury Minutes, Cm 8613, pages 34-38.

Department for Transport (November 2012), Treasury Minutes, Cm 8467, pages 25-29.

House of Commons Committee of Public Accounts, (July 2014), Crossrail, Eighth Report of Session 2014–15, HC 1062.

National Audit Office (May 2011), "Department for Transport: Local Authority Major Capital Schemes".

Chapter 4
TPICS, TIGER and United States experience: A focus on case-based ex-post economic impact assessment

Stephen Fitzroy, Glen Weisbrod and Naomi Stein - Economic Development Research Group, Inc., Boston, Massachusetts, USA.

This chapter presents the results of research recently conducted by the authors on ex-post analysis focused on the long-term economic impacts of transportation system investments in the United States. For a variety of reasons, the United States has had a tradition of making transport investments to address economic development goals and applying ex-post analysis to assess achievement of economic development impacts. These past studies are reviewed, as are some of the deficiencies and suggested improvements in methods for ex-post analysis.

The chapter also reviews methods to refine ex-post analysis of economic development in the United States via the Transportation Impact Project Case Studies (TPICS) system developed as a national database of land and economic development impact studies.

Introduction: Use of ex-post studies to assess achievement of economic impact goals

Ex-post (after the fact) analysis provides an important means of learning from the past. In the field of transportation planning, ex-post studies have received particular attention when focused on major investment projects, particularly cases where costs have ended up being higher and travel benefits lower than originally expected. Yet, ex-post analysis can do much more than simply assess the accuracy of estimated costs, forecast benefits and resulting cost-benefit ratios. It can also be used to observe the extent to which projects achieve desired and intended social and economic goals. Ex-post analysis can also be used on a continuous basis to monitor the long-term performance of transportation infrastructure investments.

In the United States, there is a history of establishing specific funding to achieve particular social and economic goals (e.g. safety enhancement, air quality enhancement, poverty reduction, expansion of employment opportunities, neighbourhood revitalisation (regeneration), and land redevelopment in target areas). The Federal Government has had designated funding programmes aimed at projects to achieve each of these above-cited goals (Weisbrod, 2000; US Department of Transportation, 2013). The individual states, which generally prioritise projects based on forms of multi-criteria analysis, often explicitly incorporate similar goal elements into their rating criteria (Weisbrod, 2011).

Using results of ex-post analysis of economic development impacts

There is a critical difference between ex-post and ex-ante studies of economic impact. Simply stated, forecasts of economic development impact that are included in ex-ante analyses are typically based on economic models accompanied by the simple caveat that they do not cover unforeseen circumstances that may affect future economic growth in a given study area. In contrast, ex-post evaluations face the much harder task of actually distinguishing the net effect of the transport investment from those totally unforeseen circumstances that were so conveniently dismissed as unmeasurable in the ex-ante studies. That reality has caused much of the United States experience with ex-post analysis to focus on whether or not it is possible to extract “lessons learned”. These are ways that future project design, implementation and/or impact forecasting may be improved by better identifying and anticipating risk or vulnerability to exogenous factors (that can accentuate or minimise actual economic development impacts).

Stepping back, it is useful to contrast this approach to ex-post analysis with the greater effort made in many European studies that have concentrated more on validating benefit-cost methodologies and assessment outcomes – for a variety of performance outcomes (Louwa et al., 2013; Kjerkreit, 2008; Chapulut et al., 2005; Braathen and Hervik, 1997; Florio and Vignetti, 2013). The US emphasis has been more targeted at economic impacts specifically, and tends to be more interested in developing planning insights; while less focused on assessing investment criteria. The objective for many of the US studies has been to help identify factors that mitigate negative and accentuate positive potential impacts of different types of transportation investment.

Many of the past United States case studies were originally commissioned to support future project development strategies by learning from past experience. For example, when the Pennsylvania Turnpike was considering constructing an interchange with a major interstate highway, the Pennsylvania Turnpike Commission funded a set of ex-post case studies of other built interchanges to identify how they affected the economic development of their local communities. The study found that changes in accessibility from the completion of a new interchange between major highways did generate development pressures in the area around the interchange, but that actual outcomes varied widely and were shaped more by land use limitations and regulations, local road network capacity and development incentives (Wray et al., 2000).

Similarly, when the City of Roanoke (Virginia) was considering whether a new interstate highway should be routed through the city or bypass around it, it funded a set of ex-post case studies of economic impacts in other cities that had new highways built (EDR Group, 2000).

There is a particularly rich literature on ex-post evaluations of highway bypasses in the United States. In contrast to larger corridor programmes, bypass and interchange projects can be easier to evaluate, as their impacts tend to be visible local land-use and development impacts. The bypass studies compare the levels of business activity in town centres before and after the construction of the bypass highway segments. Studies have been conducted throughout the United States: in Wisconsin (Wisconsin DOT, 1998), Kansas (Burress, 1996), Iowa (Anderson and Otto, 1991), Washington State (Gillis and Casavant, 1994), Texas (TTI, 1995), North Carolina (Blackburn and Clay, 1991), and elsewhere (Collins and Weisbrod, 2000). As an example, California DOT chose to build on this experience with bypass case studies. When the agency was considering its policy on rerouting local highways around smaller communities it chose to commission a meta-analysis of those bypass case studies (System Metrics Group et al., 2006).

The goal of encouraging economic development has historically been (and currently still is) a major motivation for a large share of transportation investments in the United States. This applies to everything from funding for major programmes to funding for individual projects. There are a variety of different examples of transportation investment programmes to achieve economic development goals – particularly programmes and projects aimed at helping to expand local and regional economies. National programmes such as the TIGER (Transportation Investment Generating Economic Recovery) Program have sought to "jump-start" the economy in areas of high unemployment (US DOT, 2014). The programme rules aim to focus funds on creating both immediate construction jobs and also greater long-term job growth in those areas. Longstanding programmes such as the ADHS (Appalachian Development Highway System) have been funded to reduce isolation in an eastern US mountain region where the closure of coal mines and other extractive industries left pervasive unemployment and little access to markets. Looking back as far as fifty years ago, to the time that the interstate highway system was designed and funded to provide a new form of network connectivity to link urban markets across the nation, there has been a widely recognised link between transportation infrastructure investment and economic development (Pfeiffer, 2006). Common to all of these programmes is the fact that transportation system investment was assumed to be the primary – and often the only – major investment required to achieve long-term social and economic goals, thereby fostering a shift in economic activity patterns. None of these transportation investment programmes were justified solely because they would reduce user travel costs for traffic moving between existing origins and destinations.

Perhaps the reason why travel time-savings was not used as justification is because of the nation's large size and constantly evolving population and economic patterns. Over time these led to congestion and fast growth in some areas, while other previously thriving areas became relatively isolated and poorly positioned for newly evolving forms of economic activity. The result, though, has been a continuing set of transportation investment programmes and projects aimed specifically at helping regional (and sometimes local) economic development by enabling expansion of job markets for area residents, customer markets for area businesses and inward investment to enable new land development and economic activities to emerge. Over time, there has been an evolving stream of ex-post studies to observe the extent to which those intended goals were achieved, though only a small portion of all transportation investment programmes have had such follow-on analysis. Together, these studies constitute a generally coherent trajectory towards broader application of ex-post analysis.

History of ex-post studies examining economic impacts

The documentation of actual (ex-post) economic development impacts of transportation investment goes back to anecdotal studies nearly two hundred years ago. One of the first large-scale US transportation projects – the Erie Canal –opened in 1820 to connect agricultural regions in the Ohio Valley with the urban population centres of the eastern seaboard. Follow up ex-post observations made in that era documented a twenty-fold drop in the price of wheat in urban markets, followed by a major movement of population and economic activity to the Ohio Valley (New York State Archives, 2014).

Regional programme impacts

Systematic ex-post analysis of the economic impacts of transport investment programmes in the modern era goes back at least as far as work documenting impacts of rural interstate highways by Miller (1979). That work examined economic growth rates among rural counties and compared counties through which new highways passed to counties that did not have new highways built. More sophisticated statistical procedures were applied to this same line of inquiry by Rephann and Isserman (1994) and later Lynch (2007). Those latter studies utilised a time-series, cross-section statistical framework with "matched pair" controls for each affected county. Each county experiencing new highways was matched to another county that had no such highway changes, yet was similar in population, economic profile, income level, distance from larger cities and access to interstate highways as of 1959. Both statistical studies found evidence of positive economic growth associated with new highways. The latter study also found differential impacts associated with interstate highways (limited access motorways) as compared to Appalachian Development Highways (that typically do not have limited access).

An alternative methodological approach was undertaken in a 1995 study that tracked changes in regional employment in the Mississippi Delta region, relative to overall national trends, in response to a series of highway, seaport and railroad improvements (US FHWA, 1995). Rather than using statistical controls, the study opted for an interview-based approach to establish causality between transport infrastructure investment and resulting higher rates of employment growth. This was an early example of relying on on-the-ground research to understand the complex relationship between transportation improvements and economic impacts.

National guidance

In 1991, the US General Accounting Office (now known as the US Government Accountability Office [GAO]) issued broad guidance on ex-post programme evaluation, which called for ex-post comparison with statistical control for exogenous underlying changes over the time period as well as an effort to establish attribution of credit (US GAO, 1991). This GAO guidance was then used when the Appalachian Regional Commission (ARC) funded 100 case studies of the impacts of its local public works (roadway and water/sewer) projects in 2000 and an additional 100 case studies conducted in 2007 (Brandow and EDR Group, 2000; Bizminer/Brandow and EDR Group, 2007). These studies were in response to a finding by the US GAO, assessing the lack of information about the impact of economic development initiatives implemented by several federal agencies, including ARC (US GAO, 1996). It is interesting to note that those studies found that a significant portion of the cases had a smaller-than-expected economic impact, though others had a higher-than-expected impact leading, in aggregate, to slightly higher economic impacts than originally expected.

In 2001, the FHWA sponsored development of a guidebook that provided public agencies and private researchers with guidance and standards for documenting the actual ex-post economic effects of highway investments (EDR Group and Cambridge Systematics, 2001). Titled *Using Empirical Information to Measure the Economic Impact of Highway Investments,* the guide offers three prototype designs for empirical studies of the actual economic development impacts of specific highways at regional, corridor and local levels. It includes a discussion of appropriate impact measures, as well as the types of data that can be used to collect information on these measures (i.e. published sources, survey and interview data, and site observation and field work).

This guide recommended methods for ex-post time analysis that included both statistical control comparison and interview methods to determine attribution of partial credit to the transport investment. It also suggested relevant metrics: employment, wages, GDP, land development and building investment. It further noted that economic impacts typically evolve over a period of ten years in a distinct sequence: land values start to rise, followed later by building development, and yet later by observations of job and wage growth. All the while, tax revenues can increase in different ways as a consequence of each of these steps.

Project-specific studies

The US Federal Highway Administration (FHWA) subsequently sponsored two major studies to demonstrate ex-post assessment methodologies. These were studies of the economic and land-use impacts of constructing the four-lane Highway 29 in Wisconsin (FHWA, 2002) and economic development impacts associated with completion of a four-lane expressway and designation of it as the new Interstate Highway I-86 in New York State (FHWA, 2003). Both evaluations drew on interviews with local officials and representatives of the business communities in the respective project areas. They both found evidence that expected economic impacts were indeed emerging, though both projects were completed within five years of the study and it was too soon to definitively determine that they would reach the originally expected economic development impacts.

The FHWA continued to invest in ex-post evaluation. A 2005 study developed "economic histories" of selected rural interstates through ex-post case studies that tracked changes occurring over time in rural and small urban areas that have become connected by the new interstate highways. The study focused on "secondary interstate corridors" – i.e. those interstate highways completed after 1970 that provide regional connections for smaller communities linking to the backbone, coast-to-coast routes (FHWA, 2004). That study found evidence of significant changes in economic and land development for some regions but not others. When changes did occur, the most notable changes were inward investment and relocation of warehouses to enable more efficient regional distribution.

Development of a systematic approach to ex-post studies

Each of the studies described above was undertaken by different organisations for different purposes, but none examined ex-post results in a systematic way that would allow direct comparison of investments across regions using consistent methods of ex-post analysis.

The Second Strategic Highway Research (SHRP2) Program, authorised by the US Congress in 2005[1] recognised the need for systematic ex-post studies focused on examining the economic impacts of highway investments and sponsored research into a systematic approach to ex-post based analysis. The purpose of the research, as stated at the time, was summarised in the original research objectives:

> The project objectives are: (1) to provide a resource to help determine the net changes in the economic systems of an area impacted by a transportation capacity investment. The resource should include, in an economic context, impacts on land use, land value, and the

> environment; (2) to provide data and results from enough structured cases that project planners in the future can use the cases to demonstrate by analogy the likely impacts of a proposed project or group of projects (plan); (3) demonstrate how this fits into collaborative decision making for capacity expansion. (SHRP2, 2007)

A series of research reports were produced in conjunction with this research, as were a set of online tools, a database and, subsequently, a set of analytic tools designed to help assess the wider economic benefits of transportation investment (EDR Group et al., 2012; EDR Group et al., 2013). The work products of this research are being transitioned to joint sponsorship of the US DOT's Federal Highway Administration (FHWA) and the American Association of State Highway and Transportation Officials (AASHTO). This transition and the proposed implementation programme are discussed later in this chapter. The US Transit Cooperative Research Program (TCRP) has funded an extension of the case-based approach to ex-post analysis that is specifically targeting public transport project economic impact assessments. That effort is expected to begin by the end of 2014.

A national database of ex-post studies

Among the many initiatives undertaken by the SHRP2 capacity programme's sponsored research, the research sponsored under Program C03 funded development of a national database of ex-post case studies to document the actual, observed economic development impacts of highway investment. The database, as subsequently developed, included a number of prior FHWA case studies, as well as those sponsored or conducted by other groups. In addition, further ex-post studies were conducted to make a more well-rounded coverage of different project types and settings. These new studies also adhered to the FHWA guidance. Now known as TPICS (Transportation Project Impact Case Studies), the onine database and case study reference system went live in 2013 and it can now be viewed at www.tpics.us.

TPICS was built to support several different uses:

- Planners can mine the database for examples of applicable projects elsewhere to help establish reasonable expectations concerning the range of likely impacts for proposed new projects. This is called "analysis by analogy," and it can be particularly useful for early stage planning when project concepts are being considered but project details have not yet been fleshed out.
- Case study examples can be cited at public meetings to help limit unreasonably optimistic hopes and overly pessimistic fears about project impacts by noting the range of project impacts that have actually occurred in similar situations.
- Case study details can aid project design, planning and implementation processes by pointing out factors that have been found to accentuate or mitigate positive and negative impacts.
- A case study database can allow researchers to conduct broad statistical analysis of factors affecting economic and land-development outcomes. That information, in turn, can be used to improve the models used for ex-ante forecasting of impacts for proposed projects. Furthermore, it can also be used to inform decision-making on transportation programme funding.

This perspective of case-based ex-post assessment – to build on lessons learned – can be seen in the TPICS home page, which notes:

The Case Search feature allows you to search for specific types of projects in specific types of settings. So if a specific type of project has been proposed or suggested for your area, you can use this information to inform agency planners and public meeting attendees about past experiences with similar types of projects. The available information includes descriptions of project features and pre/post data pertaining to project impacts on the local or regional economy. It also includes detailed results from local interviews on project objectives, implementation issues and other factors affecting the nature of project impacts. Aerial photos and links to other reports are also provided. Lessons learned from these experiences can be used to improve project design and implementation processes.

The approach to and structure of TPICS ex-post studies

TPICS sought to establish standards for a national database of pre/post case studies. In doing so, it included requirements for: (a) ex-post impact comparison, (b) coverage of both local and regional level impacts, (c) a wide range of alternative perspectives for viewing and measuring impacts, (d) comparison of local changes over time relative to reference sources such as state and national trends, and (e) reliance on both quantitative data and qualitative observations regarding local economic conditions. In this way, the case studies highlighted the multi-faceted ways in which economic development impacts can occur, depending on the type of project and its setting.

The research programme supporting the development of TPICS involved five steps:

- Step 1: preliminary research
- Step 2: data collection design
- Step 3: product design
- Step 4: develop and analyse case studies
- Step 5: develop user-based case system and supporting documentation.

The first step involved preliminary research and development of the working papers needed to frame the TPICS development effort and also maintain liaison with other SHRP2 Capacity Program initiatives that were being conducted concurrently. The working paper reviewed research relevant to the development of case-based ex-post studies and examined past methods used in case study development and ex-post assessments in United States experience. The working paper produced several recommendations for conducting the study and further research, including the addition of more case studies and expansion to include highways connecting with both passenger and freight intermodal facilities. These two recommendations were included in the final TPICS research and products.

Step 2 focused on the structure of the database and development of the cases. This included defining the parameters used to classify cases (project type, project setting and data to be collected for each case) and a review of the feasibility of gathering this data. This step in the project included identifying approximately 240 candidate case studies and surveying known information and resources for each. Then the cases were classified according to the likelihood of being able to obtain the necessary information and their ability to contribute the coverage required under the classification system proposed for the study. As a result of these compilations, recommendations were made for the testing and validation of the case study-based approach to be undertaken in Step 4.

Step 3 involved the design of the web-based product. As noted above, the intent was to create a web-based tool that was easy to use and could serve as a reference and source for early-stage assessment of the likely range of long-range economic impacts associated with a particular project type, taking into account its setting, project characteristics and purpose. Efforts in Step 3 focused on design of the user interface and the usability of the system. As part of the design, the ability to extract data from the existing set of cases and to provide a mechanism for adding new cases over time – as users began to develop their own projects and case studies – was determined to be an important element of the system. Figures 4.1 through 4.3 show a few of the images of the initial screens seen by users that emerged from the design process. The entire system is available for inspection online at: www.tpics.us.

Figure 4.1 shows the opening screen for the entire TPICS suite of tools and resources. Users can select whether they want to search cases, use the results of the meta-analysis to conduct a preliminary assessment of a project that they have in mind, or consult the various guidelines, handbooks or instructions for use that are linked through the web-site. A user's forum and provisions for users to submit new cases are also provided. A feedback button is also included so that those having questions requiring personal interaction can submit an e-mail request to an information hot-line.

Figure 4.2 shows how cases are selected and displayed. Users simply check a box indicating the type of project they are interested in reviewing and general information about the setting, region of the country, economic conditions and motivation for the project. The database displays all cases that meet their criteria and allows the user to sort, scroll through and select cases that interest them. All information available for each case is downloadable in either a narrative or database (Microsoft Excel) format. In all, 100 project types and conditions are selected, allowing the user to view all 100 cases.

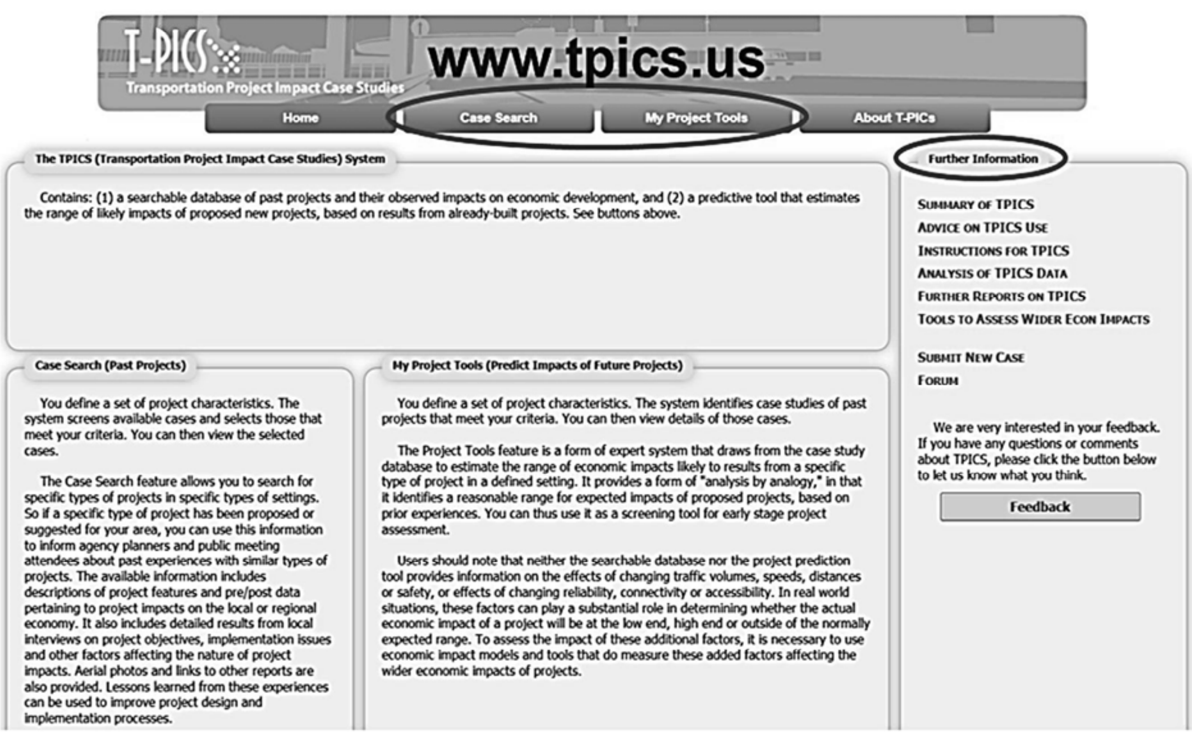
Figure 4.3 shows an example of how one particular case is presented in TPICS. A thumbnail sketch of the case appears in the upper left-hand corner. This thumbnail appears when each of the tabs (the horizontal list across the top of the screen) is clicked so that the user has some basic reference information when moving from screen to screen for a particular case. Five sets of information are available for each case:

- Case characteristics: a review of the selection parameters chosen for the case, including its location and cost.
- Setting: a summary of the demographic and economic conditions at the time the project was developed.
- Pre/Post Conditions: a table of eight measures of economic activity prior to project inception and subsequent to its completion and operation (the years of post-construction data were determined based on how each project developed and when the project was considered to have matured to the point that most of the economic impacts were likely to have been attained).
- Narrative (shown in Figure 4.3): a description of the case, findings, conditions and characteristics that comprise the body of the case study (including a list of resources, data, interviewees and other information such as links to environmental reports specific to the project).
- Impacts: estimates, based on development of each case, describing the observed effects of the project on jobs, income and output.
- Images: Google Maps® images of the project location in active mode so that users can see the project area and zoom in and out to gain street-view perspectives on the location, setting, and other information that may be relevant to their assessment of the case.

Step 4 involved the development of data for each of the 100 cases, analysis of the database to extract trends, relationships and findings supporting the research project, and the production of the working papers for the project, which appear on the TPICS website and are available in preliminary form on the SHRP2 Capacity Program website. Development of the cases was tied to the case-potential screening carried out in Step 2. Ten projects were selected as test cases. These projects were used to co-ordinate case development methods, resolve consistency issues that arose as cases were being developed, and as a mechanism for exchange of techniques, successes and failures in gathering data for cases. Regular meetings with the teams developing the cases provided an exchange of ideas and insights into how to address several of the problems in deriving pre-construction information for cases – some of which had been constructed 10 to 20 years earlier. Results of this collaboration on the pilot cases, and subsequent work on the remaining 90 cases produced a set of insights into ex-post case development that are currently being integrated into an online course for case developers. The online course will be offered to those wishing to contribute new cases to TPICS.

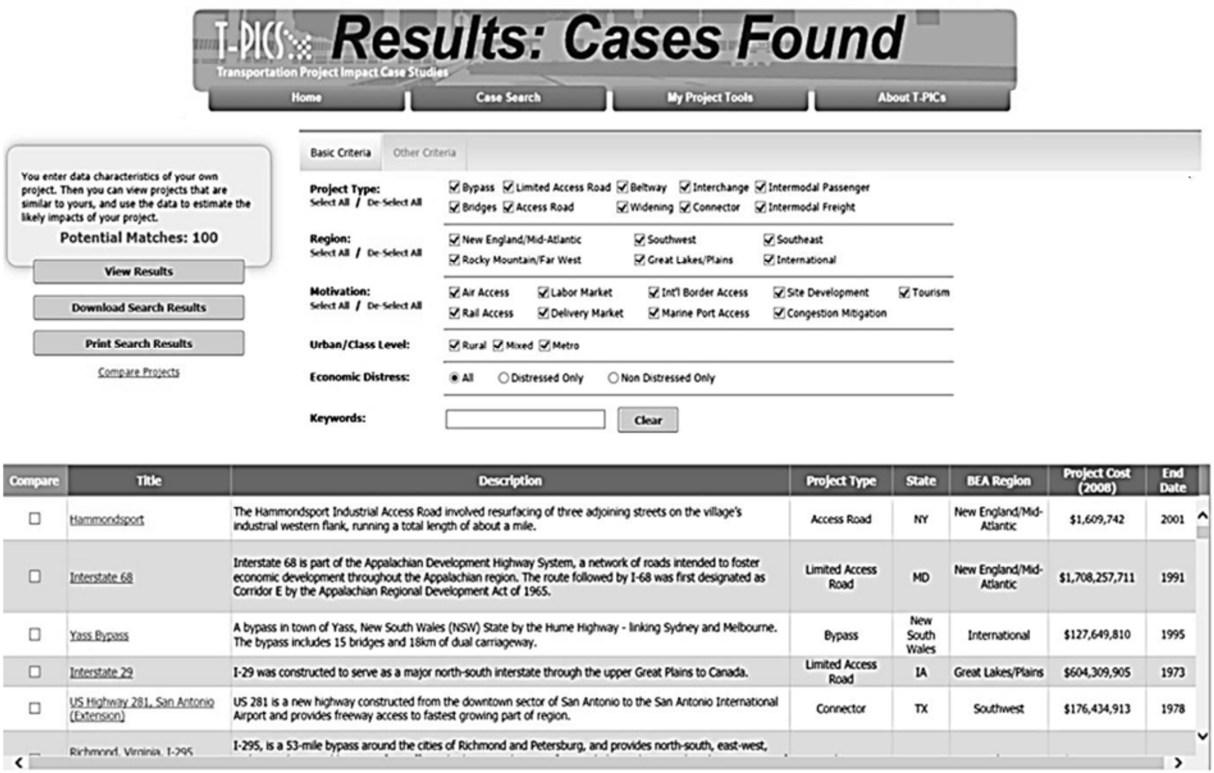
Work on Step 5 produced the final design and population of the database for the TPICS website and associated documentation. In addition to the technical documentation, design guides for the website and the 100-case database were prepared, as well as a user's guide describing how to use the website and a handbook for practitioners, designed to help users interpret and frame results from the cases effectively and as intended by the developers (e.g. as a preliminary screening tool and not as a final analysis of economic impacts).

Figure 4.1. **Introductory TPICS screen – case search, my tools and resources**



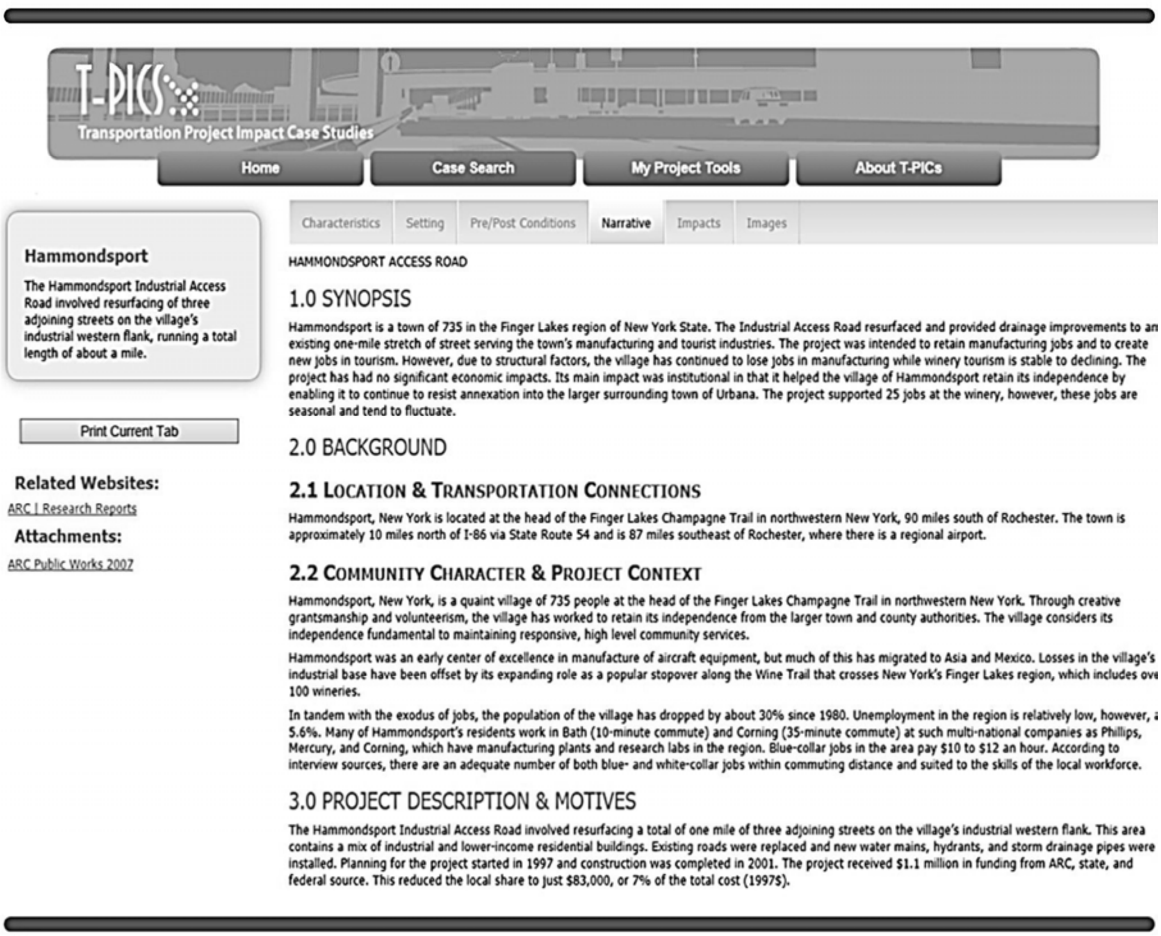
Source: TPICS website (www.tpics.us).

Figure 4.2. **Results of TPICS case search**



Source: TPICS website (www.tpics.us).

Figure 4.3. **Sample of detailed TPICS case information**



Source: TPICS website (www.tpics.us).

Characteristics of TPICS case studies

The TPICS research and deliverables were focused on US highway projects (although a few international projects were included to test the validity of the approach using data from outside the US). As described above, the total number of cases was eventually set at 100, with 11 project types and coverage for most of the broadly defined regions of the United States. Selecting the cases and designing the data and analytic systems involved a number of choices and careful balancing of the needs of the research versus the available cases and data. Some of the key characteristics of the cases and considerations related to selecting specific locations and data are described below, and include:

- project types and locations
- data collected
- project impact measures.

Project types and locations

The case study dataset was designed to cover the full range of highway-related facilities, including: intercity highways, urban beltways and local access roads, as well as local bridge and interchange projects. In addition, highway/rail projects were included, covering both intermodal freight terminals and intermodal transit terminals. Cases were selected to encompass a wide range of project types, spanning different regions of the United States and different types of urban/rural settings and economic distress levels. The initial 100 cases were distributed as shown in Table 4.1.

The selected projects represented capital investments intended to either enhance access to locations (via new routes and intermodal facilities) or expand effective capacity where it has been adversely affected by congestion or sub-standard operating conditions (via added lanes, interchanges, bypasses or intermodal facilities). These are the types of projects that have claimed economic development objectives. Other types of highway investment, such as safety projects, were not covered in the case study database because there appeared to be no point in including projects that were never intended to even have any economic development impact.

Table 4.1. **Geographic distribution of TPICS project types**

Project Type	United States Region						
	Great Lakes/ Plains	New England /Mid-Atlantic	Rocky Mtns/ Far West	South-east	South-west	Inter-national	Total
Industrial Access Road	2	2		2	1		7
Beltway	2	1	1	2	2		8
Bridge	1	2	3	2	1	1	10
Bypass	4	1	3	2	1	2	13
Connector	1	1	2	3	1		8
Interchange	4	2	1	2	3		12
Intermodal Freight Terminal	2	2	1	3	2		10
Intermodal Passenger Terminal	2	1	3	2	1		9
Major Highway (Limited Access Road)	3	4	1	4	2		14
Widening	1	1	2	3	2		9
Total	22	17	17	25	16	3	100

Source: EDR Group et al., 2011.

Table 4.2. **Distribution of settings for TPICS projects by project type**

Project Type	Metro	Mixed	Rural	Total
Industrial Access Road	2		5	7
Beltway	8			8
Bridge	4	3	3	10
Bypass	4	1	8	13
Connector	4	2	2	8
Interchange	10	2		12

Project Type	Metro	Mixed	Rural	Total
Intermodal Freight Terminal	6	1	3	10
Intermodal Passenger Terminal	9			9
Major Highway (Limited Access Road)	5	9		14
Widening	4	3	2	9
Total	56	21	23	100

Source: EDR Group *et al.,* 2011.

Data collected

For each of the 100 selected projects, data was assembled to facilitate:

- comparison of pre-project and post-project changes in economic and land development conditions
- contrast of observed project area changes with underlying state and national population and economic growth pattern trends occurring over that same period; and
- inclusion of both quantitative impact measures derived from available public sources and qualitative assessments derived from local interviews. Five categories of data were assembled for each case study:
  - project characteristics – type of transportation facility, years built, cost, size (length, lanes), and level of use (AADT, VMT)
  - project objectives – congestion reduction or access enhancement
  - pre/post change measures – employment, population, land values, building development
  - settings – region of the United States, population density, urban/rural class, topography, economic distress, market size, distance to key destinations
  - local data from interviews – land-use regulations, use of business incentives; presence and use of support programmes for economic development, other local factors enhancing or reducing observed economic changes.

Project impact metrics

Economic impacts of transportation facilities typically unfold in a sequence, affecting different impact metrics and spatial scales over time, as noted in the FHWA guide to case-study measurement of economic impacts (EDR Group and Cambridge Systematics, 2001). Acknowledging these effects, the SHRP case studies (completed in 2010) were restricted to projects that had been completed at least five years earlier. In addition, the case studies sought to measure land value and building construction effects at the level of highly localised areas, while employment, income and tax impacts were measured for both local areas and larger areas (ranging from individual municipalities to multi-jurisdictional corridors or counties). The case studies confirmed the following typical sequence of impacts:

- **Transportation impact**: Initially, a highway project is instigated to affect travel-related costs or accessibility for some areas by enabling faster or more reliable travel to and from that area, or enabling access to a broader set of origin or destination opportunities. The benefitting area may be adjacent to the project, or it may include areas well beyond the endpoints of the project corridor. There are occasionally adverse impacts on adjacent areas, which tend to be offset by benefits elsewhere.
- **Land (property) value impact**: Upon project completion, or in anticipation of it, demand starts to grow for land at benefitting locations, typically leading to higher property values and transaction prices there.
- **Building construction and investment impact**: Increased demand leads to added investment in the form of building construction. That effect is reflected initially in terms of building permits and later in terms of new or upgraded building structures.
- **Employment, income and output impacts**: Once buildings are occupied, there can be measurable increases in population or business activity. The latter can be measured in terms of added jobs, income, value added or output growth.
- **Tax revenue impacts:** The added land value, building structures, population and business activity together can show up as increases in property, income and/or sales tax collections.

The case studies confirmed two key conclusions pertaining to this list of impact measures. First, impacts unfold over time, so no single project will necessarily show every type of impact at the same time. For that reason, multiple impact measures and an appropriately broad period of observation may be needed to adequately characterise economic development impacts. Second, each of the various forms of impact can have a different spatial pattern of observation. Some may be observed at a neighbourhood level while others will be spread over a broader community or regional level. These effects also vary systematically by type of project. For instance, connectors, access roads and interchanges tend to have localised impacts, while intercity routes and bypass projects can have broader impacts with some beneficiaries hundreds of miles away.

Table 4.3 shows the extent to which each of the major categories of impacts was observed or measured in the case study process. All the 100 projects studied had some measurement of its economic development impact. In general, employment change was the measure most easily obtained because of widely accessible datasets on annual employment changes, available at the county, community and even postcode levels across the United States. For this study, the measure of employment change reported as a highway impact was defined to be whatever level of geography was deemed most relevant for that project, which varied from local community to multi-county region, depending on the project location or corridor length. Actual impacts associated with the project were determined through interviews, field investigations and, where available, consultation of building permits and records.

Quantitative metrics on building permits (square feet), property transactions (sales levels), private investment (value of new development) and incremental tax revenues generated were more difficult to obtain because they typically came from municipal or county records. These records differ widely in their format, availability and usability for time-series analysis. The most problematic metric was the tracking of property value changes over time, due to differences in local valuation rates and updating policies. For that reason, some case studies relied on "qualitative" data sources (i.e. impact estimates reported by local planners and business leaders) for estimates of local changes in property values, building investment and/or property sales patterns.

Table 4.3. **Direct observations of impact measures in TPICS cases**

Measure of impact	Number of cases in TPICS database		
	Qualitative: change observed	Some quantitative impact data	Full quantitative data in dataset
Employment	100	100	100
Income	*	*	*
Business value added or GDP	*	*	*
Building development (Sq. Ft.)	74	38	36
Direct private investment ($)	57	27	30
Property values	36	30	6
Property tax revenue	50	36	14

Note: Measures that were calculated (in the database) from employment change ratios.
Source: EDR Group et al., 2011.

Because the projects differed dramatically in their size (ranging from a single bridge or intersection to a 200-mile highway corridor), the corresponding economic development impacts also varied widely. The net employment impact per project averaged 1 355 but ranged from -10 to +93 000. The building development impact ranged from 4 200 to 50 million square feet, private investment over time ranged from USD 3 million to USD 6.3 billion, and annual property tax impact ranged from nil to USD 55 million.

Summary of key findings from case studies

Individual case studies yielded a wide array of impacts and a rich set of narrative findings. The field investigations and interviews supplemented published data and provided insights into the mechanisms by which investments in highway projects contribute to economic development and overall economic growth. However, the methods used in the research conducted for this study also allowed for additional investigations into areas not usually explored when a small number of ex-post studies are conducted. Although the research published in conjunction with this study goes into great detail about the majority of the findings tied to these cases, the following section highlights some of the issues that should be explored in greater depth as new cases are developed and as further research into ex-post analysis is conducted. These areas include:

- the role of project motivation and purpose
- the role of non-transportation factors
- projects with no apparent net employment effects.

The role of project motivation and purpose

Project motivations were classified into nine major categories. Eight are related to economic development via enhancement of various types of access. They include: improving access to terminals of air, rail, and marine modes, international borders, labour markets, delivery markets, tourism markets, or facilitating on-site development. The ninth is congestion management, which often represents an attempt to prevent further degradation in conditions rather than to enable positive enhancement compared to past or current conditions.

In the case study interviews for each project, both local planning officials and business representatives were asked to identify project motivations (e.g. reasons why the project was built). They were allowed to choose more than one reason. Table 4.4 shows the range of motivations by highway, freight intermodal and passenger intermodal projects. Overall, project motivation was obtained for 77 projects, of which 66 were motivated by at least one economic development factor and 11 were motivated by congestion management alone. The motivation to mitigate congestion was most often reported for urban highway projects, while the motivation to facilitate site development was most often reported for interchange and access road projects.

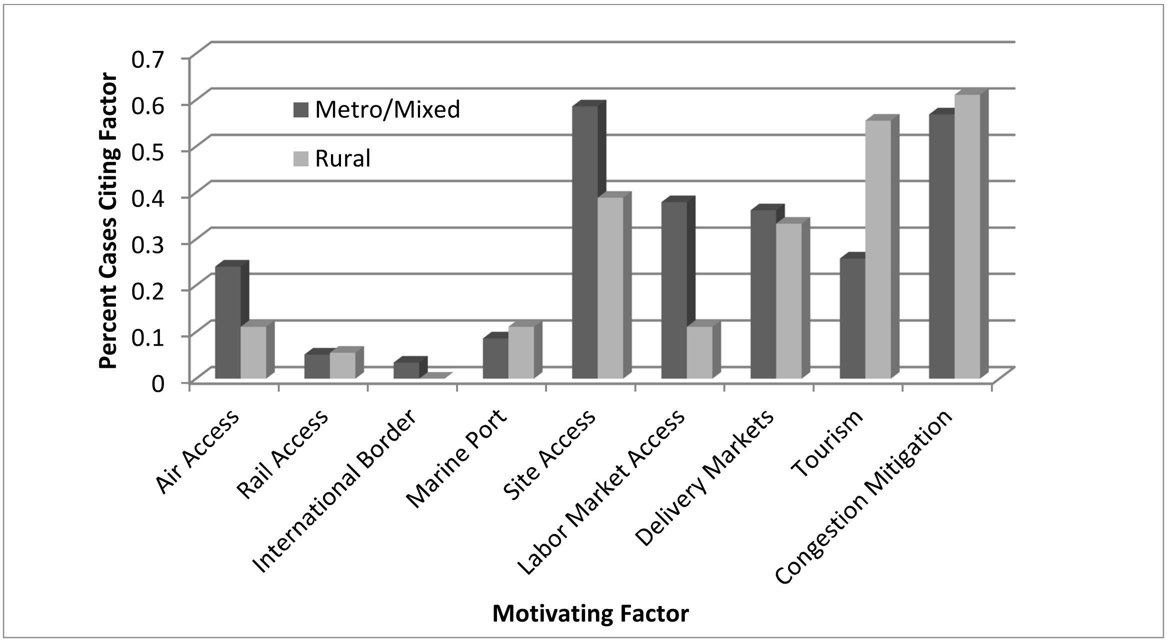
Table 4.4. **Motivation reported by project category**

Project motivation	Highway	Freight intermodal	Passenger intermodal
Improve access to airports	18	2	0
Improve access to rail	4	6	0
Improve access to international border	2	1	0
Improve access to marine port	7	2	0
Facilitate site development	42	2	8
Improve labour market access	26	0	4
Improve delivery market access	29	3	0
Facilitate tourism	26	0	0
Mitigate congestion	46	0	7
All projects reporting motivation	**58**	**10**	**9**
Total number of projects by type	**81**	**10**	**9**

Note: Many projects had more than one motivation, so the column totals in this table will not sum to the total for all projects reporting a particular reason motivating the project.
Source: EDR Group et al., 2011.

Figure 4.4 shows how project motivation varied by setting. This figure shows that the most common motivation for projects in both rural and metro areas was congestion mitigation. After congestion mitigation, site access and delivery market access were two most frequently cited motivating factors in both metro/mixed and rural settings. Tourism was an important motivator in rural areas and labour market access was a key reason motivating projects in metro/mixed areas.

Figure 4.4. **Project motivation by setting**



Note: Projects in metropolitan areas and projects in areas where there is a transition from metropolitan to rural (mixed) are grouped together for this chart.
Source: EDR Group et al., 2011.

The role of non-transportation factors

Evidence of economic development for many of the project types was attributable to leveraging project investments with other infrastructure investments, land use policies or business development incentive programmes. In some cases, effective synergy among multiple factors created a positive economic development climate that led to job creation or other economic development consequences. Yet in other cases, a lack of complementary infrastructure and supportive development policies led to disappointing economic development results. Table 4.5 shows the frequency with which these non-transportation factors affecting the long-term economic development impacts were cited in case study interviews.

Table 4.5. **Non-transportation factors that influenced economic growth**

Policy factors	Factor	Number reported
Effective synergies	Infrastructure (sewer, water, broad band, transit, etc.)–positive	33
	Land use management–positive	45
	Financial incentives/ business climate–positive	47
Lack of effective synergies	Financial incentives/ business climate–negative	5
	Infrastructure (sewer, water, broad band, transit, etc.)–negative	10
	Land use management – negative	6

Source: EDR Group *et al.,* 2011.

Projects with no apparent net employment effect

Developers of the case studies found that 15 of the 100 projects led to a zero or negative impact on job growth and economic development. Nearly all were bridges, bypasses, connectors, interchanges or passenger/freight terminals. With the possible exception of intermodal projects, these were generally projects designed more to help manage traffic flow than to generate economic growth. The finding for rural community bypass roads was also to be expected. Past bypass studies conducted for a number of different states have shown that economic impacts are either only slightly positive or negligible in most bypassed communities. That outcome is due to the offsetting positive and negative effects of shifting pass-by traffic out of local communities, which represents a potential loss for some traffic-dependent businesses, but a potential gain for others that benefit from having improved safety and a more attractive urban environment for local residents and visitors.

An important finding was that most of these 15 projects had other forms of positive economic impact despite the lack of positive employment impact. This included the following findings:

- 8 of the cases had gains in post-project business sales at the county level
- 10 of the cases had growth in local per capita income after project completion
- 6 of the cases had documented increases in local property values.

Issues in applying ex-post analysis lessons from the TPICS experience

Measurement issues

One of the most challenging aspects of casting a wide net in evaluating ex-post impacts for a variety of project types, vintages and locations – even though restricted to highway-oriented projects – is the development of consistent measures that allow for rigorous meta-analysis of case studies. Working with a team of contractors, even those with a great deal of knowledge and background in transportation and case study development, highlighted some of the measurement issues involved in case development. Development of TPICS also highlighted some of the issues related to developing data of the quality and consistency suitable for detailed ex-post research.

Addressing differences in transport mode and project scale

One of the challenges being faced in the development of TPICS has been the issue of how to span across modal and project scale differences. Among highway projects, there are major differences in the scale of impacts between interchange and access road projects that tend to have highly localised development impacts, and major highway (trunk road) projects that can have broad-based economic development impacts spanning large regions. While the latter can generate larger total impacts on jobs and income, those impacts can be widely dispersed and, therefore, hard to distinguish from underlying economic change patterns. Thus, care is needed to identify the appropriate study area that can enable measurement of post-project impacts. On the other hand, the localised projects can have highly visible impacts on land and building development, but data to measure net job impacts can be more difficult to acquire. The TPICS system thus allows for individual project cases to involve a wide range of economic impact metrics to capture changes in land values, building activity and/or employment-related measures. In that system, it is left up to the discretion of researchers to select the most appropriate study area and most appropriate impact measures for each case.

Expansion of project types to include non-highway modes

The expansion of TPICS mid-way through the development of the database to also include public transport projects raised new measurement issues because many of these projects affect either very local areas (in the case of bus and light rail stops) or specialised segments of the economy (in the case of commuter rail lines that affect labour market access for specific destinations). In addition, the sources of information are different, as the economic development impact of public transport facilities may be affected by characteristics such as the time period and frequency of service; factors that do not arise when considering highway impacts. Therefore, the public transport case studies of necessity had to be conducted with significant input from public transport operators.

Multimodal/intermodal effects

Measurement issues also become more complex to document when projects have multi-modal consequences; either because the projects lead to mode switching in a corridor or region, or because the projects are intermodal terminals. Projects involving public transport stations or terminals (whether for bus, rail, air or marine modes) are more likely to have land development impacts on the immediately surrounding area, particularly when developed as part of a mixed-use development project. Corridor projects are more likely to have broader area consequences, though that can depend on how regional market access and intermodal connectivity change for each affected mode. Freight intermodal projects can affect the supply chain of businesses that are far removed from the facility, in other states or metropolitan areas.

Connection of economic development impacts to wider economic benefits

To further understand how market access and connectivity changes can lead to economic development impacts, the SHRP2 Program funded further research to develop tools for identifying wider economic benefits of transport projects, and then tested those tools with TPICS case studies to see how they relate to actual ex-post economic development impacts (EDR, 2013). The new tools are essentially spreadsheet models designed to make it easier for planners and evaluators to measure changes in transport factors that have consequences beyond existing travellers and, thus, create opportunities for wider economic development impacts. They include: changes in labour market access, same-day travel access delivery, intermodal connectivity (highway-rail, highway-air and highway-seaport) and reliability for just-in-time business processes. These tools have been completed and tested, and are now being applied for selected TPICS pilot studies. This line of research may be of particular interest because it offers the potential to document specific drivers of wider economic benefits that are not well captured by "willingness-to-pay" studies. In addition, this on-going line of research provides a basis for demonstrating how benefits that are beyond traveller impacts lead to observable economic development impacts in a defined region.

Measurement of construction impacts

Ex-post measurement of economic development impacts usually focuses on long-term economic growth because most transport investments are motivated by long-term objectives. However, there are exceptions. Several recent programmes for transport funding in the United States have sought to stimulate economic growth in times of economic downturn. The largest was the American Recovery and Reinvestment Act (ARRA) of 2009. This programme, funded to help spur jobs and wage growth during a recession, subsequently funded nearly 13 000 highway projects and roughly 1 070 public transport project grants in a four-year period. Another was the TIGER (Transportation Investment Generating Economic Recovery) grant program.

Both the ARRA and TIGER programs were designed with provisions to ensure accountability and transparency. For the ARRA program, most attention was focused on the documentation of (short-term) construction job generation that was created by the funding of project construction activity. Each of the country's fifty states was required to collect, compile and submit to the Federal Government both monthly and quarterly reports on directly-supported construction jobs. A study of the economic impact of ARRA-funded transportation projects reviewed the data submitted by selected states. It found that the states were struggling to document construction job generation because their surveys of directly-contracted construction companies provided complete data only for prime contractors, while information from subcontractors was often incomplete and sparse. As a result the information being reported was inconsistent. Despite these problems, there were some general findings: a national average of 10.55 direct jobs per million dollars, which ranged from 9.01 to 17.03 depending on the type of project and its setting (i.e. explanatory factors such as wage rates, topography and congestion levels).

This line of ex-post research – documenting construction impacts – is of note because it serves a policy purpose totally unrelated to benefit-cost analysis (BCA). In BCA, project expenditures and the associated construction process is all part of the cost and cannot be recast as a benefit. Yet from a public policy perspective, the generation of short-term construction jobs was very much a motivation for the programme funding, and a case had been made that spurring local construction jobs would do more to jump-start local economic growth than would the alternative of refunding money to residents so that they could increase purchases of consumer products.

Recognising urban land redevelopment impacts for "mega-projects"

The Central Artery/Tunnel Project in Boston (USA) was the most costly mega project in US history. It represents an example illustrating the complexity of considering costs, benefits and economic development impacts in an ex-post study. The project involved:

- Demolishing a major elevated interstate highway (I-93) through Boston's central core and rebuilding it underground.
- Extending another existing highway (I-90) via an underground route through the old seaport (docklands warehouse) district.
- Building a new third tunnel under Boston Harbor to the airport (EDR Group, 2006a; EDR Group, 2006b).

The selected option of building all of these facilities as underground tunnels was not the lowest cost option to meet traffic flow requirements. However, that approach was chosen because of a desire by public officials to improve the region's visual and environmental quality, reknit the urban fabric of vehicle and pedestrian routes that had been split by the elevated viaduct, and facilitate new development on the eastern side of the Central Artery corridor and in the seaport district – areas that had been isolated by the existing elevated highway structure.

Ex-post studies that were carried out after completion of the project found that:

- The final project costs (USD 15 billion) were triple the initial (USD 5 billion) estimates. The reasons for this difference have been debated, but essentially come down to some combination of: (a) lack of allowance to anticipate particularly costly difficulties encountered in the complex underground construction process; (b) lack of allowance for costly changes and additions to the project specification that were added to address later stakeholder concerns; and (c) quality of management oversight.

- The user benefits were fully achieved and in fact were running about 15-20% greater than anticipated. The reason for this difference is that regional traffic growth has actually been occurring at a slightly lesser rate than originally anticipated. This is reflective of slowing national trends of growth in vehicle-miles travelled.
- The urban redevelopment benefits have been more than ten times greater than the level that was originally anticipated. The original pre-project economic study had concluded that Boston's downtown core was largely constrained from further population and employment growth because of the capacity limitation of the Central Artery highway. That study had anticipated that the completed project would enable more residential and office development, particularly in areas that had been isolated because of the highway structure - and that was correct. Yet, the actual result in terms of residential and office development was far higher than originally estimated. The reason for this difference may be due in part to the unanticipated growth of Boston's role as a world biotechnology leader, and the unanticipated shift in lifestyle decisions of the new millennial generation that are strongly opting for urban rather than suburban locations for their work and home.

The Central Artery/Tunnel Project became infamous in its early years as a failure because of the massive cost overruns, but in recent years a local consensus has emerged that the project also had a massively under-estimated value in creating growth of a vibrant urban core, while also helping to reduce suburban sprawl. Traditional benefit-cost analysis is poorly equipped to incorporate these kinds of structural changes and regional economic development effects.

Case development issues

Experience in developing cases designed to be used in ex-post analysis led the project team into several areas that should receive considerable thought, given the plans for new rounds of grants under the TIGER program, studies and expansion of the TPICS database, and similar efforts now under consideration by US governmental agencies and organisations. Should elements of the US approach be of interest to other countries, some of the methods developed in this project, and many of the insights and lessons learned may prove useful when developing their own research and development programmes.

Methodology

Even with a highly skilled team, the process of looking back in time and the unique circumstances of several of the projects encountered in this study required careful consideration of ways in which historical data needed to be developed. The most common methodological issues arose around ways in which difficult-to-obtain historical data could be reconstructed and sorting out background cyclical effects from project-related effects; especially for projects completed before many of the digitised land-use and traffic-data systems we currently employ were widely available. Guidance cited previously on the use of interviews, building permits and records, and local knowledge were helpful, but often required skilful interview techniques on the part of the case developer and cross-checking between developers to compare the kinds of response being received for similarly situated cases and case types. This was especially important given the intended use of the case data in a large-scale meta-analysis.

Consistency

As the project developed, it was important to maintain regular contact amongst the case developers to ensure that guidance was being followed and the various types of problem resolution approaches were being shared. As the project progressed, new guidelines were developed and prepared that expanded on previously published guidance. These guides and "tips" have been compiled into a multi-module course on case study development and are expected to be used as part of the on-going development of new cases.

As noted above in the measurement issues section, there are a number of scale, modal and impact measurement issues that were noted during the development and review of the case studies. At times, some of the issues confronted by the case developers (e.g. resolving inconsistencies between data sources and reconciling information obtained in interviews that differ from available data sources) required close consultation between case developers and project managers. The key was regular communications and sharing of issues at the time they arose rather than after a case was prepared and submitted for review. The high degree of consistency and review used in developing the cases contributed to the utility of the database when used for further research, meta-analysis and design of associated tools.

Identifying the correct counter-factual case

An integral aspect of ex-post analysis is setting a reasonable "base case" or "counter-factual" scenario, against which post-project economic conditions can be compared. There are two critical dimensions to consider

- Was the project's primary objective (motivation, as described in the TPICS cases):
  - to increase access through improvements enabling shorter distance, higher speed, more frequent service and/or cheaper travel costs between certain points (as provided by a new highway route or transit service)
  - to maintain an existing facility or service so that it does not degrade or close (as provided by reconstruction or replacement of facilities that face physical or functional obsolescence)
  - to address growing congestion/bottleneck conditions that increasingly constrain capacity and reduce travel time and its reliability (as provided by capacity expansion on existing roads or transit line facilities).
- Are travel demand and network models available to analyse travel volumes, speeds and conditions under three scenarios:
  - Scenario A: pre-project conditions
  - Scenario B: post-conditions for the analysis year(s) including construction of the project
  - Scenario C: post-project period for the analysis year(s) assuming that the project had not been built.

These dimensions are critical because they affect counter-factual expectations regarding transportation facility or service characteristics (in terms of capacity, distances, speed capabilities, etc.) and the ways that these expectations need to be addressed in evaluation. If the project was intended to improve access or connectivity, then we would expect the counter-factual case (Scenario C above) to be represented by continuation of pre-project (non-improvement) network characteristics (Scenario A). If the project was intended to prevent further degradation of a facility or service, then the counter-factual case for Scenario C would diminish transportation network performance. However, if the project's primary objective was to mitigate congestion impacts then a special case of counter-factual conditions would apply – a scenario dependent on engineering or transport network models reflecting the effects of increased congestion (reduced speeds) on vehicle flow (reduced through-put) without mitigation measures provided by the project.

The reason for special treatment of congestion-motivated projects is the existence of a "backward-bending supply curve" for speed/flow relationships under congested conditions. This is a phenomenon of traffic flow whereby the volume (as measured by vehicle miles travelled – VMT) that can move through a facility actually diminishes as congestion increases (Walters, 1961; May et al., 2000). More commonly, a congested road with start-stop conditions (occurring when the volume/capacity ratio exceeds 0.8) may appear to be operating near capacity but may actually be moving vehicles at less than half the speed of a road that appears half empty but has traffic moving at near free-flow speeds (such as 80 km/h or 50 mph). As a result of this congestion phenomenon, there are actually two levels of travel demand that can produce the same VMT in a given time period; one with a lower vehicle volume (vehicles/hour) but higher speed [miles (km)/hour], and one with a higher vehicle volume but lower speed. Yet while the VMT appears the same, one supports a higher level of economic activity than the other (it is worth noting that this same phenomenon can apply for passengers/hour using transit facilities).

Traffic engineers and travel modellers have recognised this phenomenon in long-range studies for both regional plans and individual road links. Traffic models for regions that are economically growing often forecast that the baseline growth in trips will, over time, lead to increasing delays and reduced throughput for major highway facilities (as congestion reduces their functional capacity). They have also forecast that expansion of the physical capacity in highways and transit facilities will enable more growth in user activity and lessen future degradation of traffic conditions, but not necessarily keep facility performance up to pre-project conditions.

A case in point, which is typical for urban transportation plans in many growing metropolitan areas, is the Portland Metro Regional Long-Range Plan (EDR Group, 2005). In that case, regional transportation and economic models were applied to evaluate two policy alternatives for increasing long-term spending on transportation infrastructure. The alternatives included both highway expansion and rail transit investments to increase regional capacity for growth. Under all future scenarios, traffic delays were forecast to worsen compared to conditions at the time of the study, yet the high investment scenario was shown to be beneficial because it would lead to less degradation in the level of service (delay conditions) than the lower investment scenario.

An individual project example that was included in the TPICS case studies and has been well-documented is Boston's Central Artery/Tunnel Project described above. An ex-ante study relied on transportation and economic models to forecast that much of the project's impact would be avoidance of a major regional bottleneck, which would otherwise stunt growth of employment and investment in the city's central core (MA DPW, 1990). Ex-post evaluation later measured actual pre/post changes in traffic speeds, volumes and economic activity compared to what had been originally forecast in the ex-ante study and noted the factors that led to underestimation of the economic effects of investments in congestion mitigation (EDR Group, 2006a).

Analysis of the TPICS case study database showed a related issue concerning ex-post evaluation of congestion mitigation projects. It found that projects intended to improve local access and connectivity often lead to observable changes in land values and business investment adjacent to the project area.

However, it found this is less likely for projects intended to mitigate congestion because they tend to affect trip ends that are more spatially diffused and farther away from the project area (EDR Group, 2012). For major highway or rail bottleneck projects, the beneficiaries may even be outside of the metropolitan area.

Thus, when conducting ex-post analysis, care must be taken to: (1) distinguish project intent between projects that are intended to improve conditions and projects that are merely intended to reduce degradation of conditions; and (2) distinguish impact areas between projects that are intended to have local benefits and projects that are intended to affect broader trade or inter-regional movements. To appropriately determine project impacts, the counter-factual (base case) scenario must be defined to reflect those differences.

Approach to TPICS implementation

The TPICS program has been selected as one of the products developed under the SHRP2 research program to advance to implementation on a wider scale. This involves transferring the data, tools and systems developed through the SHRP2 TPICS program to the joint sponsorship of the US DOT's Federal Highway Administration (FHWA) and AASHTO (FHWA, 2014). This process, including co-ordination among all three organisations, is now underway. A series of meetings have already been held and a draft implementation plan has been drawn up for review and funding. Several aspects of the implantation programme are of interest given the issues raised during the course of TPICS development.

Pilot studies

A series of pilot studies have been funded by FHWA to permit states and metropolitan governments to evaluate TPICS as a planning tool, and to begin the process of identifying projects that they are interested in developing as new cases for TPICS. Both of these objectives are vital to the continued success of the programme. TPICS was designed specifically to supplement the early-stage planning process as a way to illustrate how past projects have affected the economy of local and, depending on the size of the project, multi-county regions or states. Testing TPICS using the wide variety of current planning processes will help to identify the practical strengths and weaknesses of the system, the cases and the interface.

The second objective of the pilot programme is to test the feasibility of adding new cases to the TPICS database. As noted earlier, once a variety of project types, regions and settings have been identified, the number of cases suitable to frame the range of possible outcomes – even with 100 cases available – can be very "thin" (Table 4.1). This was evident when the meta-analysis was conducted. Development of case information by a single agency would be cost-prohibitive if several hundred cases are to be added to the database. Therefore, encouraging various levels of government to validate the current methods and contribute new cases could be a very important step towards both new case development and refinement of the information drawn from the case study database.

The selected projects represented capital investments intended to either enhance access to locations (via new routes and intermodal facilities) or expand effective capacity where it has been adversely affected by congestion or sub-standard operating conditions (via added lanes, interchanges, bypasses or intermodal facilities). These are the types of project that have claimed economic development objectives (other types of highway investment, such as safety projects, were not covered in the case-study database because there appeared to be no point in including projects that were never intended to have any economic development impact).

The pilot studies were recently awarded to the following organisations (SHRP2, 2014): Illinois DOT; Indiana DOT; Rhode Island Statewide Planning Program; and UTAH DOT.

Other implementation support

The SHRP2 Implementation Program will continue for several years as outreach and pilots implemented. The Implementation Plan being developed for the programme includes several key elements:

- Communications: promote wider practice and adoption of TPICS and other SHRP2 research products.
- Product messaging: focus on data-driven decision support and methods making use of ex-post analysis methods.
- High-level outreach: develop non-technical information for decision-makers to inform them of emerging technical advances.
- Planning resources: provide outreach to professional associations and research organisations to encourage technical exchanges and peer-level information sharing.

In addition, the SHRP2 Capacity Program is preparing several new case studies based on initial pilot testing of TPICS in Minnesota. The SHRP2 Capacity Program is also developing an online training course in conjunction with AASHTO. AASHTO is taking on the responsibility of hosting the TPICS website and providing technical co-ordination with its regional affiliates.

The Implementation Plan is also designing performance measures to assess the success of the introduction of TPICS and the tools developed as part of the programme. These measures include the following:

- number and variety of new case studies developed for inclusion in TPICS database
- number of in-person training sessions
- number of transportation investment decisions that result from use of the Economic Analysis Tools
- number of hits on website hosting the TPICS online database
- number of downloads of the Wider Economic Benefit (WEB) spreadsheet tools
- number of hits on web pages containing Economic Analysis Tools messages and information.

Potential for TIGER projects

The US DOT has established a tracking progress on all TIGER grants and is in the process of setting up a database that will include project and performance information. US DOT expects to begin funding these efforts in the current fiscal year (FY2015). They are planning to have on-going data management and tracking capabilities in place by mid-2015. Information supporting both the short-term and long-term employment effects resulting from expenditures through the TIGER awards is being required under grant agreements that have been individually negotiated with each grantee.

There are currently more than 270 projects that have received grants since the programme was first initiated in 2009. US DOT expects to award about USD 548 million in the current round of grants (applications that were submitted in late April 2014 and are due to be announced in September 2014) and are also expecting something in the order of USD 500 million to 550 million under the current authorisation. US DOT expects between 45 and 70 successful grants to be awarded from the current round, and that the relative number of awards will remain at about these levels as long as funding under this programme continues to be authorised.

This source of case studies and the potential for significant ex-post analysis of recently awarded TIGER projects provides an excellent source of information for developing information using the case study approach pioneered in TPICS. However, the wide range of project types funded by TIGER and the potential for collection of data by a wide range of agencies and staff with varying degrees of skill suggests that managing any meaningful process of data development may require extensive technical oversight and management.

If TIGER projects are to form a foundation for ex-post analysis, then many of the lessons learned from TPICS are crucial. These include; paying close attention to consistent case data development; training in developing case analysis for the organisations involved; clear specifications for data and its collection, oversight and quality control; and an active management plan that includes regular consultations between grantees and federal agencies managing data intake. Given the interest and expertise being developed as part of the TPICS Implementation Program by AASHTO and FHWA, it would seem advantageous to develop an active partnership between the FHWA's SHRP2 Solutions Program and the TIGER Management Program.

Conclusion and need for further research

The examples offered in this chapter illustrate the many ways that transportation investment can affect local and regional economic development, and the many ways in which transport projects have been conceived with the intent of supporting economic development outcomes. Transport investment can be motivated by both short-term and long-term economic impact goals, and by any combination of land-use, employment and income growth goals. Public policy decisions should be made in a way that fully recognises these types of goals and benefits in both ex-ante and ex-post studies. The interest of the planning community, the uptake by key federal agencies and transportation organisations, and continued interest by practitioners as well as state and regional planning organisations attests to the potential usefulness of ex-post analysis in supporting investment decision-making. Based on the research conducted to date and the directions in which implementation plans are heading, key elements of future research related to ex-post analysis include two key areas.

Data coverage

The individual case studies, available via the TPICS web tool (www.tpics.us), are notable for their standardised approach and attempt to isolate the incremental economic development impacts of highways. This latter objective was addressed in three main ways: (1) through inclusion of ex-post comparison of change along multiple impact metrics; (2) through inclusion of comparison of changes to reference areas (to control for external business cycles); and (3) through inclusion of interviews with local planners and business representatives (to assess the extent to which observed changes were due to the highway project vs. other factors).

However, the information in several of the case studies was limited because of the unavailability of pre/post project data on changes in traffic conditions over time. In nearly all cases, neither local nor state transportation agencies had pre-project data on traffic volumes, speeds or access conditions (the database does include post-project traffic volumes and access measures, though even pre-project speed data was limited). Since traffic congestion, performance and access conditions are frequently noted as project objectives, it is hoped that transportation agencies will be able to more systematically collect and retain measures of these conditions in the future, so that later case studies can be added with more complete pre/post transportation data.

Use of results

The case study database and analysis methodology have value in two ways. First, they provide a basis for distinguishing the extent to which the highway project was actually responsible for observed economic development impacts. Second, they serve to highlight the ways in which local economic and institutional factors served to either mute (reduce) or amplify (expand) the magnitude of observed economic development impacts. Thus, the case studies do help to establish the extent of causal connection between highway-related improvements and resulting economic impacts, though they cannot yet relate the observed magnitude of economic impacts to the magnitude of pre/post change in transportation conditions.

The most obvious application of the current case study database and TPICS web tool is to provide transportation planners with a way to search for relevant types of projects in specific settings (region location, urban/rural population density, etc.). It also allows users an option to specify a given type of proposed project and then see the range of impacts that have been actually observed in case studies to date. These features have three important uses. First, they can have value for early stage policy or strategy development, in which it can be useful to initially identify the magnitude and types of impact trade-offs to be considered. Second, they can be useful for early stage "sketch planning" processes where they may provide examples of the types of local barriers and success factors that will need to be addressed in later, more detailed planning steps. And third, the case study findings can be useful in educating elected officials, and in public hearings, as they provide a way of responding to the sometimes overly-optimistic hopes of proponents or fears of opponents, with information on the range of impacts that have actually occurred in the real world.

The case study results can also be used to provide empirical evidence to help validate the reasonableness of predictions made by economic impact forecasting models for proposed future projects. Until now, there has been a paucity of such data available for validating predictive models. However, it should also be clear that the case study database and web tool alone cannot serve as a substitute for the detailed analyses incorporated into predictive economic impact models. While predictive economic impact models forecast shifts in economic growth resulting from complex interaction of changes in transportation conditions and changes in the underlying economy, the case studies lacked both the transportation change data and the statistical controls incorporated into such models. Consequently, it is useful to view the case study database as ideal for use as a sketch planning tool for initial planning, policy or strategy development, while economic impact models are designed to be most useful in later stages of planning and prioritisation, where more details are available on the nature of proposed projects and their expected transportation system impacts.

Notes

1. SHRP2 was authorised under the Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU) through federal fiscal year 2009 and extended through March 2015 under a series of continuing resolutions. The total research authorisation for all programmes under SHRP2 currently stands at USD 232.5 million.

References

Anderson, C. and D. Otto (1991), *The Economic Impact of Rural Highway Bypasses: Iowa and Minnesota Case Studies*, Iowa Department of Transportation, Office of Advanced Planning, Ames, Iowa.

Bizminer/Brandow Company and Economic Development Research Group (EDR Group) (2007), "Program Evaluation of the Appalachian Regional Commission's Infrastructure and Public Works Projects", Appalachian Regional Commission.

Blackburn, S. and J.W. Clay (1991), "Impacts of Highway Bypasses on Community Businesses", North Carolina Division ex-post of Community Assistance and the I-40 Steering Committee, November.

Braathen, S. and A. Hervik (1997), "Strait Crossings and Economic Development – Developing Economic Impact Assessment by Means of Ex-post Analyses", *Transport Policy,* Vol. 4, Issue 4, pp. 193-200.

Brandow Company and Economic Development Research Group (EDR Group) (2000), "Evaluation of the Appalachian Regional Commission's Infrastructure and Public Works Program Projects", Appalachian Regional Commission.

Burress, D. (1996), "Impacts of Highway Bypasses on Kansas Towns", University of Kansas for the Kansas Department of Transportation.

Chapulut, J.N., J.P. Taroux and E. Mange (2005), "The new ex-post evaluation methods for large projects in France", *Proceedings of ETC 2005*, Strasbourg, France – Research to Inform Decision-making in Transport Innovative Methods in Transport Analysis – Planning and Appraisal, Appraisal I.

Collins, M. and G. Weisbrod (2000), "Economic Impact of Freeway Bypass Routes in Medium Size Cities", http://www.edrgroup.com/pdf/Urban-Freeway-Bypass-Case-Studies.pdf

Economic Development Research Group (EDR Group) et al. (2013), "Development of Tools for Assessing Wider Economic Benefits of Transportation (SHRP 2 Project C11)", Strategic Highway Research Program 2 (SHRP2), http://www.trb.org/StrategicHighwayResearchProgram2SHRP2/EconomicImpactAnalysis.aspx

Economic Development Research Group (EDR Group) et al. (2012), "Economic Impact Analysis, Interactions between Transportation Capacity, Economic Systems, and Land Use (SHRP 2 Report S2-C03-RR-1)", Strategic Highway Research Program 2 (SHRP2), http://www.trb.org/StrategicHighwayResearchProgram2SHRP2/EconomicImpactAnalysis.aspx

Economic Development Research Group (EDR Group) et al. (2011), "Highway Economic Impact Case Study Database and Analysis Findings. Interactions between Transportation Capacity, Economic Systems and Land Use: Final Report (SHRP2 Project C03)", Strategic Highway Research Program 2 (SHRP2), http://www.tpics.us/documents/SHRPC03FinalReport.pdf

Economic Development Research Group (EDR Group) (2006a), "Transportation Impacts of the Massachusetts Turnpike Authority and the Central Artery/Third Harbor Tunnel Project", Massachusetts Turnpike Authority, http://www.edrgroup.com/pdf/mta-economic-v1.pdf

Economic Development Research Group (EDR Group) (2006b), "Real Estate Impacts of the Massachusetts Turnpike Authority and the Central Artery/Third Harbor Tunnel Project", Massachusetts Turnpike Authority, http://www.edrgroup.com/pdf/mta-economic-v2.pdf

Economic Development Research Group (2005), "Cost of Congestion to the Economy of the Portland Region", Prepared for the Portland Business Alliance, Portland Metro and Port of Portland.

Economic Development Research Group (EDR Group) and Cambridge Systematics (2001), *Using Empirical Information to Measure the Economic Impact of Highway Investments*, Vols. 1 and 2, Federal Highway Administration (FHWA).

Economic Development Research Group (EDR Group) (2000), "Economic Impact of I-73 Alignments on the City of Roanoke", City of Roanoke, Office of Economic Development, http://www.edrgroup.com/pdf/Roanoke-I73-Final-Report.pdf

Florio, M. and S. Vignetti (2013), "The Use of Ex-post Cost-Benefit Analysis to Assess the Long-Term Effects of Major Infrastructure Projects", Centre for Industrial Studies, Working Paper Series, http://www.csilmilano.com/docs/WP2013_02.pdf.

Gillis, W.R. and K. Casavant (1994), "Lessons from Eastern Washington: State Route Mainstreets, Bypass Routes and Economic Development in Small Towns", EWITS Research Report #2, Department of Agricultural Economics, Washington State University.

Kjerkreit, A. (2008), "Post opening evaluation of road investment projects in Norway: how correct are the estimated future benefits?", *European Transport Conference 2008 Proceedings*.

Louwa, E., M. Leijtenb and E. Meijersa (2013), "Changes subsequent to infrastructure investments: Forecasts, expectations and ex-post situation", *Transport Policy*, Vol. 29, September, pp. 107-117.

Lynch, T. (2007), "Twin Counties Study", Chapter 3 in: *Sources of Growth in Non-Metro Appalachia*, Vol. 3, *Statistical Studies of Spatial Economic Relationships*, Appalachian Regional Commission.

Massachusetts Department of Public Works (MA DPW) (1990), "Supplemental Environmental Impact Statement/Report and Supplemental Final Evaluation", Section 4(f).

May, A.D., S.P. Shepherd and J.J. Bates (2000), "Supply Curves for Urban Road Networks", *Journal of Transport Economics and Policy*, Vol. 34, No. 3, September, pp. 261-290.

Miller, J. (1979), "Interstate Highways and Job Growth in Nonmetropolitan Areas", Venue: A Reassessment, *Transportation Journal,* Vol. 19, No. 1, pp. 78-81.

New York State Archives (2014), "1830s: The Erie Canal at Work", Archives Website.

Pfeiffer, D. (2006), "Ike's Interstates at 50: Anniversary of the Highway System Recalls Eisenhower's Role as Catalyst", *National Archives,* Vol. 38, No. 2, Washington, DC.

Rephann, T.J. and A.M. Isserman (1994), "New Highways as Economic Development Tools: An Evaluation using Quasi-Experimental Matching Methods", *Regional Science and Urban Economics,* Vol. 24, Issue 6, pp. 723-751.

Strategic Highway Research Program 2 (SHRP2) (2007), "SHRP2 Request for Proposals, Project C03, Interaction between Transportation Capacity, Economic Systems, and Land Use and Integrating Economic Considerations into Project Development".

Strategic Highway Research Program 2 (SHRP2) (2014), "Round 4 Implementation Assistance Opportunities", http://onlinepubs.trb.org/onlinepubs/shrp2/SHRP2_Round4_IAP_Awards.pdf

System Metrics Group et al. (2006), "California Bypass Study: The Economic Impacts of Bypasses", Vol. 1: *Planning Reference*, Caltrans, http://www.dot.ca.gov/hq/tpp/offices/eab/studies_files/Bypass_Final_Report_1_v3.pdf

Texas Transportation Institute (TTI) (1995), "Effects of Highway Bypasses on Rural Communities and Small Urban Areas", AASHTO.

US Department of Transportation (US DOT) (2014), "About TIGER Grants, Program Overview Website", http://www.dot.gov/tiger/about

US Department of Transportation (US DOT) (2013), "Moving Ahead for Progress in the 21st Century Act (MAP-21)", Program Overview Website, http://www.dot.gov/map21

US Federal Highway Administration (FHWA) (2014), "Easier-to-use tools for improved economic analysis: Economic Analysis Tools (C03/C11)", SHRP2 Solutions Program Webpage, http://www.fhwa.dot.gov/goshrp2/Solutions/Capacity/C03_C11/Economic_Analysis_Tools

US Federal Highway Administration (FHWA) (2004), "Economic Effects of Selected Rural Interstates at the County Level", http://www.fhwa.dot.gov/planning/economic_development/studies/

US Federal Highway Administration (FHWA) (2003), "Economic Impact Study of the I-86 Corridor", http://www.fhwa.dot.gov/planning/economic_development/studies/new_york_86/index.cfm

US Federal Highway Administration (FHWA) (2002), "Economic and Land Use Impacts Study of State Trunk Highway 29: Phase I Chippewa Falls to Abbotsford, Wisconsin", http://www.fhwa.dot.gov/planning/economic_development/studies/wisconsin_29/index.cfm

US Federal Highway Administration (FHWA) (1995), "Linking the Delta Region with the Nation and the World, December", http://isddc.dot.gov/OLPFiles/FHWA/015518.pdf

US General Accounting Office (GAO) (1996), "Economic Development: Limited Information Exists on the Impact of Assistance Provided by Three Agencies", GAO/RCED-96-103.

US General Accounting Office (GAO) (1991), "Designing Evaluations. Program Evaluation and Methodology Division", GAO/PEMD 10.4.1.

Walters, A.A. (1961), "The Theory and Measurement of the Private and Social Cost of Highway Congestion", *Econometrica*, Vol. 29, pp. 676-690.

Weisbrod, G. and D. Simmonds (2011), "Defining Economic Impact and Benefit Metrics from Multiple Perspectives: Lessons to Be Learned From Both Sides of the Atlantic", *Proceedings of the 2011 European Transport Conference,* Glasgow.

Weisbrod, G. (2000), "Current Practices for Assessing Economic Development Impacts from Transportation Investments", NCHRP Synthesis #290, Transportation Research Board, Washington, DC, pp. 41-42.

Wisconsin Department of Transportation (DOT) (1998), *The Economic Impacts of Highway Bypasses on Communities, Technical Report*, Economic Planning and Development.

Wray, S., S. Moses and G. Weisbrod (2000), "The Development Impacts of Highway Interchanges in Major Urban Areas: Case Study Findings", http://www.edrgroup.com/pdf/Hwy-Interchange-Case-Studies.pdf.

Chapter 5
Causal inference for ex-post evaluation of transport interventions

Daniel Graham - Imperial College, London, United Kingdom.

This chapter reviews methods that seek to draw causal inference from non-experimental data and shows how they can be applied to undertake ex-post evaluation of transport interventions. In particular, the chapter discusses the underlying principles of techniques for treatment effect estimation with non-randomly assigned treatments. The aim of these techniques is to quantify changes that have occurred due to explicit intervention (or "treatment").

The chapter argues that transport interventions are typically characterised by non-random assignment and that the key issues for successful ex-post evaluation involve identifying and adjusting for confounding factors. In contrast to conventional approaches for ex-ante appraisal, a major advantage of the statistical causal methods is that they can be applied without making strong a priori *theoretical assumptions.*

The chapter provides empirical examples of the use of causal techniques to evaluate road network capacity expansions in US cities and high-speed rail investments in Spain.

Introduction

Ex-ante transport appraisal has well-established theoretical and empirical roots in the consumer surplus-based calculation of Cost Benefit Analysis (CBA). Recent work on Wider Economic Benefits (WEBs) has extended "standard" CBA to incorporate some key externalities and forms of imperfect competition, again based on clearly set-out theoretical and empirical evidence. CBA provides a familiar and well-understood approach that is routinely used by civil servants, transport professionals and academics.

Considerably less attention has been paid to ex-post transport appraisal, both in theory and in practice. This is presumably because we are generally more interested in predicting how our future investments will fare than in assessing how well we have allocated resources in the past. Yet if we want to obtain a solid understanding of the impacts that transport interventions will have, a good benchmark can be established by studying previous interventions and how they performed according to some defined metrics of interest.

One way of doing this is to simply re-run the ex-ante CBA calculations some time after the project has been completed using observed rather than predicted values. This can provide useful information both on the impacts of the project itself and on how well ex-ante CBA was able to predict the benefits and costs of the scheme. Such calculations are, however, still generated within the theoretical framework of CBA, which makes a number of quite restrictive simplifying assumptions. An alternative approach is to apply statistical models to data observed before and after transport interventions and attempt to estimate impacts that were caused by the intervention. Such statistical approaches, which we refer to as methods for causal inference, rely more on empirical methods than economic theory, but have their own assumptions and properties that must be met in order to generate valid casual inference.

In this chapter we review statistical approaches that are routinely used across a range of scientific disciplines to infer cause-effect relationships from observational data. We argue that a causal inference framework is highly suitable for ex-post appraisal because it is specifically designed to estimate effects that arise when "treatments" are non-randomly assigned, as is the case with most transport interventions. The key consequence of non-random assignment is that the effect of the treatment is "confounded", implying that units in receipt of the treatment (or some particular dose of the treatment) may differ in systematic ways from units with an alternative treatment status. The objective of causal analysis is to estimate the average effect of the treatment (or intervention), net of confounding, or in other words, to uncover the marginal causal effect. We refer to this as a treatment effect estimation problem and it is within this context that we discuss possible methods for ex-post appraisal.

The chapter is structured as follows. Firstly, it discusses ex-post appraisal as a treatment effect estimation problem within the potential outcomes framework for causal inference. It outlines the implication of non-random treatment assignment and the problem of confounding, and then introduces the key assumptions required for valid causal inference. Next, it describes strategies for consistent treatment effect estimation. It discusses identification of causal effects under "ignorability" via covariate adjustment, propensity score adjustment and doubly-robust methods. It then reviews two approaches that are commonly used when ignorability is not assumed to hold. Lastly, it provides two empirical examples of the use of causal techniques for ex-post evaluation: one which evaluates the impacts of urban road network capacity expansions in the US and one which considers the regional economic impacts of High-Speed Rail investments in Spain. Conclusions are then drawn in the final section.

Ex-post appraisal as a treatment effect estimation problem

Our emphasis on causality in this chapter arises from the conviction that transport policy is fundamentally concerned with cause-effect relationships. In the United Kingdom for example, the following concerns have been highly influential in decision making in recent years:

- What effect will fuel taxation have on transport emissions?
- By how much will traffic volumes reduce under congestion charging?
- How will travel demands change as standards of living rise?
- Will investment in transport infrastructure boost the productivity of the economy?
- How will investment options affect network performance?

Each of these issues involves a cause-effect relationship and the underlying goal of policy is to attempt to shape future outcomes via public intervention. For decision makers the question of interest is: what impact, or outcome, will proposed interventions have?

Ex-post evaluation can help answer this question. By applying statistical models to historic data we can attempt to capture the key relationships of interest and can seek to evaluate the effect of past interventions on defined outcomes. There are, however, two key problems we face in obtaining a causal interpretation from observed data. First, is that we observe only what has actually occurred, not what would have taken place had we intervened in a different way. Second, the interventions we make are rarely randomly assigned and non-random assignment obscures cause-effect relationships. Taken together, these two issues in effect mean that we do not have experimental evidence upon which to base policy decisions.

In this section we outline the potential outcomes framework for causal inference which can be used to obtain a causal interpretation of observational data in the absence of experimental conditions. We discuss the defining characteristics of this approach and demonstrate how it could be used to infer cause-effect relationships for transport interventions.

Challenges in estimating the causal effect of treatments on outcomes

There are three key components that require attention in analysing cause-effect relationships from observed data: the intervention (or treatment) to be studied; the outcome of interest; and any relevant characteristics of the units of observation. For ex-post analysis we are fundamentally interested in the effect that a transport intervention (or some set of interventions) has on an outcome. We may wish to know what the outcome would have been had the intervention not been applied, or if some different intervention had been applied.

Relevant outcomes of interest could relate to traffic conditions (e.g. speeds, flow, safety, congestion), economic characteristics (e.g. output, productivity, growth), mode share, environmental consequences, social concerns, and so on.

For analytical purposes, an intervention in the transport system can be viewed as an observed realisation of random variables whose manipulation produces different outcomes. We refer to such random variables as "treatments", defined in the broadest sense to encompass any "regime" which can be manipulated to produce some effect. For instance, a treatment could involve the construction of a new link, the imposition of speed limits, changes in transport prices, changes in frequency or quality of service, allocation of subsidies, and so on. Treatment variables can be binary, multivalued or continuous. Table 5.1 gives relevant examples of transport interventions classified as treatment variables.

Table 5.1. **Transport interventions classified as treatment variables**

Binary	Multi-valued	Continuous
Tolled / untolled route	Frequency of service	Network capacity
Presence of speed camera	Speed limit	Length of segregated route
20 mph zone designation	Cars per train	Density of intersections
Peak / off-peak	No. of O-D routes	Accessibility
Pedestrianised / unpedestrianised	No. of network nodes	Tax / subsidy rates

We are interested in the effect of the treatment on the outcome, but we also recognise that the units under study will likely not have homogeneous characteristics, and these may be relevant to the fundamental relationship of interest. Depending on the purpose of the analysis, and the available data, units could comprise particular transport schemes, network links, people, households, firms, geographical zones, cities etc.

We seek to estimate causal effects using data to represent these three components. We define $z_i = (y_i, d_i, x_i)$, $i = 1, ..., n$, as a random vector of observed data where for the i-*th* unit of observation y_i denotes an outcome (or response), d_i the treatment (or exposure) received, and x_i a vector of pre-treatment covariates. As mentioned previously, the treatment can be binary (i.e. $D \in \{0, 1\}$); multivalued, in which dose d can take values in *m* categories $D \equiv (d_0, d_1, ..., d_m)$; or continuous with dose d taking values in $D \subseteq R$.

We want to estimate the effect that treatments have on outcomes. To do so we will draw on the potential outcomes framework for causal inference, which was first put forward for binary treatments in a series of papers in the 1970s by Rubin (e.g. Rubin 1973a,b, 1974, 1977, 1978), although Rubin acknowledges precursors to his approach in earlier works by Fisher (1935) and Neyman (1923). The potential outcomes framework defines the conditions under which we can estimate causal effects from observed data. These are two fundamental issues that shape the potential outcome approach.

Missing data

Ideally, we would calculate the effect of each treatment on an individual (or unit-by-unit) basis. Thus, for unit i and binary treatment $D \in \{0, 1\}$ we can define two potential outcomes: $Y_i(0)$ if $D_i = 0$, and $Y_i(1)$ if $D_i = 1$. The individual causal effect (ICE) of the treatment is then defined as:

$$\tau_i = [Y_i(1) - Y_i(0)].$$

For multivalued or continuous treatment we can define a potential outcome $Y_i(d)$ associated with each dose of treatment d, with $Y_i = \{Y_i(d) : d \in D\}$ denoting the full set of potential outcomes. The relevant ICEs would then be:

$$\tau_i = [Y_i(d) - Y_i(0)],$$

or the difference between the outcome, given assignment to dose *d* and assignment to no treatment.

A key problem for causal inference, however, is that the data available for estimation reveal only actual, not potential outcomes. For a binary treatment we observe:

$$Y_i = Y_i(1)\, I_1(D_i) + Y_i(0)[1 - I_1(D_i)],$$

where $I_1(D_i)$ is the indicator function for receiving the treatment, but we do not observe the joint density, $f(Y_i(0),\ Y_i(1))$, since the two outcomes never occur together. For multivalued or continuous treatments we observe only $Y_i(D_i)$, and outcomes at all other levels, $d \neq D_i$, are unobserved and we refer to these as counterfactual outcomes.

Thus, the problem we face is that the observed data do not provide enough information to evaluate ICEs because we do not observe the potential outcomes arising from treatment allocations that are contrary to fact. Holland (1986) refers to this as a fundamental identification problem of causal inference. A key insight of the potential outcomes approach is that if we focus on estimating average causal effects, rather than ICEs, then we do not have to observe all potential outcomes.

Average casual estimands of interest include Average Potential Outcomes (APOs) and Average Treatment Effects (ATEs). For binary treatments the APOs are:

$\mu(1) = E[Y_i(1)]$ and $\mu(0) = E[Y_i(0)]$,

and the ATE is defined as:

$\tau\ (1) = \mu(1) - \mu(0)$.

For continuous and multi-valued treatments the APO under treatment level d is denoted:

$\mu(d) = E[Y_i(d)]$,

and the ATE is:

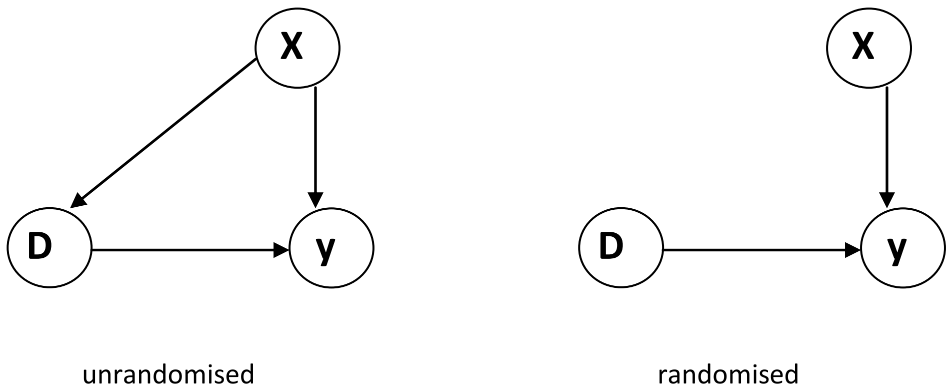
$\tau\ (d) = \mu(d) - \mu(0)$.

Other causal estimands can include ATEs on the treated, quantile effects, ATEs for a variety of sub-populations, ATEs conditional on covariates and causal odds and risk ratios. In this chapter the primary concern is with APOs and ATEs as defined above.

Non-random assignment and the problem of confounding

The conditions under which we can use the observed data to estimate APOs and ATEs depend crucially on whether the treatment is assigned randomly or not. Figure 5.1 below shows a graphical comparison of randomised and non-randomisation treatment assignments.

Figure 5.1. **Directed acyclic graph of observational data with randomisation and non-randomisation of treatment assignment**



Under a randomised assignment, unit characteristics X have no influence on the treatment received (i.e. on D). Consequently, potential outcomes are *unconditionally independent* of the treatment assignment mechanism. For binary treatments randomization implies:

$$Y_i(0), Y_i(1)) \perp I_1(Di)$$

and for multivalued or continuous treatments:

$$Y_i(d) \perp I_d(Di) \text{ for all } d \in D,$$

where $I_d(D_i)$ is the indicator function for receiving dose d of the treatment. Under a random assignment there are no systematic differences in characteristics between treated or controlled units, or in the case of multivalued and continuous treatments, between units receiving different doses of the treatment. Consequently, we can treat the unobserved potential outcomes much like data that are missing at random and consistent estimators of ATEs for binary, and multivalued or continuous treatments can then be formed as:

$$\hat{\tau}(1) = \frac{\sum Y_i \cdot I_1(D_i)}{\sum I_1(D_i)} - \frac{\sum Y_i \cdot (1 - I_1(D_i))}{\sum(1 - I_1(D_i))} \quad (1)$$

and:

$$\hat{\tau}(d) = \frac{\sum Y_i \cdot I_d(D_i)}{\sum I_d(D_i)} - \frac{\sum Y_i \cdot I_0(D_i)}{\sum I_0(D_i)} \quad (2)$$

Under non-randomisation, however, allocation of the treatment depends on a set of covariates X that are themselves important in determining outcome Y. Thus, some part of the association between the treatment and the outcome could be attributed to X rather than D. Under these circumstances, we refer to X as confounders and note that simple comparisons of mean responses across different treatment groups (as in equations 1 and 2) will not in general reveal a "causal" effect because mean outcomes across treated and control units will differ regardless of treatment status.

Identification of causal effects via the potential outcomes framework

While non-random assignment has consequences for causal estimation, consistent estimates of APOs and ATEs can still be obtained under the potential outcome framework. However, adjustment for confounding has to be made. In this subsection of the chapter, we define the conditions under which causal estimates can be identified in the presence of confounding.

There are three key assumptions required for valid APO and ATE estimation in the presence of confounding within the potential outcomes framework. These are as follows.

- **Conditional independence**: the potential outcomes for unit i should be conditionally independent of the treatment assignment, given a (sufficient) set of observed covariates X_i. For binary treatments the assumption requires that:

  $$Y_i(0), Y_i(1)) \perp I_1(D_i)|X_i, \tag{3}$$

  and for multivalued or continuous treatments, Imbens (2000) and Hirano and Imbens (2004) introduce the concept of weak conditional independence, which can be stated as:

  $$Y_i(d) \perp I_d(D_i)|X_i \text{ for all } d \in D. \tag{4}$$

  The key difference between the binary and non-binary assumptions is that, in the latter, conditional independence is required to hold for each value of the treatment (i.e. pairwise), but not joint independence of all potential outcomes.

  The conditional independence assumptions essentially require that, conditional on some set of pre-treatment covariates, assignment to treatment does not depend on the outcome. If X_i is sufficient for this to hold, then we can in effect mimic, for observational data, the assignment that would occur in a randomised control trial in which the treatment is allocated independently of pre-treatment characteristics.

- **Common support** – the support of the conditional distribution of $X_{i,}$ given a particular treatment status, should overlap with that of X_i given any other treatment status. For binary treatments, this requires that the probability of assignment to the treatment lies strictly between zero and one:

  $$0 < \Pr(I_1(D_i) = 1|X_i = x) < 1, \forall\ x. \tag{5}$$

  For multivalued or continuous treatments we require common support by treatment status in the covariate distributions within some region of dose $C \subseteq D$. A sufficient condition is that for any subset of C, say $A \subseteq C$,

  $$\Pr(D_i \in A|X_i = x) > 0, \forall\ x \tag{6}$$

  The intuition behind the common support, or overlap, assumption is that if some sub-populations observed in X_i have zero probability of receiving (or not receiving) a treatment, then it does not make sense in these cases to talk of a treatment effect since the counterfactual does not exist in the observed data.

- **Stable unit treatment values**: the relationship between observed and potential outcomes must comply with the Stable Unit Treatment Value Assumption (SUTVA) (e.g. Rubin 1978, 1980, 1986, 1990), which requires that the observed response under a given treatment allocation is equivalent to the potential response under that treatment allocation. For binary treatments we require that:

$$Y_i = I_1(D_i)Y_i(1) + (1 - I_1(D_i))Y_i(0) \quad (7)$$

for all i = 1, ..., N. For multivalued or continuous treatments we require:

$$Y_i \equiv I_d(D_i)Y_i(d) \quad (8)$$

for all $d \in D$, for all $Y_i(d) \in Y_i$, and for i = 1, ..., N.

The SUTVA requires that the outcome for each unit be independent of the treatment status of other units or, in other words, there should be no interference in treatment effects across units (Cox, 1958). It also implies that there are no different versions of the treatment. The no-interference assumption is generally satisfied when the units are physically distinct and have no means of contact. Violations of the assumption can occur when proximity of units allows for contact and this presents a particular concern for transport applications.

The three assumptions defined above, which are together referred to by Rosenbaum and Rubin (1983) as strong ignorability, allow for identifiability of causal effects from observational data. In the case of binary treatments the ATE can be derived as:

$$\tau = E_i(Y_i(1) - Y_i(0)) = E_X [E_i(Y_i(1)|X_i = x) - E_i(Y_i(0)|X_i = x)] \quad (9a)$$

$$= E_X [E_i(Y_i(1)|X_i = x, I_1(D_i) = 1) - E_i(Y_i(0)|X_i = x, I_1(D_i) = 0)] \quad (9b)$$

$$= E_X [E_i(Y_i|X_i = x, I_1(D_i) = 1) - E_i(Y_i|X_i = x, I_1(D_i) = 0)] . \quad (9c)$$

Conditional independence justifies the equality of (9a) and (9b), the SUTVA allows the substitution of observed for potential outcomes to give (9c), and overlap ensures that the population ATE in (9c) is estimable since there are units in both the treated and untreated groups. Note that the ATE is defined as an expectation over covariates X. If we do not take this expectation, but instead simply use the integrand, we obtain an estimate of the causal effect of D within strata of X. In other words, we get the conditional treatment effect, that is, the average treatment effect for units with characteristics X = x. By integrating X out of this distribution we get the average causal intervention distribution.

For continuous or multivalued treatments, the APO under a given dose D = d, $\mu(d) = E[Y_i(d)]$, or the dose-response function, can be derived as:

$$E[Y_i(d)] = E_X [E(Y_i(d)|X_i)] = E_X [E(Y_i(d)|I_d(D_i), X_i)] = E_X [E(Y_i|I_d(D_i), X_i)] , \quad (10)$$

where the second equality follows from conditional independence, the third from the SUTVA, and the overlap assumption ensures that the APO is estimable since there are comparable units across treatment levels.

Causal methods for treatment effect estimation

The literature on methods for causal estimation is vast and growing at a rapid rate. Consequently, a truly comprehensive review of the field is outside the scope of this chapter and, in fact, would make little contribution since excellent up-to-date reviews already exist, such as those by Hernan and Robins (2012), Imbens and Wooldridge (2009), Tsiatis (2006) and van der Laan and Robins (2003). Instead, we outline the general principles under which the construction of estimators proceeds. We do so first for methods that assume ignorability and then we consider two popular approaches that are used when ignorability is thought to be violated.

Treatment effect estimation under ignorability

Using the notation of Tsiatis and Davidian (2007), we define joint densities of the observed data of the form:

$$f_Z(z) = f_{Y|D,X}(y|d, x) f_{D|X}(d|x) f_X(x).$$

When ignorability is assumed to hold, estimation of APOs and ATEs generally proceeds in one of the following ways:

- **Direct covariate adjustment**: leave $f_{D|X}(d|x)$ and $f_X(x)$ unspecified and posit a model for $E[Y_i|D_i, X_i]$; the expectation of the conditional density of the response given treatment and covariates. This is typically achieved via an outcome regression (OR) model such as a Generalised Linear Model (GLM), a Generalised Linear Mixed Model (GLMM), a Generalised Additive Mixed Model (GAMM), or other spline-based approach. ATEs can be estimated directly from these OR models. This regression approach is commonly used in transport analyses.
- **Propensity Score adjustment**: leave $f_{Y|X}(y|x)$ and $f_X(x)$ unspecified but assume a model for $f_{D|X}(d|x)$ and use these to form Propensity Scores (PS), which measure the probability of assignment to treatment, given the set of observed pre-treatment covariates. An important result, due to Rosenbaum and Rubin (1983), is that the conditional independence assumption (i.e. equations 3 and 4) can be restated by replacing the covariate vector X_i with the scalar PS. Rosenbaum and Rubin (1983) proved this result in the case of binary treatments and Imbens (2000) and Hirano and Imbens (2004) generalise the PS to cover the case of multivalued and continuous treatments.

  The PS are to be used to form a number of different non-parametric and semi-parametric estimators, via weighting, matching, stratification, blocking and regression (for details, see Imbens and Wooldridge, 2009). A key advantage in using the PS is that it avoids the need to condition on a potentially high dimensional covariate vector and it is this dimension-reducing property that allows for effective implementation of flexible estimators .Another advantage of the PS is that it is highly effective in isolating the region of common support, a task that is difficult using multiple covariates (for discussion, see Joffe and Rosenbaum, 1999).

- **Doubly-Robust estimation**: leave f(x) unspecified but assume both an OR model and a PS model and form an estimator that combines both models. This is usually achieved by weighting or augmenting the OR model with covariates derived by inverting the PS. The key feature of doubly-robust estimators is that APO and ATE estimates are consistent and asymptotically normal when either the OR or the PS model are correctly specified, but we do not require both models to be correct (e.g. Robins, 2000; Robins *et al.*, 2000; Robins and Rotnitzky, 2001; van der Laan and Robins, 2003; Lunceford and Davidian, 2004; Bang and Robins, 2005; Kang and Schafer, 2007). The rationale for doubly-robust estimation is that the analyst effectively has two chances at getting the model specification right.

Estimation given a non-ignorable treatment assignment

The validity of the estimation approaches discussed in the previous subsection requires us to maintain that ignorability holds. When this assumption is no longer tenable we have to look at other approaches. In this subsection we review two popular estimators that are used when ignorability is not assumed: differences-in-differences and instrumental variables.

Differences-in-differences

A common problem in identifying causal effects is that there may be unobserved differences between the treated and untreated units which affect potential outcomes and are also influential in treatment assignment. In addition, we may suspect that temporal trends affect the outcome variable due to events that are unrelated to the treatment.

The differences-in-differences (DID) estimator addresses such potential sources of bias by using information for both treated and control groups in both pre- and post-treatment periods. The DID estimator approximates:

$$\tau DID = \{E[Yi(1)|D = 1] - E[Yi(1)|D = 0]\} - \{E[Yi(0)|D = 1] - E[Yi(0)|D = 0]\} \quad (11)$$

The "double-differencing" of the DID estimator removes two potential sources of bias. Firstly, it eliminates biases in second period comparisons between the treated and control groups that could arise from time-invariant characteristics. Secondly, it corrects for time-varying biases in comparisons over time for the treated group that could be attributable to time trends unrelated to the treatment.

It is important to note two potential limitations with the DID approach. Firstly, it relies on the strong identifying assumption that the average outcomes for the treated and control groups would have followed parallel paths over time in the absence of the treatment.

Secondly, the model is sensitive to error specification and, in particular, it has been shown that the existence of correlation within groups or over time periods can adversely affect the performance of the DID estimator (Bertrand et al., 2004).

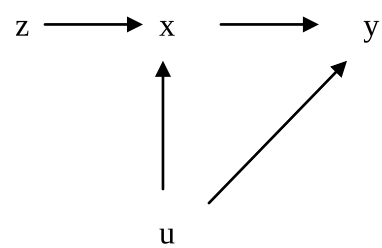
Instrumental Variables (IV)

The IV estimator is well known and widely used and for that reason we do not provide an extensive review here. The key principles of IV estimation are:

- Find a set of instruments which are exogenous and highly correlated with the covariates;
- Use the instruments to enforce orthogonality between the error term and an instrument transformed design matrix.

The relationships assumed in IV estimation are shown graphically below in the context of the linear regression model $y = X\beta + u$ with instrument matrix Z.

Figure 5.2. **Relationships in instrumental variables estimation**



The defining characteristics of the IV model are that: changes in z are associated with changes in x, but do not lead to changes in y other than through x; z is causally associated with x but definitely not with u; z would not be found in the regression model for y.

A common method used to obtain IV estimates is two-stage Least Squares (2SLS):

- Regress each column of X on the instrument matrix Z.
- Regress y on the predicted values from the first stage.

IV can be used to establish causal effects under a non-ignorable treatment assignment and is particularly useful when endogeneity via bi-directionality is present. However, it is crucial that the two key assumptions of exogeneity and relevance are met, and in practice such instruments can be hard to find. When instruments are only weakly correlated with the endogenous regressors, or when the instruments themselves are correlated with the error term, IV estimation can produce biased and inconsistent estimates. This problem is further confounded by the fact that the available diagnostic statistics do not provide a foolproof means for detecting an inadequate instrument specification. To quote Hahn and Hausman (2003), even using standard tests for instrument validity, "the researcher may estimate 'bad results' and not be aware of the outcome". In addition, it is worth noting that the IV estimator can be much less efficient than OLS.

Applications

In this section we describe two applications of the treatment effect approach for ex-post evaluation of transport interventions. The first relates to an ex-post evaluation of urban road capacity expansions in the US. The objective is to estimate ATEs from road capacity expansion in relation to induced traffic demand, traffic densities and productivity. The study uses a PS-based methodology for dose-response estimation for continuous treatments proposed by Graham *et al*. (2014). The second application considers the regional economic impacts of high-speed rail investment in Spain using a DID estimator.

Ex-post evaluation of urban road capacity expansions in US cities

Objective

The objective of this study is to use available longitudinal data to assess the impacts of urban road network capacity expansions on traffic volumes, traffic densities and productivity.

Method

The study uses a GPS-based regression methodology to control for confounding and estimate APOs and ATEs. The APOs of interest are defined by $\mu(d) = E[Y_{it}(d)]$ and the ATEs by $\tau(d) = E[Y_{it}(d)] - E[Y_{it}(0)]$, where d is some dose of capacity expansion, i indexes units and t indexes time. Calculations are made for several doses of interest and a dose response curve is formed using penalised spline regression.

As mentioned above, the GPS is defined by $\pi(D_{it} = d|X_{it}; \hat{\alpha})$ and for valid causal inference we require conditional independence and common support:

$$Y_{it}(d) \perp D_{it} = d|X_{it} \text{ and } \Pr(D_{it} \in A|X_{it} = x_{it}) > 0 \ \forall\ x_{it}, A \subseteq C$$

where C is a region of common support (e.g. Hirano and Imbens, 2004).

Consistency requires that X_{it} is sufficient to represent confounding. This is, however, effectively an untestable assumption. In longitudinal applications we often assume the existence of unobserved covariates U_i or W_i, say, which could enter our causal model such that $D_{it} = f(X_{it}, U_i, W_i)$ and $Y_{it} \perp D_{it} = d|X_{it}, U_i$. Thus, U_i is a time-invariant unobserved confounding covariate while W_i is a time-invariant unobserved non-confounding covariate. To address this issue, Graham *et al.* (2014) specify a longitudinal mixed (LM) model for the GPS:

$$D_{it} = X_{it}^T\theta_1 + b_i + H_{i,t-p}^{yT}\theta_2 + \varepsilon_{it} \text{ with } \quad b_i \sim N(0, \sigma_b^2$$

which, in addition to observed time-varying confounders X_{it} conditions on unobserved unit level effects b_i and some lag of the response variable $H_{i,t-p}^{y}$. Thus the paper proposes a GPS approach to ATE estimation which allows for measured time-varying confounding, unobserved time-invariant confounding, and bi-directionality between response and treatment.

Analytical results and simulations presented in the paper show that under given conditions a LMGPS approach will yield unbiased estimates of the dose-response function, but more extensive conditioning can adversely affect efficiency and can render the task of finding overlap in support of the covariate distribution more challenging.

The algorithm for ATE estimation under the LMGPS approach is as follows:

- Estimate $f_{D|X,U}(D_{it}|X_{it}, b_i; \alpha)$ using a mixed model;
- Use $\hat{\alpha}$, with some appropriate density function, to calculate LMGPSs for observed $\pi(D_{it}|X_{it}, b_i; \hat{\alpha})$ and unobserved $\pi(d|X_{it}, b_i; \hat{\alpha})$ treatments;
- Isolate a common support region

  $\Pr(D_{it} \in A|X_{it}, b_i) > 0 \ \forall\ x_{it}, A \subseteq C$;

- Estimate E $[Y_{it}|D_{it}, \pi(D_{it}|X_{it}, b_i; \widehat{\alpha})]$ using a flexible model;
- Average over predicted values from 4., evaluated at dose d, to obtain a point estimate of the APO at d: μ (d);
- Repeat for all dose of interest, form the dose-response curve, and estimate ATEs:

  $\tau(d) = \mu(d) - \mu(0)$
- Use a single (block) bootstrap re-sampling scheme over 1. to 6. to obtain standard errors.

Data

The data available for estimation are taken from the Texas Transportation Institute (TTI) urban mobility data, which describe traffic conditions for 101 US cities over the period 1982-2007.

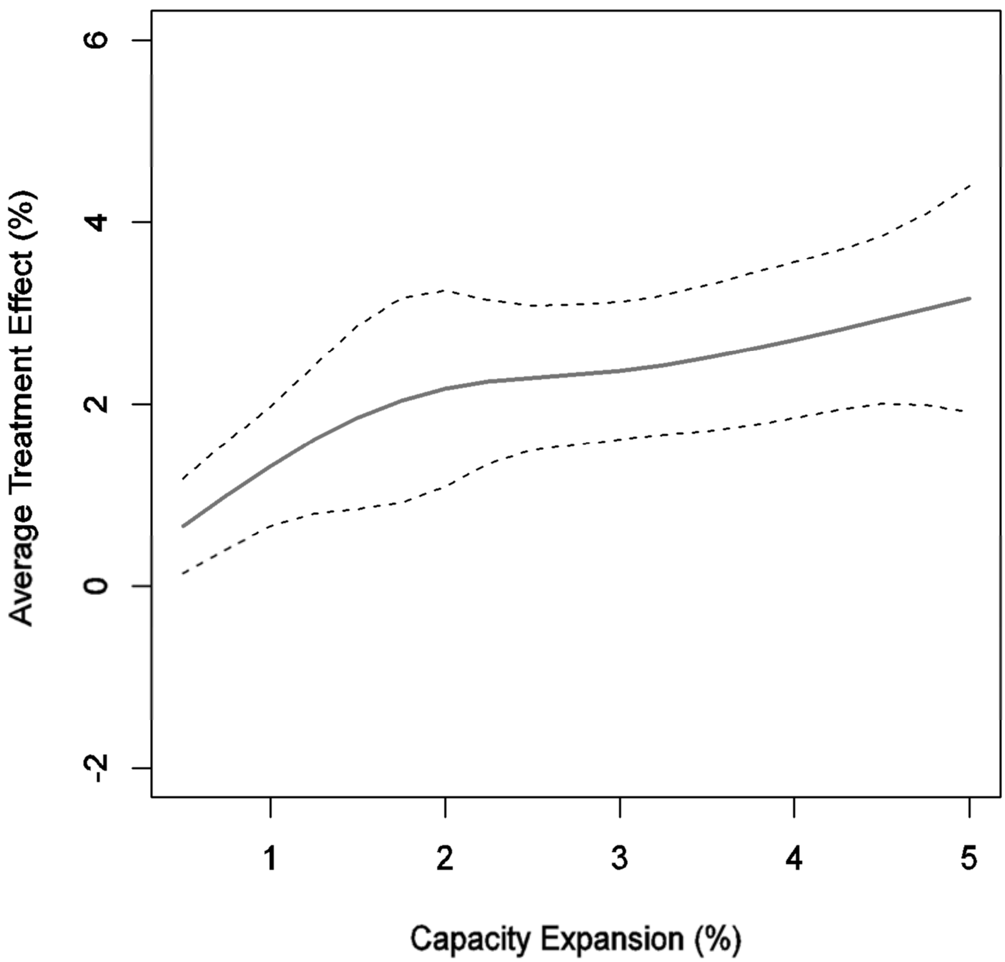
- **Responses:** Annual proportional change in demand (vmt), network performance (delay per vmt), and productivity (average wage).
- **Treatment:** Annual proportional change in network lane miles.
- **Pre-treatment covariates (confounders):**
  - lagged responses: to capture reverse causality
  - congestion and traffic volume: measured by delay and vmt
  - network scale and mix: network length, mix of freeway/arterial
  - traffic mix: volume on freeway/arterial
  - mode characteristics: public transport patronage, state fuel price
  - economy: productivity, income and economic structure
  - employment and population distribution and growth.
- **Unobserved (unknown) confounders:**
  - zone/area/region characteristics, road network design, activity/travel behaviour.

Results

The results for our three responses are shown graphically below. In each figure the y-axis shows the ATE and the x-axis shows the corresponding dose of capacity expansion.

First we consider results for traffic volumes as measured by vmt.

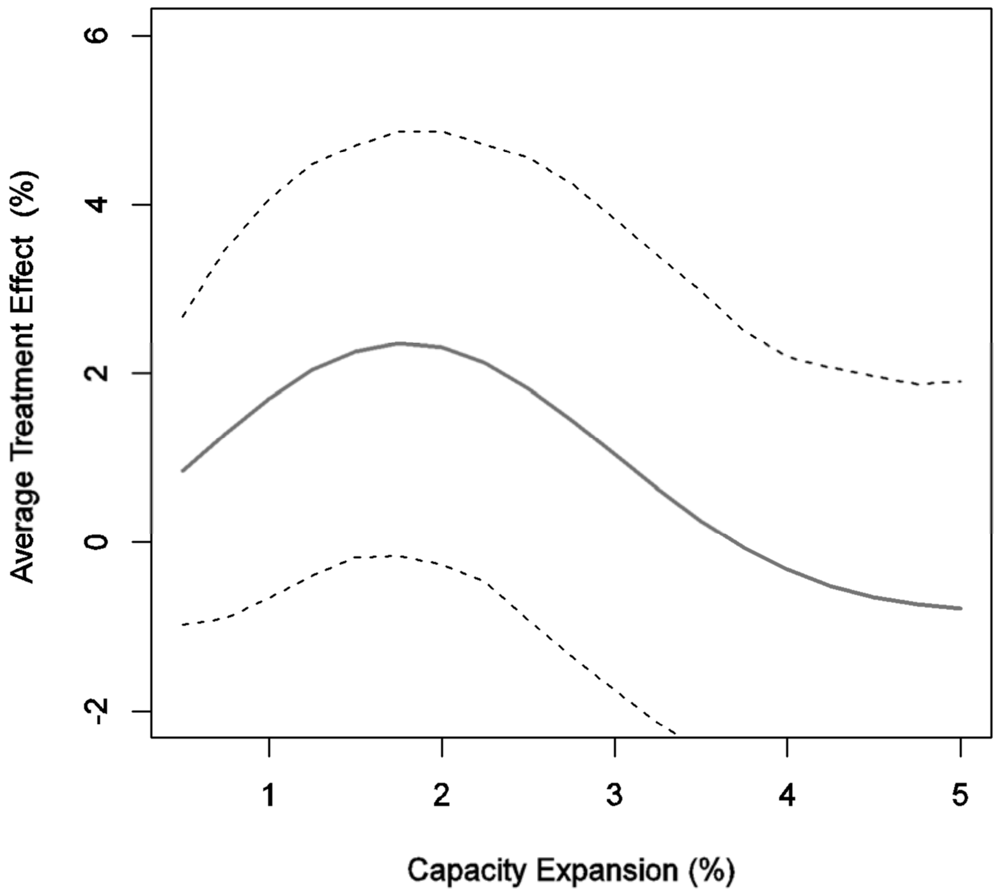
Figure 5.4. **Dose-response relationship for traffic volumes**



The results show evidence of induced demand over the range of dose having adjusted for confounding. The ATE is growing faster than capacity for doses of up to 2% increase in capacity. On average we find that a 10% increase in lane miles is associated with a 9% increase in vmt net of "natural growth" (estimated 1.4% per annum). As a consequence of this, we find that capacity expansions in the range considered have not, in general reduced traffic density (i.e. the traffic volume to capacity ratio).

Next, we look at impacts on network performance as measured by delay per vmt.

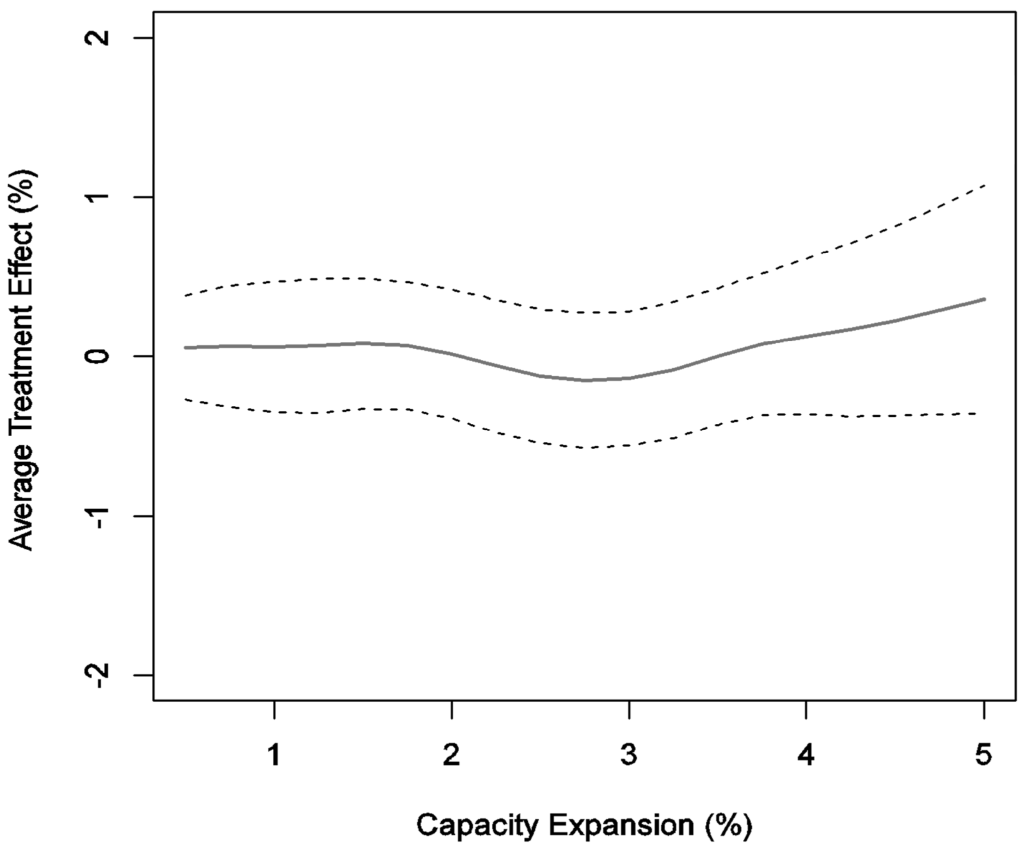
Figure 5.5. **Dose-response relationship for network performance**
(delay per vmt)



The results indicate that capacity expansions have not ameliorated urban congestion. The average road user has not experienced change in delay from capacity expansions as there have been no statistically significant effects on delay per vmt, and we find this to be the case even for large capacity expansions. In fact, due to natural growth, congestion has worsened by approximately 3% per annum and because there is now more traffic, total urban delay increases over the range of dose.

Finally, we look at effects on productivity as represented by the average urban wage rate.

Figure 5.6. **Dose-response relationship for productivity (average wage)**



The results indicate that urban road network expansions have not induced higher productivity. If we run a "naïve" regression of productivity on treatment we do find a positive association between capacity growth and wages, but we do not find significant ATEs having adjusted for confounding and isolated a region of common support.

Thus, our causal analysis finds that urban road network expansions have induced demand but have not ameliorated congestion or raised productivity. These results do not imply that there are no economic benefits from road capacity expansions per se. The results are specific to marginal changes on mature congested urban networks. While capacity expansions have allowed for increased mobility, in the sense that there is more traffic, network generalised costs have not improved and total urban delay has risen. The scale (increased traffic) effect does not appear to have influenced productivity (either +ve or -ve).

Ex-post evaluation of regional economic impacts of high-speed rail in Spain

Objective

Between 2000 and 2010, the Spanish Government carried out the largest high-speed rail construction programme in Europe, such that by 2011 the Spanish HSR network had become the largest in Europe, exceeding France and Germany. By 2020, it is planned that 90% of the country's population will live within a 50 km radius to the nearest high-speed rail station. Investments in high-speed rail projects are frequently justified on the basis of projected positive effects on regional and national economic growth. In this application, a DID analysis is undertaken to study impacts on economic output arising from construction of the Madrid-Barcelona high-speed rail corridor. This is PhD work which is still in progress, and is presented here as an illustration of ex-post evaluation rather than as a definitive statement on the impacts of HSR in Spain.

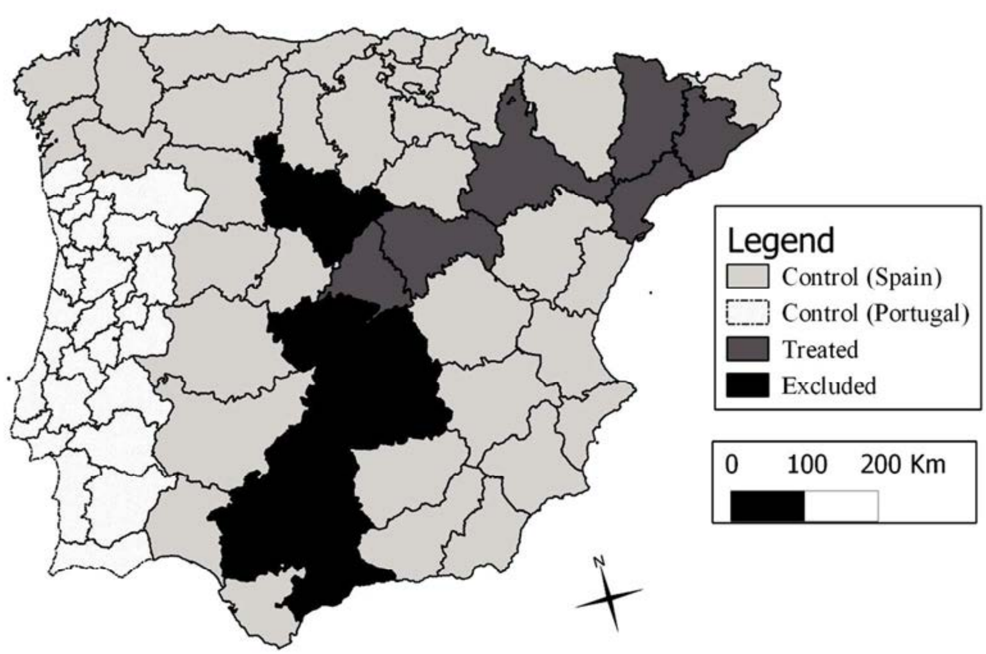
Method

To assess the effect of the HSR corridor linking Madrid with Barcelona, we divide the Spanish economy into 47 peninsular provinces and treat access to HSR as a binary treatment. Provinces receiving/not receiving a high-speed rail connection are called treated/untreated and provide the basis for the DID analysis. These are shown in Figure 5.7. We use GVA per capita (GVApc) as the outcome variable in province *i* at time *t*. We run a regression of this response on the year variable, a binary variable for each of the groups (east, west and north) and the three interaction terms between year and the three potential control groups north, east and west (year*east, year*west and year*north, respectively) where the base case is the group of treated provinces. We also estimate the same equation adding two covariates to account for the economic structure of the different provinces: share of jobs in manufacturing, energy and construction sector (share industry), and share of jobs in the services sector (share services).

Results

The results indicate that predictions of a positive impact on the economic performance of regions receiving HSR have not taken place, at least in the short to medium term. In the case of the Madrid-Barcelona HSR corridor, our results show that there are no significant differences in the pattern of regional economic growth before and after the HSR corridor between the treated and untreated provinces.

Figure 5.7. **Control and treated provinces for the north-east corridor of Spain**



Conclusions

In this chapter we have reviewed methods that seek to draw causal inference from observed data and have shown how they can be applied to undertake ex-post evaluation of transport projects. We argue that a causal inference framework based on potential outcomes is highly suitable for ex-post appraisal because it is specifically designed for instances in which treatments are non-randomly assigned and experimentation is not possible, circumstances that characterise the allocation of transport interventions. The methods we review have been used extensively for casual analysis across a range of scientific disciplines, but to our knowledge have received little attention in transport analyses. We provide two applications of ex-post evaluation based on causal techniques: one which evaluates the impacts of urban road network capacity expansions in the US, and one which considers the regional economic impacts of high-speed rail investments in Spain.

A major advantage of the statistical causal methods is that we can use them to analyse the impact of interventions without making strong *a priori* theoretical assumptions about underlying economic behaviour, as is required in ex-ante or ex-post CBA. However, valid causal inference from observational data has its own set of rather stringent assumptions, which in many instances may not hold in the available data.

References

Bang, H. and J.M. Robins (2005), "Doubly robust estimation in missing data and causal inference models", *Biometrics*, 61, 962-972.

Bertrand, M., E. Duflo and S. Mullainathan (2004), "How much should we trust difference-in-differences estimates?", *Quarterly Journal of Economics*, 119 (1), 249-275.

Cox, D.R. (1958), *Planning of Experiments*, London: John Wiley & Sons.

Fisher, R.A. (1935), *The Design of Experiments*, Edinburgh: Oliver and Boyd.

Graham, D.J., E.J. McCoy and D.A. Stephens (2014), "Quantifying causal effects of road network capacity expansions on traffic volume and density via a mixed model propensity score estimator", *Journal of the American Statistical Association*, DOI:10.1080/01621459.2014.956871.

Hahn, J. and J. Hausman (2003), "Weak instruments: diagnosis and cures in empirical economics", *American Economic Review*, 93, 118-125.

Hernan, M.A. and J.M. Robins (2012), "Causal Inference", *Monographs on Statistics and Applied Probability*, London: CRC Press.

Hirano, K. and G.W. Imbens (2004), "The propensity score with continuous treatments", in: A. Gelman and X.L. Meng (Eds.), *Applied Bayesian modeling and causal inference from incomplete data perspectives*, pp. 73-84, New York: Wiley.

Holland, P.W. (1986), "Statistics and causal inference", *Journal of the American Statistical Association*, 81 (396), 945-970.

Imbens, G.W. and J.M. Wooldridge (2009), "Recent developments in the econometrics of program evaluation", *Journal of Economic Literature*, 47 (1), 5-86.

Imbens, G.W. (2000), "The role of the propensity score in estimating dose-response functions", *Biometrika*, 87 (3), 706-710.

Joffe, M.M. and P.R. Rosenbaum (1999), "Propensity scores", *American Journal of Epidemiology*, 150 (4), 327-333.

Kang, J.D.Y. and J.L. Schafer (2007), "Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data", *Statistical Science*, 22 (4), 523-539.

Lunceford, J.K. and M. Davidian (2004), "Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study", *Statistics in Medicine*, 23, 2937-2960.

Neyman, J. (1923), “On the application of probability theory to agricultural experiments. Essay on principles”, Section 9. *Statistical Science*, 5 (4), 465-480. Translated in 1990.

Robins, J.M. and A. Rotnitzky (2001), “Comment on ‘Inference for semiparametric models’: some questions and an answer”, *Statistical Sinica*, 11, 920-936.

Robins, J.M. (2000), “Robust estimation in sequentially ignorable missing data and causal inference models”, in: *Proceedings of the American Statistical Association, Section on Bayesian Statistical Science*, pp. 6-10, Alexandria, VA: American Statistical Association.

Robins, J.M., A. Rotnitzky and M.J. van der Laan (2000), “Comment on the Murphy and Van der Vaart article ‘On profile likelihood’”, *Journal of the American Statistical Association*, 95, 431-435.

Rosenbaum, P.R. and D.B. Rubin (1983), “The central role of the propensity score in observational studies for causal effects”, *Biometrika*, 70 (1), 41-55.

Rubin, D.B. (1990), Neyman (1923) and causal inference in experiments and observational studies, *Statistical Science*, 5 (4), 472-480.

Rubin, D.B. (1986), “Comment: which ifs have causal answers?”, *Journal of the American Statistical Association*, 81 (396), 961-962.

Rubin, D.B. (1980), “Comment on ‘Randomization analysis of experimental data in the Fisher randomization test’”, by Basu, *Journal of the American Statistical Association*, 75 (371), 591-593.

Rubin, D.B. (1978), “Bayesian inference for causal effects: the role of randomization”, *Annals of Statistics,* 6 (1), 34-58.

Rubin, D.B. (1977), “Assignment to treatment group on the basis of a covariate”, *Journal of Educational Statistics*, 2 (1), 1-26.

Rubin, D.B. (1974), “Estimating causal effects of treatments in randomised and non-randomised studies”, *Journal of Educational Psychology*, 66 (5), 688-701.

Rubin, D.B. (1973a), “Matching to remove bias in observational studies”, *Biometrics*, 29, 159-183.

Rubin, D.B. (1973b), “The use of matched sampling and regression adjustments to remove bias in observational studies”, *Biometrics*, 29, 185-203.

Tsiatis, A.A. and M. Davidian (2007), “Comment: Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data”, *Statistical Science*, 22 (4), 569-573.

Tsiatis, A.A. (2006), *Semiparametric theory and missing data*, Berlin: Springer.

van der Laan, M. and J.M. Robins (2003), *Unified methods for censored longitudinal data and causality*, Berlin: Springer.

List of Participants

Chair:

Mr. Tom Worsley
University of Leeds
Institute for Transport Studies
36 University Road
LS2 9JT Leeds
United Kingdom

Rapporteurs:

Mrs. Geraldine Barker
Director, Transport VFM National Audit Office
157-197 Buckingham Palace Road
London SW1W 9SP
United Kingdom

Prof. Alain Bonnafous
Professeur émérite de l'Université de Lyon
Laboratoire d'Economie des Transports (LET)
LET- UMR du CNRS 5593
ISH - 14 avenue Berthelot
69007 Lyon Cedex 07
France

Dr. Stephen Fitzroy
Senior Vice President
Economic Development Research Group (EDR)
155 Federal Street
Suite 600
02110 Boston, MA
USA

Dr. Daniel Graham
Senior Research Fellow
Imperial College London
Centre for Transport Studies
Civil and Environmental Engineering
South Kensington Campus
SW7 2AZ London
United Kingdom

Participants:

Mme Julie De Brux
Responsable d'Études et Prospectives
VINCI
9 place de l'Europe
92851 Rueil-Malmaison
France

Prof. Yves Crozet
Laboratoire d'Économie des Transports (LET)
Université de Lyon
LET-ISH
14 avenue Berthelot
F-69363 Lyon Cedex 07
France

Prof. Jonas Eliasson
Professor of Transport Systems Analysis
Royal Institute of Technology
Centre for Transport Studies
Teknikringen 72
100 44 Stockholm
Sweden

Dr. Thomas Kranz
Federal Highway Research Institute (BASt)
Brüderstr. 53
D-51427 Bergisch Gladbach
Germany

Mr. Martin Köning
IFSTTAR
SPLOTT
14-20 Boulevard Newton
Cité Descartes, Champs sur Marne
77447 Marne La Vallée Cedex 2
France

Dr. Niels Buus Kristensen
Director DTU Transport
Technical University of Denmark
Bygningstorvet 116 Vest
DK-2800 Kgs. Lyngby
Denmark

Professor Rosario Macario
Lisbon Technical University
Av. Rovisco Pais 1
P-1049-001 Lisbon
Portugal

Dr. David Meunier
Laboratoire Ville Mobilité Transport
Ministry of Ecology, Sustainable Development and Energy
Ecole des Ponts (UMR LVMT)
6-8 Avenue Blaise Pascal
F-77455 Marne La Vallée Cedex 2
France

Prof. Christopher Nash
University of Leeds
Institute for Transport Studies
36 University Road
LS2 9JT Leeds
United Kingdom

Mr. Sarbojit Pal
Associate Fellow
The Energy and Resources Institute (TERI)
Darbari Seth Block
India Habitat Centre Complex
Lodhi Road
New Delhi 110 003
India

Prof. John Preston
University of Southampton
School of Civil Engineering and the Environment
Transportation Research Group
Highfield
SO17 1BJ Southampton
United Kingdom

Dr. Panos Tzieropoulos
Directeur
Ecole Polytechnique Fédérale de Lausanne
ENAC-IIC-LITEP- GR-TZ
Bâtiment GC-C2 390, Station 18
CH-1015 Lausanne
Switzerland

Prof. Roger Vickerman
Director
University of Kent
Centre for European, Regional and
Transport Economics
Keynes College
CT2 7NP Canterbury
United Kingdom

Mr. Morten Welde
Researcher
Norwegian University of Science and Technology
Department of Civil and Transport Engineering
Concept Research Programme
N-7491 Trondheim
Norway

Ms. Pauline Wortelboer-Van Donselaar
Ministry of Infrastructure and the Environment
KiM Netherlands Institute for Transport Policy Analysis
P.O. Box 1511
Europlan 40
3526 KS Utrecht
The Netherlands

Observers:

Mr. Nathan Bounie
IFSTTAR
SPLOTT
14-20 Boulevard Newton
Cité Descartes, Champs sur Marne
77447 Marne La Vallée Cedex 2
France

Mr. Pablo Gutierrez
Policy Analyst
Permanent Delegation of Mexico to the OECD, Paris

Mr. Rudy Brage Ardao
Imperial College London
Centre for Transport Studies
Civil and Environmental Engineering
South Kensington Campus
SW7 2AZ London
United Kingdom

International Transport Forum Secretariat

Prof. José Viegas
Secretary-General

Joint OECD/ITF Transport Research Centre

Mr. Stephen Perkins
Head of Research and Policy

Dr. Dejan Makovšek
Economist

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where governments work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, Chile, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The European Union takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

OECD PUBLISHING, 2, rue André-Pascal, 75775 PARIS CEDEX 16
(74 2017 04 1 P) ISBN 978-92-82-10814-7 – 2017

International Transport Forum

Ex-Post Assessment of Transport Investments and Policy Intervention

Ex-post evaluation is important to improving the delivery of transport policy objectives. It can be used for multiple purposes at the core of which is the improvement of ex-ante assessment. A small number of jurisdictions employ ex-post evaluation systematically and leading experience is reviewed in this report.

One reason ex-post analysis is not more widely employed is a number of methodological and data challenges. Several approaches have been developed to tackle these, including advanced statistical approaches, establishment of transport observatories to retain data that would otherwise be lost and a case study approach to compare similar locations subject to different investment and policy treatments.

This report examines examples of best practice, discussing the potential of different approaches and how they can be used to complement each other.

OECD publishing
www.oecd.org/publishing

International Transport Forum
2 rue André Pascal
75775 Paris Cedex 16
France
T +33 (0)1 45 24 97 10
F +33 (0)1 45 24 13 22
Email : itf.contact@oecd.org
Web: www.internationaltransportforum.org

9 789282 108147
(742017041P1)
ISBN 978-92-821-0814-7