International Transport Forum

Quantifying the Socio-economic Benefits of Transport

Roundtable Report

OECD

Quantifying the Socio-economic Benefits of Transport

Roundtable Report

Please cite this publication as:
ITF (2017), *Quantifying the Socio-economic Benefits of Transport*, ITF Roundtable Reports, OECD Publishing, Paris.
http://dx.doi.org/10.1787/9789282108093-en

ISBN 978-92-82-10808-6 (print)
ISBN 978-92-82-10809-3 (PDF)

Series: ITF Roundtable Reports
ISSN 2074-3378 (print)
ISSN 2074-336X (online)

Photo credits: Cover © Maciej Noskowski/iStockphoto.com.

Corrigenda to OECD publications may be found on line at: *www.oecd.org/about/publishing/corrigenda.htm*.

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

Socio-economic cost-benefit analysis (CBA) provides a quantitative measure of the extent to which, over its lifetime, a project or initiative will bring the community benefits that exceed its costs of construction and operation. This report describes efforts to improve the quality of transport CBA and its applicability to decision making. Three areas are addressed in detail: strategies for making the most of CBA, valuing and forecasting reliability benefits, and capturing wider economic impacts. The report is based on the papers and discussions at a roundtable meeting of 30 experts held in Paris in November 2015.

Key findings

CBA is a powerful framework that can be very useful to governments making investment decisions. However the standard application of transport CBA faces three major challenges that have attracted the attention of practitioners and researchers. First, there can be a mismatch between the information most sought by decision makers (such as the impact on jobs and regional growth) and what is supplied by a standard CBA (impact on national welfare and resource gains). Second, the scope of benefits captured in a standard CBA is generally constrained by practical limitations of forecasting and valuation capability. Third, fundamental changes to the quantity or locations of businesses, investments, households and employment that are anticipated from transport investments are not captured within standard CBA. Roundtable participants took the view that a multi-faceted approach is needed to address these shortfalls; CBA theory and practice need to be gradually expanded to incorporate more impacts in the rigorous valuation and forecasting framework; and CBA results need to be more effectively linked to other criteria in the broader decision-making framework, including by bringing in a more diverse evidence base.

Main recommendations

CBA guidelines can be expanded to include reliability and some wider impacts

The current evidence base on the valuation and forecasting of reliability benefits, agglomeration benefits and labour supply benefits provides a sufficiently rigorous basis for inclusion within the core CBA of major transport projects. If properly applied, based on local evidence, the formal inclusion of these benefits is better than either excluding them or applying simple mark-up rules.

Further research into reliability benefits is needed to improve confidence in results

There is significant variation among transport users in forming expectations about travel-time reliability and in responding to it. Current approaches to valuing and forecasting reliability benefits take a simplified approach in the interests of practicality. However, more research that disaggregates results and examines the linkage between the standard of reliability and the transport choices made by users will improve accuracy and build confidence in the results. The behavioural feedback can range from changes in transport mode choices through to fundamental reorganisation of housing and business locations. Closer international collaboration of researchers to share techniques, data and results is a promising avenue for accelerating progress in this regard.

Wider economic impacts should be examined in cases where they are expected to be significant

The relocation and reorganisation of businesses and households (a change in "economic geography") is a major motivation for some transport projects, such as regeneration schemes or transit-oriented developments at rail stations. Wider economic impacts from such changes need to be communicated to decision-makers. Formal inclusion of these impacts within CBA requires the identification of relevant "market failures" in the project area (such as monopoly power) and the nature of the activity the project affects. In many cases, there is no way of confidently predicting the economic geography outcomes when co-ordinated actions among actors are required, such as in property development, so the best approach is usually the application of scenarios. These scenarios can be informed by *ex-post* analysis of similar projects as these can reveal success factors and the reasonable range of impacts that may materialise. For any project though, a strong and scrutinised case for the inclusion of wider benefits is required to overturn the traditional assumption that user benefits adequately account for the whole economic impact of a project.

Further research into the impacts and tools for capturing wider impacts is needed

Practitioners currently face a choice between appraising the wider economic impacts of a transport project by taking a "user benefits (CBA) plus wider benefits" approach or using a "big model" (such as a Land Use-Transport Interaction model) to capture all impacts. Roundtable participants generally favoured the former approach. This is because no big models are yet able to adequately capture all the relevant impacts of a transport project, they require a very large amount of data and the complexity of the modelling required undermines transparency. However, given that LUTI or general equilibrium model, at least in theory, should be able to produce answers to the most relevant questions these may ultimately be best placed to address the limitations of CBA. With further research, there may come a point where the big models become responsive, accurate, and cheap enough to apply as the preferred project appraisal approach. Most roundtable participants though were of the view that that time has not yet arrived, so there is a strong motivation to keep improving the practice of CBA.

CBA can play an important role in decision making, but need not dominate

CBA is valuable yet imperfect. Appraisal is most useful to decision-makers when the CBA approach is clearly aligned with the objectives sought, when it draws on the best local evidence available and when the shortfalls and uncertainties are clearly highlighted in the analysis. Available evidence will not always be of sufficient quality to justify inclusion within the formal CBA. In such cases, supporting frameworks and alternative evidence will be useful to communicate possible project impacts. Two options in particular were highlighted at the roundtable. First, by drawing quantitative and qualitative insights from similar past projects, *ex-post* analysis can give vivid insights into potential economic geography changes and their driving forces. Second, complementary tools, such as economic impact analysis and qualitative explanation of non-quantifiable impacts, can help address shortfalls inherent in CBA. Presenting such diverse information to decision-makers is better than producing a single performance measure, since the latter can generally only be achieved either by including bold and unfounded assumptions or by ignoring impacts altogether.

Chapter 1

Improving transport cost-benefit analysis: Overview and findings

Daniel Veryard[1]

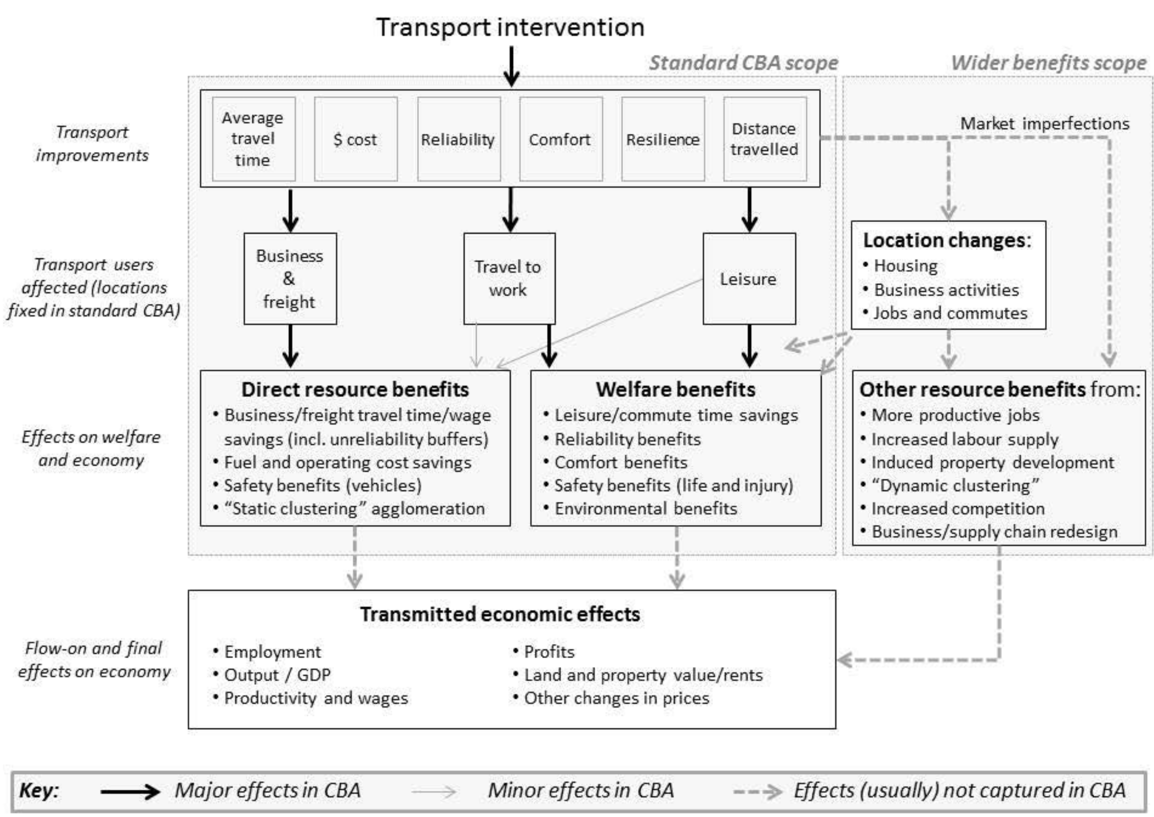
Socio-economic cost-benefit analysis (CBA) is a powerful framework that can be very useful to governments making investment decisions. However the standard application of transport CBA has room for improvement. This chapter describes efforts to improve the quality of transport CBA and its applicability to decision making. Three areas are addressed in detail: strategies for making the most of CBA, valuing and forecasting reliability benefits, and capturing wider economic impacts. The chapter is based on the papers and discussions at a roundtable meeting of 30 experts held in Paris in November 2015. Roundtable participants took the view that a multi-faceted approach is needed to address the shortfalls; CBA theory and practice need to be gradually expanded to incorporate more impacts in the rigorous valuation and forecasting framework; and CBA results need to be more effectively linked to other criteria in the broader decision-making framework, including by bringing in a more diverse evidence base.

1 International Transport Forum, Paris, France

Introduction

For most transport economists working in the field of public policy making the preferred tool for project prioritisation and selection is socio-economic cost-benefit analysis (CBA). The standard practice of CBA involves the estimation and valuation of direct benefits to transport users when the transport system is incrementally improved. Some direct external impacts are also often accounted for, such as impacts on congestion and the environment. Under the assumptions of constant returns to scale and perfect competition, this framework (if well applied) captures the ultimate economic impacts of expanded production, wages and employment (Dodgson, 1973; Jara-Diaz, 1986). That is, under these assumptions, the shaded area labelled "standard CBA scope" in Figure 1.1 is fully consistent with the lower box in the figure ("transmitted economic effects").

Figure 1.1. **Relationship between standard CBA and final economic effects of a transport project**



However, the standard application of transport CBA faces challenges that have attracted the attention of practitioners and researchers. These can be described with further reference to Figure 1.1 and broadly fall into three related themes:

- Relevance – There is often a mismatch between the information wanted by decision makers compared to what is supplied by a standard CBA. For instance, CBA supplies measures of resource benefits and social welfare benefits from the perspective of the nation. But decision makers may wish to understand the final (transmitted) impacts on jobs and economic activity in their region (the lower box in Figure 1.1). CBA results are not able to be directly interpreted in terms of final economic impacts, except in the broadest terms.

- Sophistication – The scope of benefits captured in a standard CBA is generally constrained by practical limitations of forecasting and valuation capability, even within the grey shaded "standard CBA scope" area in Figure 1.1. In this chapter we discuss in particular the limitations and recent improvements in the estimation of reliability benefits.

- Coverage – Transport improvements can encourage fundamental changes to the quantity or locations of businesses, investments, households and employment that are not captured within standard CBA (the right hand side of Figure 1.1). When, as is often the case, the theoretical assumptions underpinning standard CBA don't fully apply, the equivalence between direct benefits and final economic impacts breaks down. There are genuinely additional benefits that conceptually need to be added to a CBA to be complete. These wider economic impacts are discussed further in this chapter.

In relatively straightforward transport investment decisions, such as choosing between alignment options for roads, standard CBA is likely to provide sufficient relevant evidence. However, there will be cases where relying only on the estimation of impacts that are "easily" forecast and valued will not be sufficient. Focusing only on well-established benefits, such as direct travel-time savings to existing users, could bias investment towards projects that are well suited to narrow assessments (e.g. road expansion) and disadvantage projects that address particular objectives such as measures to encourage people to shift to public transport or strategies to address freight travel-time variability and could miss the critical aspects of projects designed to have a more "transformative" effect such as regeneration of depressed regions or enhancing the potential for growth through, for example, deepening and thickening the labour market.[1]

Progress has been made in the past decade to improve the quality of CBA and its applicability to decision making. This report reviews the state of the art in two areas in particular: reliability benefits and wider economic impacts. It summarises four papers on these issues commissioned for a roundtable meeting, held in Paris in November 2015, and brings together the discussions among leading experts at the roundtable. Edited versions of the four input papers comprise the remaining chapters of this report.

Strategies to improve the practice and relevance of transport CBA

Transport CBA is a powerful framework that provides a quantitative measure of the extent to which, over its lifetime, a project or initiative will bring the community benefits that exceed its costs of construction and operation. The framework is sufficiently flexible to be used to support a wide range of decisions. For instance, CBA can be used to filter out poor projects from consideration, or can be used to optimise a relatively promising project (e.g. refining alignments). The particular role of CBA can depend on the quality of the portfolio of transport projects that come under consideration (ITF, 2011).

The role that CBA plays in the overall decision making about transport investment is also affected by the relevance of the criticisms that CBA does not capture all of a project's expected impacts, or is unable to provide all the information that is relevant to decision makers. In some jurisdictions, such as several US states, CBA is not mandated or considered at all in decisions over transport investments (Weisbrod, Chapter 4). In most countries, and particularly in northern Europe and Australasia, CBA is a central, although not always dominant, part of the overall decision-making framework (Mackie and Worsley, 2013).

CBA results tend to be supplemented with other quantitative and qualitative information when presented to a decision maker. Where CBA is applied, the quantified (and monetised) estimates of project benefits are typically presented in a business case alongside descriptions of impacts that are more difficult to value (such as heritage), and information about how the direct (user) impacts of a project are transmitted through the economy into changes in employment and output. The latter effects can be estimated using a range of economic modelling techniques discussed later in this section. In jurisdictions where CBA is not applied, the local and regional economic impact estimates (as opposed to national welfare benefits) are given more prominence in decision making. Regardless of the technical approach to quantifying expected impacts, it is almost always the case that the final decision is taken based on a judgement over quantitative and qualitative information (Mackie and Worsley, 2013).

Roundtable participants identified and discussed a range of alternative and complementary strategies that could be pursued to improve the quality of transport CBA and its usefulness for decision makers:

- improved strategic alignment and communication of results
- applying complementary appraisal frameworks to capture effects outside traditional CBA
- drawing evidence from previous projects (case studies and *ex-post* analysis)
- extending the toolkit and scope of accepted CBA practice
- tailoring each CBA to the project's context and objectives.

Improved strategic alignment and communication

Ideally projects should be proposed on the basis of a careful strategic planning exercise that starts from the overall objectives or mission for a jurisdiction (supranational, national, regional or local). Projects are, however, often the result of more "instinctive" proposals from politicians or public officials. Several participants at the roundtable reported having been asked to appraise a project for which there was no clear statement of what problem it was trying to address, or what objectives it was trying to achieve. In such circumstances, where there is no "narrative", it is difficult to predict and advise on whether the project is likely to be a success, regardless of the analytical framework applied.

Practitioners can apply the standard CBA framework to estimate travel-time savings, safety benefits and environmental improvements regardless of the nature of the project and its objectives. However, it is unlikely that the results on their own will be sufficiently meaningful for political decision makers or their constituents. Instead, there is a need to first place the project into the overall strategic context and to develop a qualitative case for which transport, social, environmental and economic variables the project is likely to have the greatest effects.

Clarifying the strategic intentions of the project allows the CBA to align the assessment of benefits with the achievement of objectives. Clarity and alignment with project objectives will not only ensure the CBA is relevant to the project but will also ensure the results can be explained to decision makers in relevant language, with conclusions that are relevant to the project's original objectives.

Applying complementary appraisal frameworks

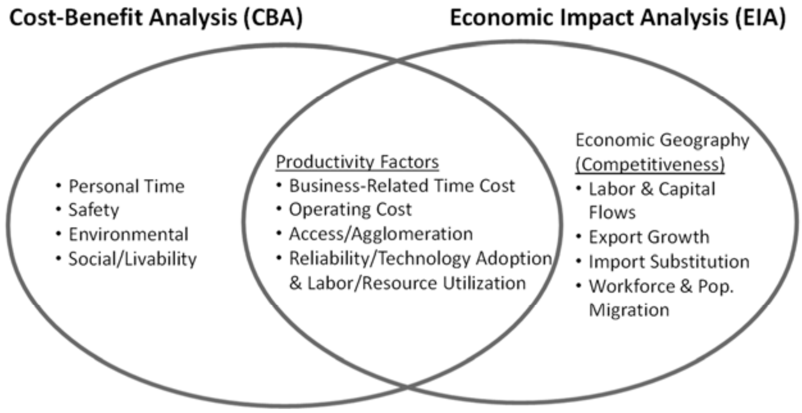
An obvious strategy to overcome the limitations of CBA is to complement it with an alternative appraisal framework. Two broad categories were discussed at the roundtable: economic impact analysis, which measures changes in business revenue, profits, wages and jobs for a specified area; and scoring methods, such as multi-criteria analysis (MCA).

Economic impact analysis can be used for *ex-ante* assessment to directly forecast the final economic impacts in the lower box of Figure 1.1 above. Economic impact analysis uses macroeconomic methods, such as econometric regressions, computable general equilibrium (CGE) models, spatial CGE models or mesoscopic models, and aims to forecast the increase in output or employment from an increase in the stock or quality of transport infrastructure (see Chapter 4 or Vickerman (2008) for a review). Economic impact analysis takes a different perspective to CBA but is somewhat overlapping in scope (Figure 1.2). Participants at the roundtable recognised the potential value of these models for two primary tasks: communicating economic impacts in terms relevant to decision makers (jobs and economic activity), and for highlighting the regional and socio-economic distribution of impacts.

CGE models in particular have the capability to account for some of the market imperfections that fall outside the scope of traditional CBA and can be used to provide information on the extent to which resources are displaced from one location to another (discussed further in the final section of this chapter) and more generally on who benefits and who loses from the investment. However, CGE models are data-intensive, costly to run and are difficult to critique due to their mathematical complexity. The question is whether the transport assets in such macro models could be specified with sufficient accuracy to account for improvements in dimensions such as travel-time variability and end-to-end public transport journeys that are relevant in many contemporary transport policy decisions.

Weisbrod (Chapter 4) cites two systems regional of economic models used by some US state transport departments for *ex-ante* assessment of projects. These CGE-type formulations do incorporate many detailed transport characteristics as inputs, and provide outputs of regional economic impacts. As an alternative to this system of models approach, one participant argued that the theoretical foundations of stand-alone CGE models could be re-built to include a more detailed representation of the transport system and its quality in consumer and producer functions. While this approach would be useful for understanding the effects of major transport-sector technological improvements, other participants argued that most transport projects under consideration are more marginal, so marginal appraisal approaches are sufficiently accurate.[2] Economic impact analysis methods were therefore considered most appropriate as a (non-additive) complement to CBA for major projects (or programmes of projects) where non-marginal effects on the economy are expected and where the expense of the modelling is justifiable.

Figure 1.2. **Scope of CBA versus economic impact analysis**



Source: Adapted from Weisbrod (Chapter 4).

Scoring methods, such as multi-criteria analysis (MCA) or scorecards, require an analyst to assign scores that assess the expected achievement of the project against pre-defined objectives. In the various US states, transport-related objectives can be inferred from transport departments' mission statements, with strong emphasis given to economic development, environmental improvement, mobility and safety (Volpe Transportation Systems Center, 2012). In MCA, objectives are generally weighted according to their priority for the decision maker (and to an extent their degree of overlap with other objectives). While guidelines are available to support MCA decision processes, economists are generally sceptical of its application due to issues of double-counting, arbitrary implicit valuations,[3] lack of viability threshold, and susceptibility to "gaming" (Dobes and Bennett, 2009).

Weisbrod (Chapter 4) argues that the level of government making the transport investment decision can influence the choice of appraisal framework. For example, national decision making in the UK has traditionally preferred CBA over CGE and MCA approaches as it focuses more on the overall efficiency of investments as a use of national funds, with less emphasis on the distributional impacts that approximately "net out" across winners and losers.[4] In contrast, where decisions are made at state or regional level, these local decision makers are very concerned with the potential for redistribution of economic activity into their jurisdiction, but are less concerned if these gains are made at the expense of other jurisdictions. CGE and other economic impact analysis models may therefore support this kind of decision making (though some participants noted that CBA can also be used to answer at least some of the questions about the spatial distribution of benefits). As some member countries devolve decision making from national to lower levels of government, such considerations may come further into focus (Mackie and Worsley, 2013).

The question of trust and empowerment was also raised in the context of the choice of appraisal framework. The choice between MCA and CBA can be viewed as a question of who the community prefers to make judgements on their behalf. A community that has strong trust that elected politicians and their planners will make trade-offs in the public interest might prefer the simplicity, comprehensiveness and immediacy of MCA. In contrast, a community that is suspicious that an MCA might be manipulated for political objectives that do not align with social welfare might prefer the reassurance of the more technocratic approach of CBA.

MCA can be useful for bringing together expected impacts that cannot be appropriately valued in CBA, such as irreversible heritage and environmental effects, alongside those effects that can be valued with CBA techniques (Weisbrod, Chapter 4). Several participants argued that in such instances, it is preferable to instead highlight these relevant impacts qualitatively alongside the CBA result, as is the practice in the UK's Appraisal Summary Table, rather than adding a false sense of precision with the integrated result of the MCA approach, particularly given concerns about double counting.

Drawing evidence from previous projects

CBAs of major urban transport projects usually estimate the project's future impacts using strategic transport forecasting models. Even with major advances in the modelling techniques applied in strategic models since they were developed in the 1950s, there are still many real-world behaviours of firms and households that are not reflected in these models and hence the CBAs that rely solely on them. A crucial missing element is the redistribution and reorganisation of businesses and households that might happen in response to a significant improvement in accessibility in a with-project case (the right-hand side of Figure 1.1). In other words, in standard CBA transport activity develops under the "fixed land use" assumption that ignores changes in "economic geography". But changes in economic geography are a major motivation for some transport projects, such as regeneration schemes or transit-oriented developments at new rail stations. In such instances, evidence on these changes is critical for:

- describing to decision makers the extent to which resources and activity may be redistributed across the economy, and what factors might influence these outcomes

- understanding the mechanisms and magnitudes of the economic geography changes that can give rise to very specific additional benefits not already captured in standard CBA (discussed in the final section of this chapter).

Roundtable participants discussed two strands of research that provide evidence from previous projects to explain the potential impacts of a proposed project: case studies and *ex-post* statistical analysis. Both types of approaches seek to infer relationships between transport improvements (the top of Figure 1.1) and their final impacts on the regional or national economy (the bottom of Figure 1.1).

Case studies

Weisbrod (Chapter 4) discusses the US approach to learning from previously implemented projects. Evidence is sought on the many economic and social effects that occur after different types of projects have been implemented, such as the development of industry clusters in areas with improved accessibility. Over 100 case studies are available in the *EconWorks* database hosted by the AASHTO, the association of state transportation departments. Across the cases, different forms of accessibility are found to be important to each industry sector. For example, professional services require a large commuter catchment and an international airport, while manufacturing businesses need to be able to access markets and suppliers in a one-day truck turn-around. Analysis of the wide range of projects and contexts also allows different types of clustering to be identified with respect to its supporting transport infrastructure.

One cluster was discussed at the roundtable in some detail: a linear automobile supply chain cluster along the I-65 and I-75 highways in Kentucky and Tennessee. Here, suppliers can traverse the full extent of the cluster by truck in a single day to allow "just-in-time" supply chains and "lean production" processes (Weisbrod, Chapter 4). The rural location simultaneously gives access to urban markets and to low cost local labour. One roundtable participant stressed that although this may be an interesting result of the transport infrastructure, it was not necessarily delivering a social welfare gain above alternative spatial patterns. It may even be possible that by allowing employers to exercise monopsony power in local labour markets there could be a social loss. This kind of granular insight is important to the narrative presented to decision makers. It illustrates the risks inherent in scaling up the results of economic impact analysis from the firm or specific area scale to drawing conclusions on socio-economic welfare changes from a national perspective.

Ex-post statistical analysis

There are significant difficulties in using case study evidence in *ex-ante* analysis of other projects. Chief among the challenges is the difficulty of separating the effect of a specific project from other factors that would help explain the development outcomes following a transport investment. Empirical *ex-post* analysis of US case study data attempts to control for these factors by asking the analyst to identify a counter-factual reference case (e.g. state-wide average changes over the study period) and to qualitatively attribute (less than 100% of) the observed changes in the study area (relative to the reference case) to the project (Weisbrod, Chapter 4). However, the data gathered and techniques recommended are generally not yet sufficient to do so in a statistically robust way.

A recent roundtable on *ex-post* analysis identified several challenges that need to be overcome to be able to robustly attribute observed outcomes to a particular transport project or initiative (ITF, Forthcoming). The first challenge is that, unlike controlled trials in health or education, the location that receives a "treatment" (a transport investment) is not randomly selected from a set of options. Instead, investments tend to be made precisely where for example congestion or crash rates are the worst, or economic activity is supressed. A second challenge in *ex-post* transport analysis is that each transport project takes place within a different transport network and socio-economic context. This means that an appropriate counterfactual or "control" cannot be identified or specified to allow for the effects of the project to be neatly identified (Worsley, 2014).

Several, but not all, participants at the present roundtable were optimistic that the data and tools required to delineate the effects of projects from other factors could be brought together in the near term. *Ex-post* data being collected – for example, in the US (Weisbrod, Chapter 4) and France (Bonnafous, 2014) – continues to expand in quantity and quality (though some participants were sceptical that the projects selected for analysis may not always be a neutral selection of successes and failures). Graham (2014) describes the use of statistical inference methods to remove the influence of "confounding effects" by simulating a random assignment of investments across alternative locations. These techniques have been applied with success in the *ex-post* analysis of road safety projects. "Crash modification functions" have been developed for different project types and contexts that allow safety impacts from a project to be estimated *ex-ante* (ITF, 2012). It may therefore be possible in the future to get to the same point for economic impacts from transport projects: the development of a range of "economic impact factors" for different project types and contexts that could be applied in *ex-ante* assessments.

Extending the toolkit and scope of accepted CBA practice

In northern Europe at least, the majority of effort by researchers to improve the usefulness of CBA has been to extend the framework, rather than to replace it with an alternative (e.g. UK, France and Sweden). This can be considered as either adding to, or improving on, the items in "direct resource benefits", "welfare benefits" or "other resource benefits" in Figure 1.1.

Several participants emphasised that the "burden of proof" required to incorporate additional effects in CBA is very high, with significant scepticism from national oversight bodies, such as Treasury departments. In practice, the inclusion of a "new" benefit to the accepted CBA framework requires researchers and practitioners to robustly demonstrate that the additional effect or benefit:

(1) is theoretically *additional* to other benefits captured in CBA

(2) can be *valued* in a way that is robust and does not overlap with the valuation of related effects that are already included

(3) can be adequately *forecast*, with and without the project intervention.

The most topical of the extensions to the CBA framework in recent years are those classified as "wider economic benefits" or "wider impacts". Before treating these in detail (in the final section), we will discuss reliability benefits (the next section), whose additionality is less theoretically contested (point 1 above), but are challenging to forecast and value (points 2 and 3).

Tailoring each CBA to the project's context and objectives

The discussion at the roundtable emphasised the diversity of possible impacts from major transport projects (and how these could vary by location) as well as the complexity of the tools that can be deployed to forecast and value these appropriately in CBA. A natural question arises: is it better to have a standard toolkit (or model) that is applied equally to every project assessed by a government agency, or should some parts of the toolkit be deployed only where the relevant benefits were expected to be significant for the project? The majority view on this question was that a scalable and modular approach was best in practice, though there were reservations to this view. This issue will be taken up further in the final section.

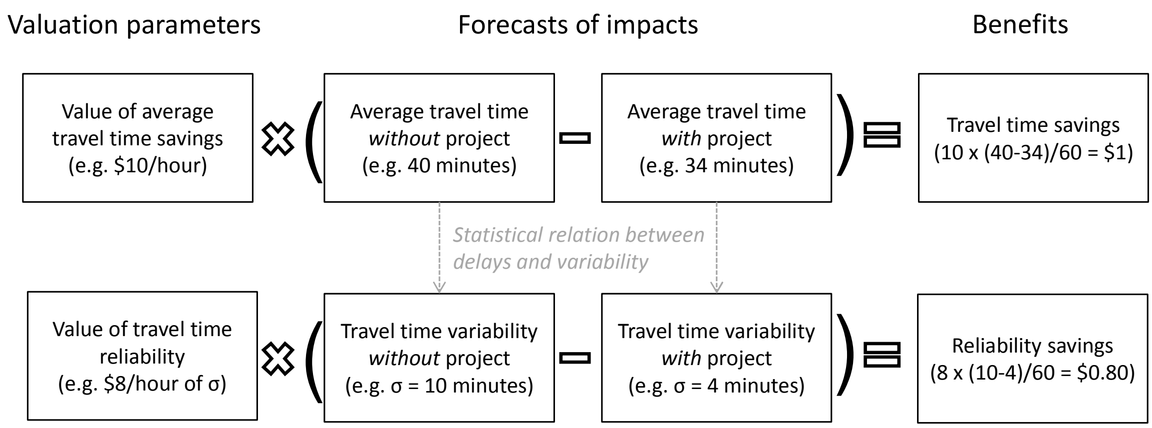
Incorporating reliability benefits in CBA

Travel-time variability is inherent in all modes of transport. This variability often leads to significant costs to travellers and the economy. One of the recurring themes of the roundtable discussions was the diversity of the experiences of travellers with unreliability: from the mild inconvenience of arriving unexpectedly late at a holiday destination, through to complete spoilage of time-critical freight or missed business meetings. Measures that improve reliability are therefore certain to be valuable to the community, so the ability to effectively include reliability benefits into transport CBA is paramount to ensure that projects that particularly improve reliability are properly reflected in project prioritisation.

The diversity of trip purposes and responses to variability make its measurement and valuation particularly difficult. Roundtable discussions centred on the three items required for reliability to be included in CBA, as noted in the previous section. That is, we need first a clear definition and measure of reliability that does not overlap with other items in CBA; then a unit value for the cost of unreliability (i.e. the benefit of an improvement in reliability); and finally an approach to forecasting reliability with and without a project intervention. These three points can be demonstrated with reference to an example (Figure 1.3). In the figure, additionality of average travel-time benefits and reliability benefits is ensured if the valuation parameters (on the left hand side) are distinctly estimated in the same study. The approach to forecasting future variability described later in the section relies on a relation between the average levels of delay or congestion and travel-time variability. In the example, reliability benefits are 80% of the magnitude of average travel-time savings.

The consensus view at the roundtable was that reliability is a critical dimension of transport system performance that should be included within CBA if a sufficient evidence base was available to value and forecast improvements. Implemented approaches range from relatively sophisticated valuation and forecasting approaches (e.g. the UK and France), to percentage mark-ups on the average travel-time savings (e.g. currently in the Netherlands) to excluding it altogether. Outside the US, where the focus has been on freight reliability, much of the research focus has been on road transport for passenger travel. Extensions of the research to cover public transport were also discussed at the roundtable.

Figure 1.3. **Example of approach to estimating travel-time savings and reliability benefits**



Defining reliability

The definition of travel-time reliability is not straightforward compared with the definition of average travel time. While the latter is widely understood by the public and is a concept routinely estimated by strategic transport models, the same cannot be said for reliability. Three elements of the definition of reliability were discussed at the roundtable: the nature of unreliability, travellers' perceptions of this unreliability, and responses to unreliability.

The nature and causes of unreliability

Data on travel times in many different contexts demonstrates systematic variability across the day, week and year (ITF, 2010). These variations are generally due to widely understood peaks and troughs in passenger demand relating to work and education schedules. On a given day, other variations in travel times can be due to somewhat regular occurrences, such as traffic signals, rain, crashes or network maintenance. A third type of cause is what can be called "extreme events", such as flooding, severe incidents or network closures (de Jong and Bliemer, 2015). Roundtable participants were somewhat divided on whether a hard distinction should be drawn between more routine types of travel-time variability and that caused by extreme events. Put another way, the question is whether to separately treat outliers (very long travel times) of travel-time distributions when we estimate our reliability metric (or metrics).

How travellers perceive unreliability

In distinguishing between average travel-time savings and travel-time reliability savings, we are necessarily allowing travellers some imperfection in their *ex-ante* knowledge of travel-time outcomes. At one extreme, with perfect knowledge of their travel times for upcoming trips, transport projects could produce only average travel-time savings since there would be no *ex-ante* uncertainty about travel time. In practice, travellers will have some knowledge of the likely travel time for their upcoming trip, but they will recognise there is some risk of the actual travel time being longer than this "likely" time. That is, travellers implicitly have a distribution of travel times in mind for their trip *before the trip is completed*. The discussion at the roundtable raised a number of questions about when these expectations are formed and what information they are based upon. Two main sources of information about upcoming travel times are:

- personal experience travelling on the same route or service

- external information about current travel conditions, such as traffic reports, "smart" navigation on smartphones or GPS units, real-time information signage en route, or even weather reports.

The difficulty in defining an *ex-ante* measure of reliability is that different people will be armed with widely differing information about the likely travel time associated with essentially the same trip. For example, two people may be about to travel on the same stretch of highway departing at 8:20 am on Tuesday. The first is from out of town and has only a paper map to help direct him to his destination. The second person regularly commutes along the highway, has reviewed the traffic incident reports over breakfast and she found yesterday that the week-long strike on the rail network meant that travel times were higher than usual for a Monday. The challenge is to relate the expectations from such diverse travellers to observable information about travel-time variability, as well as available information about how travellers value improvements in reliability.

Responses to unreliability

A further aspect that is relevant in defining reliability is the ways that travellers may respond to unreliability. One response to *expected* unreliability of an upcoming trip is for a traveller to build in an additional time margin above her expected (point estimate) travel time and depart slightly early (Fosgerau, Chapter 2). People may choose to change their departure time altogether to a period where travel times are more predictable. Travellers with repeated experience in using a route or service that find it to be unacceptably unreliable may change route, service or travel mode. Even once the trip is underway, if a traveller finds the trip is taking longer than was expected, they can sometimes change route – especially if they have real-time information (either from road-side signs or an on-board device).

People and businesses may make more fundamental changes in response to unreliability (a change in "economic geography" discussed in the final section below). For example, if a regular commuter finds travel times too unreliable, she may move house to be closer to her workplace or closer to a more reliable transport node, such as a rail station. Enterprises have an even wider range of choices in response to unreliable transport services. While initially, late-running deliveries can result in service penalties and driver overtime costs, freight companies may decide to purchase additional fleet (and pay additional drivers) to meet customer service levels at a given level of network reliability. More fundamentally still, businesses may choose to reorganise their entire supply chains and production processes to be more local (Weisbrod, Chapter 4), though some have argued there is not much evidence to support these changes taking place (McKinnon et al., 2008). This kind of reorganisation would trade off the economies of scale from fewer production points against the high costs of transport (including unreliability).

Selecting a reliability metric

The discussion at the roundtable suggested there was no single correct metric for reliability to apply in CBA. Any approach taken would either have to be highly disaggregated to reflect travellers having different information and trip purposes, or must involve strong homogenising assumptions about information, preferences and behaviour.

In practice, the metric currently typically applied in CBA is the variability (generally measured by the standard deviation) of a distribution of origin-destination travel times that is implicitly known to the traveller *ex-ante*. The traveller selects her departure time in advance (when the travel-time distribution is estimated or "observed" by her). In the current practice, the mode, route/service and the origin/destination of the trip are assumed fixed. No approaches to valuation or demand forecasting are yet available that account for (some) travellers incorporating real-time "unexpected" reliability information (while the trip is underway) into their route choices. The traveller may use information about "expected" variations (time of year or day) to choose their departure time but once she is "on the road", is assumed to be an otherwise uninformed and passive recipient of travel-time outcomes.

Valuing reliability

Fosgerau (Chapter 2) outlines a theoretical structure for the rational decision making of a traveller faced with an uncertain travel time for her upcoming trip. The models are in the scheduling model tradition of Vickery (1969) and Small (1982) where utility from the trip depends on both the mean and standard deviation of travel time. This definition is important as it shows how the traveller will value reductions in this variation, even where the expected travel time (i.e. the mean travel time) is unchanged, so the reliability measure is conceptually distinct from travel-time savings captured elsewhere in CBA. The two variants are argued to be applicable to different types of trips:

- The step model, where the traveller attributes a high utility to being at the destination by a specific hour (but not otherwise), is relevant particularly for travel to employment with a fixed start time or travel to appointments. Travellers described by this model will be motivated to include a travel-time margin by departing earlier than would be implied by the average travel time and their preferred arrival time.

- The slope model, where travellers have more gradually shifting preferences to be at the destination rather than the origin, is more relevant to travellers for whom the specific arrival time is not critical, such as leisure travel or travel to employment with a flexible starting time. Travellers described by this model would not incorporate a travel-time margin when deciding their departure time.

Mathematically, each alternative model implies a different metric of variability that should be used for valuation. The step model gives a valuation in terms of standard deviation (and Fosgerau (Chapter 2) shows that the model is compatible with several alternative measures of dispersion, such as the mean lateness, if the shape of the travel-time distribution is fixed). The slope model by contrast leads to a valuation in terms of variance. This distinction can be important in the practical application, as variance is simpler to sum across composite links in a journey to get an aggregate variability measure.

Valuations for average travel time and travel-time reliability are generated by fitting model equations (either in structural form or reduced form) to a dataset. These are typically thought of as the costs internalised by the traveller themselves, but other parties (e.g. a fellow meeting participant) may also incur costs (Fosgerau et al., 2014). The two types of datasets that can be used in the empirical estimation are stated preference (SP) drawn from survey responses to choice exercises and revealed preference (RP) data drawn from observed travel outcomes. These approaches have strengths and weaknesses that were debated at the roundtable (Table 1.1.).

Fosgerau (Chapter 2) highlights some fundamental challenges to the rationality assumptions underlying the scheduling models when they are used to estimate the value of reliability with SP data. While these were acknowledged, not all roundtable participants agreed with his suggestion that SP-based valuation should be abandoned in favour of RP methods. What participants did agree on is that RP methods should be pursued with renewed vigour as the amount of high quality RP data available to researchers expands rapidly with the growth of smartcard public transport ticketing systems, GPS units in vehicles and smartphone location data. There may also be opportunities to draw on the strengths of both data sources through combined SP and RP estimation (Ben-Akiva and Morikawa, 1990).

Table 1.1. **Approaches to valuing reliability benefits**

Approach	Strengths	Weaknesses
Stated Preference (SP)	• Relatively low cost to gain large sample • Control over the range of variation • Repeat observation of each respondent allows control of personal/local influence on choices • Can frame questions to align with assumptions (e.g. *ex-ante* known distribution)	• Respondents may not be able to process the questions effectively to make choices reflecting true behaviour • Different valuations for reliability depending on whether structural or reduced-form estimation • Very difficult to use for freight sector – which part of the supply chain to interview?
Revealed Preference (RP)	• Observing real behaviours and decisions • More comprehensive coverage • Data increasingly available • Feasible to provide valuations for freight sector (though these may understate true values)	• Requires huge amounts of data (and processing power) to generate valuation results • Can be difficult to delineate the effects of average travel time and variability on travel choices • Can be difficult to introduce monetary dimension

Source: ITF based on roundtable discussion and Fosgerau (Chapter 2).

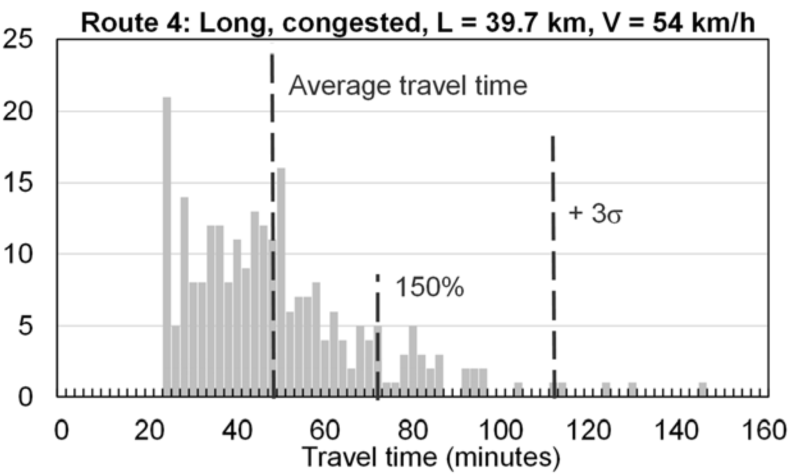
Extending the valuation approach from the case of passenger car travel is likely to be achievable with further research. In the case of public transport service reliability it may be that, as with *average* travel time, travellers may experience disproportionate disutility from *variation* in the components of the journey (wait time, in-vehicle time) and conditions (seating versus standing). For instance, there may be greater disutility from unexpectedly standing on a bus for 10 minutes, compared with 10 minutes extra on-board in a seat. Valuation for freight trips was seen by some participants as actually easier to model and estimate, given the incentives are clear-cut. However, the issues with freight are perhaps more fundamental. Freight operators are likely to be well informed about levels of variability, and so are likely to have adapted their operations in advance of the "marginal trip" being considered (in either an SP or RP context). The costs of unreliability (and hence the benefits from unreliability) will come more from the reorganisation of operations, the changes in fleet and staff levels, etc. that will be undertaken to best provide the required level of service to their customers. The marginal CBA framework will not easily be extended to capture such costs, yet they are likely to be large.

Forecasting reliability

Coherently measuring, let alone forecasting, the *ex-ante* travel-time distributions faced by road users between point A and point B is challenging. As described above, for any departure time on any given day, each traveller would have their own *ex-ante* travel-time expectation, which might be a range or a point estimate, is likely to vary considerably from person to person, and is not observable. *Ex post*, a single travel time can be observed from number-plate recognition systems or loop counters.[5] However, in the absence of any empirical evidence to develop a theoretical model of the formation of expectations, any approach needs to relate the real observed data to the expectations of the travellers. The approach of Kouwenhoven and Warffemius (Chapter 3) is to assume travellers form a time expectation based on times observed at the same time of day in the recent past (and even incorporating values from the near future). The *ex-ante* unreliability perceived by the traveller is assumed to align with the *ex-post* travel-time distribution (around this expected value).

In developing *ex-post* distributions in 15-minute time slices Kouwenhoven and Warffemius (Chapter 3) exclude selected travel-time observations as "extreme event" outliers as recommended by de Jong and Bliemer (2015) (Figure 1.4). This approach polarised participants: should extreme events be captured by a separate valuation and forecasting framework or within a single framework? In the Dutch case, the argument is that the valuation study, which used SP surveys, did not cover extreme delays. Moreover, the speed-flow curves that are the base for the travel-time calculations in the Dutch transport model do not include extreme events either. For Kouwenhoven and Warffemius (Chapter 3) then, the focus is on retaining consistency, rather than ignoring these real and important observations. Including the extreme events in this or in a separate tool is one of the longer term improvements proposed for the Dutch post-processing reliability model. In the meantime, a roundtable participant suggested retaining the outliers in their sample and applying an alternative measure of dispersion as their reliability metric (and later transforming the results to accord with the standard deviation in the valuation study).

Figure 1.4. **Example travel time histogram**



Note: Extreme events excluded three standard deviations above the average

Source: Kouwenhoven and Warffemius (see Chapter 3).

In most contexts, the only basis for forecasting *ex-post* unreliability is by relating it to outputs from a strategic transport model. Since these models are generally based on equilibrium travel times across routes in a given time period in the future, their outputs are point estimates rather than distributions. While a delay on any given day may be completely unrelated to traffic flow and speeds, the principle is to understand the systematic part of unreliability that can be predicted. This further supports the removal of extreme event observations, where delays are likely to be completely unrelated to traffic volumes.

Kouwenhoven and Warffemius (Chapter 3) relate their observations on unreliability to contemporary observations on average travel times. There is a wealth of international experience and options for functional forms for this exercise. Roundtable participants had some suggestions to improve the logical coherence and fit of their dataset, such as the importance of standardising mean delays by the route length, but fundamentally it was recognised that this approach is best practice with the current tools available. A strong positive relationship was found in the Dutch motorway data between the mean delay on any given origin-destination journey and the variability of that journey, though the specifics of the relationship varied depending on whether data was regressed across routes or across time periods on a given route (with the former approach strongly preferred by the authors).

Equipped with an empirical relationship between average travel times (and other outputs of the strategic model), researchers are able to make predictions about how reliability may improve if transport interventions affect the nature of the routes and the associated mean delay. Further investigation is required to understand how to forecast the reliability impacts of projects such as ramp metering or bus lanes that specifically target unreliability (perhaps at the expense of average travel times). Roundtable participants suggested that, in general, projects that *indirectly* improved reliability tended to represent a movement along the same trade-off curve between reliability and mean delay, but that projects specifically *targeting* reliability may involve a shift in the curve. One example is a ramp metering project implemented on the A6W motorway near Paris, which suggested reliability benefits were around the same magnitude as travel-time savings (Bhouri and Kauppila, 2011). Further *ex-post* evidence should give greater insights into the incorporation of impacts from such projects.

By combining valuation parameters and a forecasting model in a test undertaken for the Kouwenhoven and Warffemius chapter (Chapter 3), reliability benefits were estimated to range from 15% to 60% of travel-time savings across three projects. This suggests that even though a relatively straightforward relationship between average travel times and variability applies, the overall approach is responsive to other dimensions of the projects. It also underscores the importance of capturing the diversity of project impacts in a more sophisticated way than is possible with the simple (25%) mark-up approach currently in place in the Netherlands.

Ways forward

This section discussed the methods currently being applied that, while not perfect, represent a major improvement compared to excluding reliability benefits (or applying simple travel-time savings mark-ups). The discussions at the roundtable highlight some of the limitations and potential ways forward to continue to improve the theory and practice:

- Learning about expectations – the rationality challenges identified by Fosgerau (Chapter 2) still need to be resolved before practitioners can be confident with reliability valuations from SP studies. The role of information availability (e.g. real-time traffic reporting) is also likely to have a major role in understanding the real implications of unreliability for travellers.

- Theoretical underpinning – several participants suggested that the confidence in reliability forecasting would be improved by developing a theoretical basis for the interactions between transport network (supply) and travellers (demand), i.e. by developing physical and rational behaviour relationships.

- Incorporating demand feedback – a specific limitation identified at the roundtable and by de Jong and Bliemer (2015) was the absence of feedback between the reliability expectations and behaviour (beyond scheduling). Reliability is certainly a part of a traveller's generalised cost, so how does route choice, mode choice and even origin/destination choice respond to different levels of reliability?

- The role of reliability in supply chain reorganisation and residential location – understanding the non-marginal responses of households and firms to different levels of reliability will certainly be useful in constructing the strategic narrative for reliability improvements.

- Disaggregation and extension – the discussion above and at the roundtable highlighted the diversity in the types of travellers, their trip purposes, and their access to (and use of) real-time information. These dimensions will affect the way that travellers perceive and respond to unreliability, which will ultimately affect valuation and forecasting. So future work will need to focus on disaggregating into relevant groups to improve the accuracy of inputs into CBA. Approaches will also need to be extended to capture freight and public transport to ensure completeness and comparability of CBA results across project types.

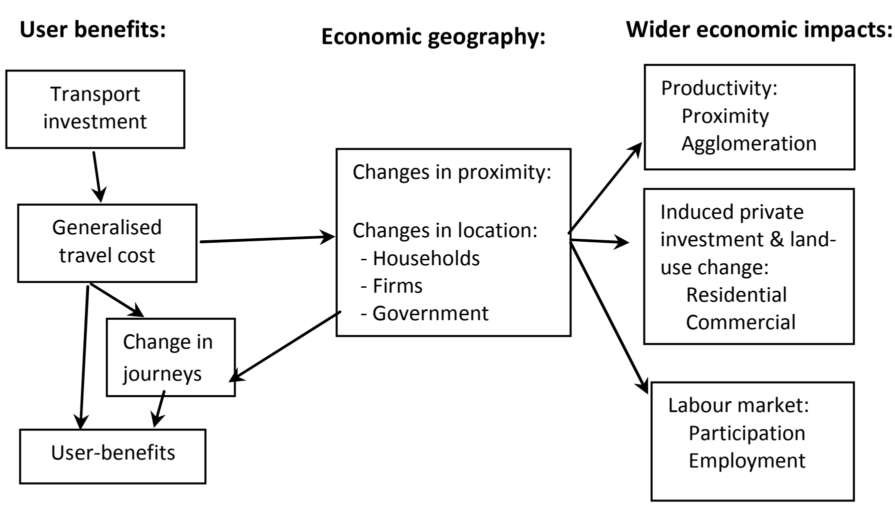
- Sharing experience and data – several roundtable participants voiced an interest in closer international collaboration of researchers to share techniques, data and results. The roundtable itself was seen as a useful step in this direction.

Several of these areas for improvement are somewhat long-term research projects, but the importance of reliability to travellers and the economy generally mean that they are worth targeting so that projects with an emphasis on improving reliability receive appropriate priority. In the meantime, the consensus view of the experts at the roundtable was that the state of practice as it currently stands is sufficiently well developed to start including reliability in CBAs of major transport projects.

Incorporating wider economic impacts in CBA

The previous section described how reliability benefits are difficult to include in CBA because of practical challenges with valuation and forecasting. In this section the focus is on wider economic impacts, which are difficult to include in CBA because their existence is due to violations of the working assumptions of standard CBA. This raises conceptual issues of additionality on top of the practical challenges of valuation and forecasting.

Figure 1.5. **Framework for separating user benefits from wider economic impacts**



Source: Adapted from Venables (Chapter 5).

Venables (Chapter 5) proposes a framework for conceptualising the interplay between the direct user benefits of a transport improvement and the wider economic impacts (Figure 1.5). The left hand group of impacts are estimated using demand forecasting techniques; impacts are valued using standard welfare economics approaches. Conceptually, these direct user benefits account for the full welfare improvement from a marginal transport improvement provided industries have constant returns to scale and markets are complete and perfectly competitive.[6] The past decade has seen a gradual incorporation of insights and techniques from "new economic geography" (among other fields) into more advanced transport CBA frameworks. The central box in the figure can be characterised as changes in economic geography or land use, which are often ignored by strategic transport models, and hence standard CBA. The recognition of real-world violations of the standard CBA assumptions (e.g. increasing returns to scale and market imperfections) and the changes in economic geography arising from a transport project open the possibility that there are changes in economic welfare not captured by direct user benefits (and external costs). These potential benefits are identified in the right hand column of the figure as wider economic impacts.

Changes in economic geography

Changes in accessibility – whether defined in terms of average road travel times, improved transport reliability or an expansion of public transport options – will encourage some households, firms and government departments to re-think their location decisions. For example, a household might choose to locate closer to a new rail station to improve their commuting options or avoid the need to own a second car. A member of a different household may decide to commute further to a higher paying job which is now within a 45-minute trip following the upgrading of a road. A logistics firm may respond to improved road travel times or reliability by relocating to lower cost land that allows them to service a large market in a single day round-trip (Weisbrod, Chapter 4). Other firms may reorganise their internal activities, moving some functions (e.g. relatively low skilled work) to lower cost areas away from the city centre if connectivity improves (see for example section 5.3.3 in de Rus, 2009). These kinds of microscopic adjustments within decision-making units are increasingly illuminated by *ex-post* analyses (such as Weisbrod (Chapter 4)) but are not generally integrated into strategic demand forecasting models. They may occur with significant time lags and be difficult to confidently attribute to particular transport improvements.

In any case, roundtable participants and case study evidence emphasised that private decisions of households and firms do not always relate directly and predictably to the transport improvements (Weisbrod, Chapter 4). The effects of transport on economic geography depend also on:

- local governments, whose plans and rules constrain development (land available for residential or commercial uses, or maximum building densities)
- local communities, who can influence local government land use decisions by protesting proposed developments or insisting on local improvements as part of plans to accommodate extra activity
- property developers, who ultimately need to make early and sometimes risky investments to provide floor space in large residential and commercial buildings.

Regeneration or new cluster development is likely to require co-ordinated actions beyond the transport investment. Actual land use change outcomes from a transport project are therefore difficult to predict, and nearly impossible to model without a strong knowledge of local conditions. The US case study approach includes interviews with many of the important actors (such as developers and business chambers) to understand this bigger picture (Weisbrod, Chapter 4). The interdependencies inherent in land use change require the use of scenarios (e.g. "high" scenarios where actors co-ordinate well and a new cluster flourishes, against a "low" scenario where little new activity is generated) and evidence on magnitudes of changes in any scenario. Two evidence sources were discussed: bottom-up local information and formal modelling.

Venables (Chapter 5) argues for the use of "bottom-up" local information, such as planning documentation (e.g. about the number of additional dwellings proposed), interviews with stakeholders and technical information about the transport scheme (e.g. additional peak-hour commuting capacity) to understand the likely location, scale and timing of the changes in employment and population with the project. *Ex-post* evidence from similar projects would clearly be a useful evidence base in developing these local plans. While this approach may be appropriate, particularly in smaller schemes with localised impacts, some at the roundtable thought that the approach could suffer from biases and inconsistency in the appraisal.

Formal modelling, such as a land use transport interaction (LUTI) models, seek to integrate land use responses into a strategic transport model across a large spatial area. Unfortunately these models are expensive, complex and their predictions have not been extensively tested (Worsley, 2014). Conceptually though, LUTI models provide much richer information than bottom-up approaches since, for example, many small displacements of activity from various areas to the project area can be clearly represented (Venables, Chapter 5). Several participants noted that progress is being made in the development and application of LUTI models, for instance in the appraisal of the Grand Paris transport and urban development project (de Palma, 2014).

Market imperfections and externalities

Changes in economic geography can simply reflect the direct transport impacts on users rippling through prices (especially land prices and rents) and quantities across the economy. If there were no external costs or other market imperfections, these spatial reorganisations would be of interest to decision makers and planners, but would not affect the overall results of the CBA (Venables, Chapter 5). In the case of the logistics firm that joins an emerging cluster, it would move to a more distant location on a new highway, adding to demand in one origin-destination pair (a "new user"), and subtracting from demand in another pair (a "lost user"). So *if this shift can be forecast accurately*, the social gain will be measured through the "rule of half" in standard CBA (Venables, Chapter 5).[7] It is arguable that standard forecasting models do not fully capture the range or extent of these responses, so there may be a need for some "out of model" adjustments to adequately reflect future outcomes.

The difficulties of directly measuring the aggregate economic benefits of a transport improvement using *ex-post* evidence, particularly at a project level were explained in the second section of this chapter. Instead of attempting this, Venables proposes to continue to rely on CBA's measurement of direct user benefits, and supplement these with wider economic impacts where specific market imperfections can be identified and robustly quantified. Three main types of market imperfections were highlighted at the roundtable:

- Agglomeration economies, where the total social gains from additional workers in an industry relocating to the same cluster are greater than the gains made by the relocating firm or workers themselves. The cluster as a whole benefits from greater opportunities to interact, access to a thicker labour market, and improved opportunities for labour and firms to specialise, co-ordinate and co-operate. These effects continue to exist, even as information technology makes some kinds of face-to-face contact unnecessary.[8]

- Taxes on labour (and government unemployment payments) that mean that perceived private benefits from working (net of tax paid and government payments withdrawn) are lower than the social gains from an increase in labour supply (i.e. there is a tax wedge since the additional tax revenue is valued by society).

- Price-cost divergences, due to market power (e.g. of property developers), government restrictions (e.g. planning controls) or economies of scale in the provision of services and utility networks. These divergences mean that transport improvements that lead to increased land development in some locations can generate additional social gains. Transport may also lead to social gains if the volume of trade increases (or market power is reduced) in sectors where prices diverge from marginal costs.

These imperfections mean that transport improvements *can* give rise to net benefits to the national community beyond the direct user benefits.

Benefit identification and valuation

Venables (Chapter 5) emphasises that the market imperfections described above are a necessary, but not sufficient, condition for identifying the wider economic benefits to society from transport improvements. We must demonstrate the extent to which employment or investment induced by the transport improvement would not have occurred elsewhere in the country in the absence of the project. In other words, we need to show that the activities induced by the transport project are of higher social value than the activities that would have happened without it (Venables et al., 2014). Venables argues that in developed economies, an appropriate starting assumption is that 100% of new activity in a specific area affected by a transport project would have occurred elsewhere in the counterfactual case.

Demonstrating, valuing and forecasting wider economic impacts requires a strong evidence base, in addition to a robust theoretical framework, before they can be included in a CBA. The level of maturity of this evidence base differs across benefit type and across member countries. The state of practice for three particular benefits was surveyed by Venables and roundtable participants.

Agglomeration economies

The approach to quantifying agglomeration benefits from a transport improvement typically involves two stages: (1) developing an empirical relationship between a worker's productivity and her access to jobs (or, more technically, access to "economic mass" or "effective density"); and (2) the extent to which a transport project improves this access to jobs.

Elasticities are estimated based on cross-sectional data that demonstrate that larger industry clusters and cities have higher output per worker. Substantial effects have been found: estimates suggest that productivity in a city of 5 million is 12% to 26% higher than a city of 500 000 (Venables, Chapter 5). To be useful in transport CBA, elasticities need to isolate the incremental component of productivity that is *caused by* better access. Substantial progress has been made in recent years to improve this relation by controlling for other factors that explain productivity differences, such as more skilled workers themselves choosing to work in larger cities. Other confounding variables still need to be addressed, such as public transport network quality, which is correlated with cluster size but may not be fully captured in the access metric used in the empirical work. *Ex-post* project evidence may be helpful in identifying the incremental effects of transport in this context. Several participants also cautioned against transferring elasticities from one context to another due to the range of context-specific factors that are embedded in results. Nevertheless, there was broad recognition that the techniques available for estimating agglomeration elasticities were relatively mature, even if the ideal data is not available in all countries and contexts.

In the UK in particular, a distinction is drawn between "static clustering" and "dynamic clustering" (Venables, Chapter 5). Static clustering is where transport improvements increase access to and within an employment cluster, even if employees do not shift their home or work location. Static clustering is relatively easy to estimate using the standard (fixed land-use) outputs from the strategic transport model. Even in this case, there are complications that arise in practice, such as the exact definition of accessibility in a multi-modal context. For example, if a new public transport mode is introduced, is the relevant measure of accessibility the forecast, mode-weighted average travel time, or a relation of the "logsum"[9] of time savings (that captures the option value of additional modes)?

Dynamic clustering is an intensification of an employment cluster when workers (or firms) relocate in response to improved access. This intensification brings both positive (e.g. productivity spillovers) and negative effects (e.g. congestion) for the cluster that need to be accounted for. This relocation of jobs and firms may materially weaken the original cluster relative to the without-project case. Venables (Chapter 5) cites the case of London's Crossrail project, which seeks to increase the size of the financial services cluster in the City of London, and argues that without the project these workers would not have worked in an alternative cluster, so the agglomeration effect is unambiguously positive in this particular case.

Labour market effects

If a transport improvement encourages workers to either switch to a more productive job (e.g. one that is further from home in the city centre) or to increase their labour supply (working additional hours) then there is a *gross* increase in the value of output produced. In welfare terms, the only net gain is the improvement in the government budget balance. When workers trade off income, leisure and travel costs, they make decisions based on post-tax income gains, while their output is valued at the pre-tax amount paid by their employer. This "tax wedge" (that also reflects employment benefits withdrawn) is the relevant valuation parameter. Although the tax wedge will be different for each economy and the mix of the employees that expand their labour supply, it can be estimated based on readily available information.

Perhaps of more interest to policy makers is the question of whether a project induces genuinely additional economic activity and hence wholly new jobs, either during construction or operations. The benefits from a worker moving out of *involuntary* unemployment will be very large and certainly important for that person. However, Venables (Chapter 5) argues that it is likely that in most developed economies the labour required for a transport project is most likely to be drawn from alternative employment (so of no additional social value) or from people working additional hours (captured by the tax wedge calculation described above).

Induced property investment

While the two previous groups of benefits are well established in CBA guidelines and practice in several countries (e.g. the UK and New Zealand), Venables (Chapter 5) considers some additional areas where wider impacts could arise due to transport projects inducing property investment. Two types of situations where induced investment might be associated with social gains are considered. The first is dependent residential development, where a transport improvement allows planning restrictions to be relaxed in a situation where prices are below marginal social costs. The additional benefit is proportional to the price-cost gap.

Economies of scale in utilities and social service delivery can create large price-cost gaps that should be reflected in some cases. In Sydney, Australia, for example, the relative resource costs of accommodating incremental dwellings in a medium density land use pattern have been found to be significantly lower than for low density development (CIE, 2012). This can be due to scale economies in network infrastructure (making use of any trunk capacity rather than extending networks) or environmental and health savings from less car-dependent travel patterns. To an extent households pay a price for housing and service delivery that does not reflect the full social costs of alternative development patterns (Langer and Winston, 2008). Transport projects that encourage lower social cost development patterns can give rise to resource cost savings that can be attributed to the transport project. Venables (Chapter 5) argues that if the planning change and transport improvement are both necessary (but not sufficient) it is not possible to allocate the benefit between the two policies.

The second benefit considered by Venables (Chapter 5) is large-scale retail developments.[10] Developers often have market power that gives rise to a price-cost gap, so an expansion in floor space will give rise to similar proportional benefits as in the residential case above. Alternatively a price-cost gap may emerge due to co-ordination failure, whereby the success of a development would only be assured if other actors (households, developers or firms) invested in the same location. Transport investment can focus investment on a specific location to break the deadlock. In the retail case, Venables

(Chapter 5) argues that the scale of developments may also give rise to non-trivial expansions in product *varieties* that, when taken together, make the entire retail area more attractive (a specific form of agglomeration benefit).[11] The extra varieties generate greater consumer surplus than would be the case if the same floor space were distributed more evenly. (Of course, close attention needs to be paid to the loss of varieties in the alternative locations to isolate the net effects.) Venables argues there is empirical evidence to help make the valuations as a proportion of floor space or additional local expenditure that may be available from local planning data.

A modular versus comprehensive approach

Venables (Chapter 5) and a number of other attendees argued for a proportional and scalable modular approach to CBA. This is because the available “single model” tools, such as LUTI models, can be expensive to deploy and cannot yet satisfactorily forecast all relevant impacts. So it makes sense to have an appraisal framework that allows smaller projects, or projects with straightforward effects, to be appraised with a targeted set of tools. Potentially “transformational” projects may need a substantially different set of tools which respond to the economic context (e.g. in a developing economy where access to markets is highly constrained) and the anticipated impacts of the project (e.g. regeneration of a depressed local area). This requires a clear overarching appraisal framework, good judgements by practitioners on which effects are relevant, and strong quality control, probably at a national level.

The alternative position was argued by some participants. Providing (well resourced) project proponents with the discretion to choose their tools, they would attempt to “game the system” by including every benefit that added to the assessed net present value; less well-resourced project proponents would be at a disadvantage in project rankings. A common approach would provide fairness and direct comparability of results in project prioritisation. Some participants also cautioned on the need to ensure a consistent theoretical perspective, such as on the nature of market competition, across all modules. This is required to avoid double-counting of benefits.

Case studies and academic research continue to uncover new market imperfections and effects of transport investments. Researchers, practitioners and oversight bodies need to be open to the addition of new and case-specific benefits if they can be rigorously demonstrated. At the same time, there should be an on-going healthy scepticism of the application of complex modelling to estimate a wide range of project benefits, particularly when the motivation for doing so cannot be clearly traced back to the project objectives.

Ways forward

This section discussed the approaches to forecasting changes in economic geography and identifying the resulting (potential) social gains that are not already reflected in user benefits in CBA. A number of conclusions can be drawn from the discussions at the roundtable.

The approaches to estimating agglomeration benefits and labour supply benefits are relatively mature (though the depth of valuation evidence available varies across countries). Further research work, particularly on *ex-post* identification and attribution of the effects of transport access improvements to productivity and labour supply, will help to improve the practice and will allow such effects to be rigorously challenged. But policy makers should take comfort from the fact that if properly applied, the current evidence base provides a sufficiently rigorous basis for including these benefits within the core effects of major transport projects. The size and importance of these benefits means that the alternatives of either excluding them or applying simple mark-up rules were considered unacceptable to roundtable participants.

For any wider impact under consideration, the analyst needs to identify a specific departure from the perfect markets assumption in the core CBA and the source of any "new" activity with the project. Venables (Chapter 5) provides such a justification for some aspects of transport-induced development to be included, though further scrutiny is still required. In the absence of a strong and scrutinised case for the inclusion of wider benefits of any given project, the long-standing position that user benefits adequately account for the whole economic impact of a project should remain in place.

Perhaps the largest challenge is in forecasting changes in economic geography. In this task, participants emphasised the importance of conducting scenarios (even in formal modelling tools like LUTI models) as there is no way of confidently predicting the outcome when coordinated actions among actors are required, such as in property development. Clearly, forecast changes in economic geography need to flow through consistently in all parts of the CBA, e.g. having the same employment location profile for estimating office development patterns and labour supply changes.

The choice between an appraisal using the "user benefits plus wider benefits" approach or using a "big model" (such as SCGE or LUTI model) currently favours the former, mainly because no big models are yet able to adequately capture all the relevant impacts of a transport project. No single strategy works for all projects yet, so decision makers are given partial information from several inconsistent frameworks. However, given the big models, at least in theory, can produce answers to the most sought-after questions, these may ultimately be better placed to address the "relevance" criticism of CBA described in the opening section of this chapter. So there may come a point where the big models become responsive, accurate, and cheap enough to apply as the preferred project appraisal approach. Most roundtable participants though were of the view that that time has not yet arrived.

Acknowledgements

The author is grateful to the rapporteurs, Chair (Jonas Eliasson) and other participants at the Roundtable for their insights and open discussion. Helpful comments and suggestions on this paper were received from Peter Mackie, Marco Kouwenhoven, Stephen Perkins, Jari Kauppila, David Meunier, Dejan Makovšek, Jagoda Egeland, Pim Warffemius, Phil Manners, Mogens Fosgerau, Mark Ledbury, Emile Quinet and Clifford Winston. All remaining errors are the responsibility of the author.

Notes

1 The Eddington Transport Study in the UK found that "accounting for social and environmental effects tends to increase the relative returns of public transport interventions", supporting the possibility of such biases affecting prioritisation if a wide set of benefits is not considered (Eddington, 2006: p. 139).

2 An alternative suggestion was to use these same foundations and to estimate relationships between transport quality (e.g. road congestion) and EIA metrics (e.g. employment) in reduced form. One participant identified the issue of attribution with this approach, that is, can the impact of each project on transport quality be identified? If not, it is not possible to use such an approach for project prioritisation.

3 Careful *ex-post* analysis of decisions made can reveal the values implicitly ascribed to each increment in score on the scale for each objective (Nellthorp and Mackie, 2000).

4 Indeed, Weisbrod (Chapter 4) notes that when the national government in the US began investing in infrastructure after the financial crisis, it relied heavily on CBA to prioritise projects.

5 Notwithstanding the moderate variation in times across particular drivers or vehicle types.

6 In addition to these, relatively straightforward departures from the latter assumption (external costs of noise, pollution and crashes) have been included in standard CBA techniques and guidelines for decades.

7 The logic of the rule of half is that the maximum private gain that could be made by a firm that moves location (including lower land costs) after a project is undertaken is the reduction in trip costs (otherwise they would have already moved); the minimum gain that would induce a firm to move location converges on zero. The average of these two extreme cases is *half* the reduction in trip costs.

The total gain by "new users" from the lower trip costs is therefore conventionally estimated as *half* multiplied by the *reduction in trip costs* multiplied by the *number of new trips*.

8 Larger populations in cities can also be associated with greater demand for high-quality services, which further improves the attractiveness of the city to mobile (high-skilled) labour, which can further enhance the productivity spillovers.

9 Where a project involves introducing a new mode of transport into an area, it is not possible to apply the simple rule of half algorithm to estimate user benefits as there is no "base case" travel time defined for that mode to compare the "project case" travel time with. Instead, user benefits must be calculated directly from the mode choice module of the demand model using a logsum. The logsum gathers together the overall utility from a set of options defined in part of the mode choice "tree". Providing travellers with a new travel option will increase the utility of the bundle, and some (but not all) of whom will be expected to select the new mode (Williams, 1977).

10 He argued that a highly analogous benefit may emerge in the case of large-scale office development.

11 Venables (Chapter 5) emphasises that the varieties approach should not be applied *in addition to* the agglomeration calculations as they are measuring the same thing in different ways.

References

Ben-Akiva, M. and T. Morikawa, (1990), “Estimation of switching models from revealed preferences and stated intentions”. *Transportation Research Part A*, 24A/6, pp. 485-495.

Bhouri, N. and J. Kauppila (2011), “Managing highways for better reliability: Assessing reliability benefits of ramp metering”. Transportation Research Record: *Journal of the Transportation Research Board*, Issue 2229, pp. 1-7.

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.

CIE (2012), “Costs and benefits of alternative growth scenarios for Sydney, focusing on existing urban areas”, prepared with Arup for the NSW Department of Planning and Infrastructure, Sydney.

de Jong, G.C. and M.C. Bliemer (2015), “On including travel time reliability of road traffic in appraisal”. *Transportation Research Part A*, Vol. 73, pp. 80-95.

de Palma, A. (2014), “The Grand Paris Project: Tools and Challenges”, in *Major Transport Infrastructure Projects and Economic Development*. Paris: OECD Publishing, pp. 71-98. DOI: http://dx.doi.org/10.1787/9789282107720-en

Dobes, L. and J. Bennett (2009), “Multi-Criteria Analysis: "Good Enough" for Government Work?”. *Agenda: A Journal of Policy Analysis and Reform*, Vol. 16(3), pp. 7-29.

Dodgson, J.S. (1973), “External Effects and Secondary Benefits in Road Investment Appraisal”. *Journal of Transport Economics and Policy*, Vol. 7/2, pp. 169-185.

Eddington, R. (2006), *The Eddington Transport Study: Transport’s Role in Sustaining the UK’s Productivity and Competitiveness*. H.M. Treasury, London.

Fosgerau, M. (2015), “The valuation of travel time variability”, International Transport Forum. Discussion Paper, Paris. http://www.itf-oecd.org/sites/default/files/docs/valuation-travel-time-variability.pdf

Fosgerau, M., L. Engelson and J.P. Franklin (2014), “Commuting for meetings”. *Journal of Urban Economics*, Vol. 81, pp. 104-113.

Graham, D. (2014), Causal influence for ex-post evaluation of transport inteventions, International Transport Forum, Paris.

ITF (2010), *Improving Reliability on Surface Transport Networks*, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/9789282102428-en

ITF (2011), *Improving the Practice of Transport Project Appraisal*, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/9789282103081-en

ITF (2012), *Sharing Road Safety: Developing an International Framework for Crash Modification Functions*, OECD Publishing, Paris.
DOI: http://dx.doi.org/10.1787/9789282103760-en

ITF (Forthcoming), *Ex-post assessments of transport investments and policy interventions*. OECD Publishing, Paris.

Jara-Diaz, S. (1986), “On the relation between users’ benefits and the economic effects of transportation activities”. *Journal of Regional Science*, Vol. 26/2, pp. 379-391.

Kouwenhoven, M. and P. Warffemius (2015), “Forecasting Travel Time Reliability in Road Transport: a new Model for The Netherlands”, International Transport Forum, Paris.

Langer, A. and C. Winston (2008), “Toward a comprehensive assessment of road pricing accounting for land use”, in G.a.R.P.J. Burtless, (ed.) *Papers on Urban Affairs*. Brookings Institution Press, Washington, D.C., pp. 127-175.

Mackie, P. and T. Worsley (2013), *International comparisons of transport appraisal practice: overview report*, University of Leeds, Leeds.

McKinnon, A., A. Palmer, J. Edwards and M. Piecyk (2008), Reliability of road transport from the perspective of logistics managers and freight operators, Report prepared for the Joint Transport Research Centre of the OECD and the International Transport Forum, Paris.

Nellthorp, J. and P.J. Mackie (2000), “The UK Roads Review—a hedonic model of decision making”. *Transport Policy*, Vol. 7, pp. 127-138.

Rus, G.d. (2009), “Economic analysis of high-speed rail in Europe”. http://www.fbbva.es/TLFU/dat/inf_web_economic_analysis.pdf (accessed 10 December 2015).

Small, K. (1982), “The scheduling of Consumer Activities: Work Trips”. *American Economic Review*, Vol. 72/3, pp. 467-479.

Venables, A. (2015), *Incorporating wider economic impacts within cost-benefit appraisal*, International Transport Forum. Discussion Paper, Paris. http://www.itf-oecd.org/incorporating-wider-economic-impacts-within-cost-benefit-appraisal

Venables, A., J. Laird and H. Overman (2014), *Transport investment and economic performance: Implications for project appraisal*, Department for Transport, London.

Vickerman, R. (2008), "Recent Evolution of research into the Wider Economic Benefit of Transport Infrastructure Investments", in *The Wider Economic Benefits of Transport: Macro-, Meso- and Micro-Economic Transport Planning and Investment Tools*, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/9789282101834-3-en

Vickery, W.S. (1969), “Congestion theory and transport investment”. *American Economic Review*, Vol. 59/2, pp. 251-260.

Volpe Transportation Systems Center (2012), Trends in statewide long-range transportation plans: core and emerging topics. https://www.planning.dot.gov/documents/State_plans_report_508_A.PDF (accessed 10 December 2015).

Weisbrod, G. (2015), *Estimating Wider Economic Impacts in Transport Project Prioritisation using Ex-Post Analysis*. International Transport Forum, Discussion Paper, http://www.itf-oecd.org/estimating-wider-economic-impacts-transport-project-prioritisation-using-ex-post-analysis-0

Williams, H.C.W.L. (1977), “On the formation of travel demand models and economic evaluation measures of user benefit”. *Environment and Planning A*, Vol. 9/3, pp. 285-344.

Worsley, T. (2015), *Summary and Conclusions of the Roundtable: Ex-post Assessment of Transport Investments and Policy Interventions*. International Transport Forum, Discussion Paper,

http://www.itf-oecd.org/quantifying-socio-economic-benefits-transport-roundtable-summary-and-conclusions

Chapter 2

The valuation of travel-time variability

Mogens Fosgerau[1]

This chapter provides an overview of some alternative conceptual definitions of travel-time variability, discusses their implications about behaviour, and puts them into a broader context, including deviations from the underlying assumptions regarding rational behaviour. The chapter then discusses the empirical basis for assigning a value to travel-time variability. This discussion leads to the conclusion that a fair amount of scepticism is appropriate regarding stated preference data and that attention should turn to the possibilities that are emerging for using large revealed preference datasets. The bottom line is that travel-time variability is quantitatively important and cost-benefit analysis should account for it, using the best values we can get, in order not to imply a bias towards projects that do not reduce travel-time variability. Omitting the cost of travel-time variability is not the neutral option.

1 Technical University of Denmark, Kongens Lyngby, Denmark

Introduction

Many commuters across the world find their travel time between home and work to be rather unpredictable. In addition to systematic variation by time of day and by day of the week, travel time has a considerable random component. Random travel-time variability is a significant issue. The annual total time and money expenditures on transport in the US are worth more than USD 5 trillion (2007), which corresponds to more than 30% of the US GDP (Winston, 2013). For the US in 2011, road congestion for work trips alone caused an estimated 5.5 billion hours of travel delay and 2.9 billion gallons (11 billion litres) of extra fuel consumption with a total cost of USD 121 billion (Schrank et al., 2012). Nonrecurring traffic congestion (due to accidents, bad weather, special events, and other shocks) contributes 52-58% of total delay in US urban areas (Schrank et al., 2011). This indicates that accounting for travel-time variability would add a very significant amount to the accounted cost of traffic congestion.

The variability of travel time is often large as a proportion of travel time on a given trip. Looking for example at a range of routes in central Stockholm, the standard deviation of travel time is on average 25% of mean travel time during commuting peaks with values ranging up to 75%, taking into account the systematic variation of travel time over the peak (Fosgerau et al., 2014). Accounting for travel-time variability would then add very different costs to different routes, which would matter for traveller route choices, and also for project selection.

Cost-benefit analysis of transport projects builds on an assessment of the generalised travel costs to travellers with and without projects under consideration. The generalised travel costs include monetary costs as well as a range of time-related costs accounting for the value of time spent on various parts of trips. Changes in the generalised travel costs induce changes in demand, which in turn are associated with changes in the consumer surplus that enters the cost-benefit analysis. To be able to include travel-time variability into transport project cost-benefit analysis, we then need to include the cost of travel-time variability into the generalised travel costs. This requires three things: defining a unit of measurement for travel-time variability, predicting the quantity of travel-time variability, and determining the cost to travellers per unit of travel-time variability.

We have in mind a setup where a traffic model predicts travel demand and computes travel-time distributions, including means and variability of travel times. The output from the traffic model is subsequently used in a cost-benefit analysis. Values of time and of travel-time variability are applied to the corresponding quantities in the traffic model output to compute changes in the generalised cost of travel. Travel-time variability would also be accounted for in traffic model description of travel demand and thereby impact behaviour.

This chapter first provides an overview and discussion of alternative conceptual definitions of travel-time variability, emphasising models that have a foundation in micro-economic theory. Such a foundation is a strong advantage. It embodies rationality assumptions that constrain the behaviour implied by the models, thereby ensuring that the behaviour implied by the models makes sense. The information that we can obtain from empirical data will probably never be sufficient to pinpoint the best model. The restrictions implied by theory on the range of possible models are therefore useful to complement the empirical evidence that we can find.

Having established some conceptual models for travel-time variability, we next discuss some broader issues. The first is the fact that travel-time variability will often be costly to people other than the traveller, including other meeting participants. However, the standard models used to assign a cost to travel-time variability take the perspective of a single agent and do not represent other people. The second issue is the role of information and expectations. What counts as random travel-time variability depends on the information available to travellers and it is important that we take this into account. Very little is known, however, about how travellers actually form expectations of travel times. Third, the models that we use are based on neoclassical rationality assumptions, but these assumptions can be systematically violated, especially in stated preference experiments. The chapter discusses what to do about this issue.

The chapter continues to review some empirical evidence, indicating the main challenges we need to confront and arguing that we should move towards using revealed preference data to infer the cost of travel-time variability to travellers.

This chapter is not a literature review and does not mention all the relevant literature. Li et al. (2010) review both theoretical and empirical aspects of the value of reliability. Carrion and Levinson (2012) provide a review with much historical detail and a broader focus than presented here. Small (2012) reviews the broader literature on the valuation of travel time. Concerning freight transport, a recent review is provided by Feo-Valero et al. (2011). The reader may refer to these papers for a more comprehensive overview of the literature.

Conceptual models

The basic model

We consider the simplest possible models that meet the standards of classical micro-economics, while allowing us to talk about the cost of travel-time variability in a meaningful way. This section outlines a general model structure with some specific instances presented in sections below. We discuss the assumptions involved in some detail such that we know how to interpret the numbers that arise. Some technical details are omitted in order to focus the exposition on intuition; these details may be found in the original papers cited above. For concreteness, we talk about the case of a commute trip from home to work, but the models are applicable to any trip. The models describe passenger transport but may just as well be used to describe freight transport trips. Empirically, it is very important to account for the diversity of trip purposes as some trips are much more sensitive to delays than others: think for example about freight trips with perishable goods that may lose their value due to delays or about passenger trips accessing a flight connection at an airport or urgent trips to the hospital.

The first step in defining a model that is useful for valuing travel-time variability is to consider what the outcomes of interest are. The minimum we have to consider is travel time itself; we shall consider models below where outcomes are described just in terms of travel time. Here we shall use a little more structure, assuming that what matters about a trip is *when* it takes place, using notation t_{dep} for the time of departure and t_{arr} for the time of arrival. This will allow us to distinguish trips of the same duration at different times of day. The travel time T is the difference between the departure time and the arrival time; from the point of view of the traveller, the travel time is random and we assume that the traveller knows the distribution of travel time. The traveller selects the departure time and then the arrival time is given by $t_{arr} = t_{dep} + T$.

The next step is to think about a utility function $U(\,t_{dep}, t_{arr})$ that ranks all possible outcomes. We call it a scheduling utility, since it concerns the scheduling of the trip. We shall consider several possibilities below with the thing in common that scheduling utility is money-metric, such that the utility difference between two potential outcomes is a monetary value. At this stage, we could equally well have formulated utility in terms of travel time and either departure time or arrival time, since travel time, departure time and arrival time are related by an identity. Formulating utility in terms of departure time and arrival time emphasises that it is these things that matter to travellers, rather than the travel time itself. We assume that travellers always prefer to depart later and to arrive earlier, *ceteris paribus*.

In order to arrive at tractable expressions for the cost of travel-time variability, we will make the simplifying assumption that utility is separable, splitting into a part that depends only on departure time and another part that depends only on arrival time, $U(t_{dep}, t_{arr}) = U_1(t_{dep}) + U_2(t_{arr})$. A natural way to think of the part of utility that depends on departure time is to say that utility is accumulated at home at some potentially time-varying rate $h(t)$ until the time of departure

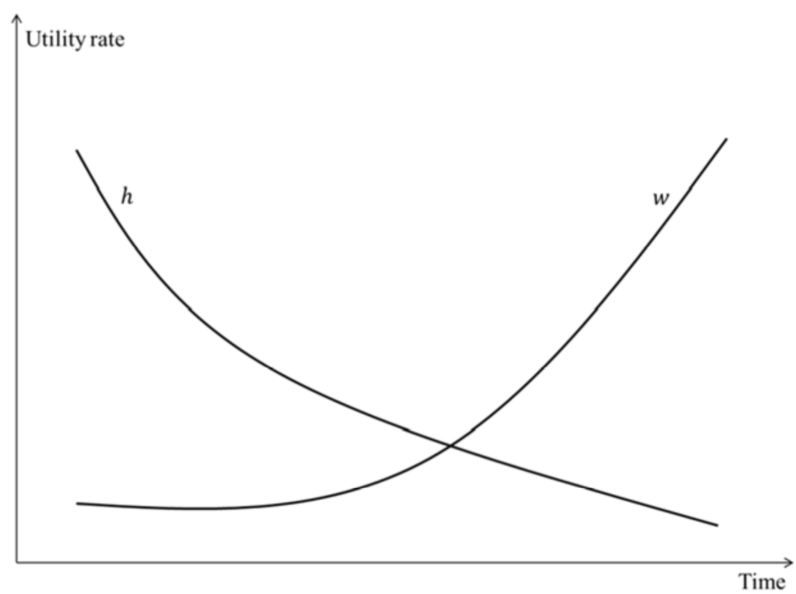
$$U_1(t_{dep}) = \int^{t_{dep}} h(t)dt, \tag{1}$$

and similarly that utility is accumulated at work at another time-varying rate $w(t)$ after the time of arrival

$$U_2(t_{arr}) = \int_{t_{arr}} w(t)dt. \tag{2}$$

The utility rates should be understood as differences from the utility rate that is achieved while travelling (Oort, 1969). This generic form for scheduling utility was first formulated by Vickrey (1973). It is illustrated in Figure 2.1.

Figure 2.1. **Utility rates**



The third step is to say that travellers choose departure time before observing the corresponding travel time. We assume that he or she chooses departure time to maximise the expected utility:

$$E\left(U_1(t_{dep}) + U_2(t_{arr})\right) \tag{3}$$

The final element in the basic model is a technical assumption needed to ensure analytical tractability. We assume that the random travel time T has a distribution that is independent of the departure time. Strictly speaking, this is mostly inaccurate: in the morning peak, the travel time will be longer on average at the height of the peak compared to traveling at the shoulders of the peak. The model is in any case a useful approximation (Fosgerau and Karlstrom, 2010; Fosgerau and Fukuda, 2012).[1]

We can then rewrite travel time as $T = \mu + \sigma X$, where μ is the mean travel time, σ is the standard deviation of travel time and X is the standardised distribution of travel time, having mean zero and standard deviation 1 that is independent of departure time. This is useful to give some structure to the distribution of travel time. We denote by f the density of standardised travel time and by F the corresponding cumulative distribution of standardised travel time.

The expected utility is

$$E\left(U\left(t_{dep}, t_{dep}+T\right)\right) = \int^{t_{dep}} h(t)dt + E\left[\int_{t_{dep}+\mu+\sigma X} w(t)dt\right]. \tag{4}$$

The traveller is assumed to choose departure time to maximise this expected utility. We have in mind trips where the departure time can be freely chosen, in which case the derivative of the expected utility with respect to departure time must be zero at the optimal departure time t^*_{dep}, i.e.

$$h\left(t^*_{dep}\right) = Ew\left(t^*_{dep} + \mu + \sigma X\right). \tag{5}$$

Given specifications of h and w, this equation could be solved to find t^*_{dep}. So in principle, this determines the optimal departure time t^*_{dep} as a function of the mean travel time, its standard deviation as well as the distribution of the standardised travel time, i.e. $t^*_{dep}(\mu, \sigma, f)$. Through this function, the scheduling model explicitly takes into account that travellers will schedule their trips in response to the distribution of travel time, and this is what makes the scheduling model different from models where utility depends only on travel time and not on the timing of trips. Plugging the optimal departure time into the expected utility shows, in principle, how the optimal expected utility depends on the distribution of travel time.

We can define the *value of travel time* as (minus) the derivative of optimal expected utility with respect to mean travel time and the *value of travel-time variability* as (minus) the derivative of optimal expected utility with respect to some measure of dispersion. The term “the value of time” is a tradition in the transportation economics literature; a more precise term would be the “marginal expected disutility of travel time”. Similarly, a traditional term for the value of travel-time variability is the “value of reliability” and it would more precisely be called the “marginal expected disutility of travel-time variability”. We may define the value of travel-time variability with respect to the standard deviation of travel time, in which case we will have the value of travel-time standard deviation. Another possibility is to use the variance of travel time and consider the value of travel-time variance. Which measure of dispersion is more appropriate depends on the model as we shall see below. In order to compare studies from different contexts we may consider the reliability ratio, which is the ratio between the value of travel-time variability and the value of time. Usually, the value of travel-time variability that enters the reliability ratio is expressed as the value of travel-time standard deviation.

We have now arrived at a generic micro-economic model that comprises the minimal number of elements and assumptions, while accounting for the timing of trips. The generic model allows us to go some way in analysing the cost of travel-time variability, but it is necessary to impose additional assumptions to make it operational. There are several attractive possibilities for these additional assumptions and they will lead to different measures of the cost of travel-time variability that we may use in applications. They also imply different behaviour, which enables them to be distinguished empirically.

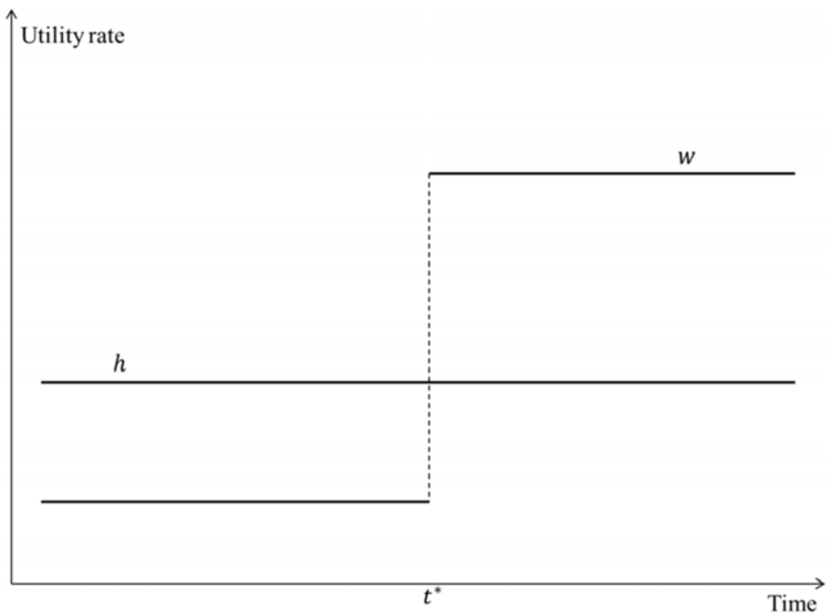
Step model

The most common specification of scheduling utility was established by Vickrey (1969), who included scheduling preferences in his now famous bottleneck model of commuting in a congested demand peak. Later, Small (1982), in a parallel development, formulated the same scheduling utility in an analysis of the timing of commuter work trips. It specifies the utility rate at home to be constant $h(t) = \alpha$, while the utility rate at work is a step function

$$w(t) = \begin{cases} \alpha - \beta, t \leq t^* \\ \alpha + \gamma, t > t^*, \end{cases} \tag{6}$$

where t^* is a preferred arrival time and α, β, γ are positive constants. The model based on this specification of scheduling utility is known as the step model. The utility rates are shown in Figure 2.2.

Figure 2.2. **Utility rates in the step model**



The optimal departure time can then be found to be

$$t^*_{dep} = t^* - \mu - \sigma F^{-1}\left(\frac{\gamma}{\beta+\gamma}\right). \tag{7}$$

This has an intuitively appealing interpretation: The traveller will depart in advance of the preferred arrival time, allowing for mean travel time μ as well as some additional head start or safety margin $\sigma F^{-1}\left(\frac{\gamma}{\beta+\gamma}\right)$ that is proportional to the standard deviation of travel time. The proportionally factor $F^{-1}\left(\frac{\gamma}{\beta+\gamma}\right)$ depends on the shape of the travel-time distribution as well as on preference parameters β, γ that express the cost of arriving early or late. The traveller will arrive later than t^* with probability $\frac{\beta}{\beta+\gamma}$, which depends only on the preference parameters and not on the distribution of travel time.

Inserting the optimal departure time into the expected utility and differentiating, shows (after a bit of work) that the cost of mean travel time μ is a constant α per time unit, while the cost of travel-time variability is $(\beta + \gamma) \int_{\frac{\gamma}{\beta+\gamma}}^{1} F^{-1}(s)ds$ per unit of standard deviation σ. The reliability ratio is

$$\frac{\beta+\gamma}{\alpha} \int_{\frac{\gamma}{\beta+\gamma}}^{1} F^{-1}(s)ds, \tag{8}$$

which depends on preference parameters and on the standardised travel-time distribution but not on the mean μ and the standard deviation σ of travel time.

These convenient results were first established for some special travel-time distributions (Noland and Small, 1995; Bates et al., 2001) and then later for a general travel-time distribution (Fosgerau and Karlstrom, 2010).

To put a ballpark number on the reliability ratio, we may use the stylised values $\alpha = 2, \beta = 1, \gamma = 4$ based on Small (1982) and a value of 0.3 for the integral expression in the reliability ratio (Fosgerau and Karlstrom, 2010). Then the reliability ratio becomes 0.75.

The standard deviation may be replaced as a measure of travel-time variability by any other statistic that is proportional to the standard deviation, provided that the shape of the travel-time distribution may be considered to be constant. To see this, note that for any positive number ρ, we may rewrite the cost of travel-time variability into the following, simply dividing and multiplying by ρ.

$$\left(\frac{\beta+\gamma}{\rho}\right)\left(\rho\sigma\int_{\frac{\gamma}{\beta+\gamma}}^{1} F^{-1}(s)ds\right). \tag{9}$$

Any measure of travel-time variability that is proportional to the standard deviation has the form shown in the second parenthesis here. The first parenthesis is then the corresponding unit value of that measure of travel-time variability.

Given a fixed shape of the distribution of travel time, the standard deviation is proportional to many other measures of the dispersion of travel time that have been used. This includes:

- the difference between two specific quantiles of the travel-time distribution used, notably, by Small, Winston and Yan (2005)
- the difference between a quantile and the mean travel time
- the buffer time index (Texas Transportation Institute and Cambridge Systems, Inc., 2006)
- the mean lateness, $\sigma\int_{\frac{\gamma}{\beta+\gamma}}^{1} F^{-1}(s)ds$.

All such measures are proportional when the shape of the standardised travel-time distribution is constant. Hence, they all share the above micro-economic foundation in the step model.

A drawback of the step model is that the standard deviation is not additive across parts of a trip having independent travel times. If the travel time for the first part of a trip has standard deviation σ_1 and the second part has standard deviation σ_2, then the travel time for the combined trip has standard deviation $\sqrt{\sigma_1^2+\sigma_2^2}$, which is strictly smaller than $\sigma_1+\sigma_2$ with a difference that can be large. If, for example, $\sigma_1=\sigma_2$ then $\sqrt{\sigma_1^2+\sigma_2^2}=\sqrt{2}\cdot\sigma_1$. This is inconvenient when working with network based traffic models, since the expected travel cost for a trip then cannot simply be added up from the level of links.

The step model may be used in several ways. First, it may be used as a structural model, where preference parameters α, β, γ and t^* are combined with an observed travel-time distribution to compute cost measures. This is feasible in particular using stated preference data where the preferred arrival time t^* may be observed or inferred. It may also be used in a reduced form where the expected travel cost for a trip is simply described as being linear in the mean travel time and in the standard deviation (or some other measure proportional to the standard deviation).

Slope model

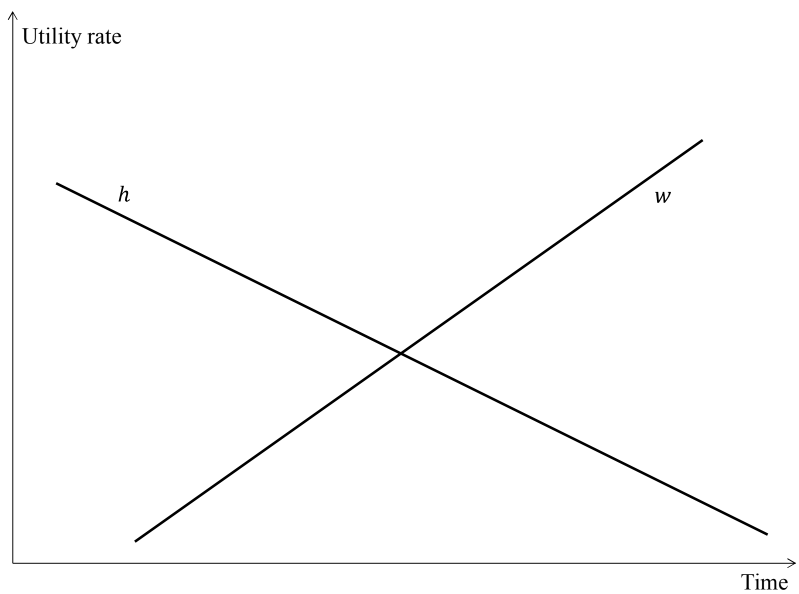
We now elaborate another version of the general model, but this time based on utility rates that are time-varying with constant slopes (Fosgerau and Engelson, 2011). The slope model specifies utility rates as

$$h(t) = \alpha - \beta \cdot (t - t^*) \tag{10}$$

$$w(t) = \alpha + \gamma \cdot (t - t^*), \tag{11}$$

where again α, β, γ are positive parameters while t^* plays the same role as the preferred arrival time in the step model of locating scheduling preferences in time. These utility rates are illustrated in Figure 2.33.

Figure 2.3. **Utility rates in the slope model**



Source: Fosgerau and Engelson, 2011.

Proceeding in the same way as before leads to the optimal departure time

$$t_{dep}^* = t^* - \frac{\gamma}{\beta+\gamma}\mu. \tag{12}$$

The traveller will depart $\frac{\gamma}{\beta+\gamma}\mu$ before time t^*. On average they will arrive $\frac{\beta}{\beta+\gamma}\mu$ after time t^*. The amount of random travel-time variability affects, of course, the distribution of arrival times, but has no effect on the departure time or on the average arrival time. The traveller in the slope model does not allow an additional safety margin when faced with random travel time. This distinguishes the slope model from the step model.

The value of mean travel time is $\alpha + \frac{\beta\gamma}{\beta+\gamma}\mu$ per unit, which depends itself on the mean travel time. Thus the value of time is higher for long trips than for short trips. This is because long trips begin at a time when the utility rate at home $h(t)$ is higher. When $\beta = 0$ such that the utility rate at home is constant, then also the value of mean travel time is constant and equal to α.

The cost of travel-time variability can be expressed as the cost per unit of the variance of travel time and it is then a constant $\gamma/2$. Then the cost of travel-time variability has the advantage, in contrast to the step model, of being additive across links if the random travel times on links are independent.

The value of the standard deviation of travel time is not constant but is proportional to the standard deviation. The reliability ratio in terms of the standard deviation is then not constant and is not immediately comparable to the reliability ratio from the step model.

Engelson and Fosgerau (2011) extend the slope model to the case where the utility rate at work is an exponential function and show that this exhausts all possibilities for slope models that are additive across links.

Non-scheduling models

The step and the slope models describe the timing of individual trips and the preferences concerning the times of departure and arrival, and combine this with expected utility maximisation. This structure has advantages; in particular it ensures that the behaviour predicted by the model makes sense. It leads to predictions regarding how departure times relate to the distribution of travel time, which may be used to evaluate models against empirical evidence. But the structure may also be a constraint if the models do not match actual behaviour. It is alternatively possible to use models that impose less structure.

Such models would ignore the timing of trips and just assume that travellers have preferences regarding travel time where less is better. The equivalent monetary cost of a trip would then be expressed as a convex function of travel time $C(T) = C(\mu + \sigma X)$, and the value of travel time and variability would be derived from the expected cost function, with the expected cost per unit of mean travel time becoming $EC'(T)$ and the expected cost per unit of standard deviation becoming $E[C'(T)X]$.

There are several convenient forms that can be used for the cost function. For example a quadratic cost function $C(T) = \mathrm{bT} + \mathrm{cT}^2$ leads to (Polak, 1987):

$$\frac{\partial EC(T)}{\partial \mu} = b + 2c\mu, \frac{\partial EC(T)}{\partial \sigma^2} = c, \tag{13}$$

which is a different model from the linear slope model presented above.

We are not free to specify any conceivable cost function: it may turn out that some forms are not consistent with underlying scheduling preferences. We would have less faith in such models as they would then not be consistent with any underlying rational scheduling behaviour.

Scheduled services

The scheduling models presented above assume that travellers are able to select their departure time optimally, as is the case for car drivers. Travellers who use scheduled services are constrained in their choice of departure time and this affects their value of reliability. In the case of the slope model it turns out that the value of travel-time variability is unchanged relative to the case of car drivers for most of the models (Fosgerau and Engelson, 2011; Engelson and Fosgerau, 2011).

In the case of the step model, a similar result for the variability of total travel time is not available (Fosgerau and Karlstrom, 2010). Something can be said, though, for the case of a frequent scheduled service. When departures are sufficiently frequent, travellers do not aim for a specific departure, but arrive at the station to catch the next departure, not knowing specifically when that will be. The waiting time until the next departure is then random from the perspective of travellers. In this case, the general result from the step model applies with the arrival time at the station replacing the self-selected departure time (Benezech and Coulombel, 2013). Then travel-time variability may be accounted for, due to random irregularity of departures as well as due to random variability of travel times.

Some broader perspectives

With the above basic theory in place, which allows values to be assigned to travel-time variability, we shall now discuss this theory from a somewhat broader perspective. We will begin by looking at the framing of the scheduling models in terms of a single individual.

External delay costs

The models that we have considered take the perspective of a single individual who travels from home to work or, more generally, between two activities. They consider the cost to the individual related to the time of departure from the first activity and the time of arrival to the second activity. In many cases, however, there will be someone waiting at the second activity and it may have consequences for them if the traveller experiences an unanticipated delay. Such costs are not included in the models we have discussed if the traveller does not include them fully in her scheduling preferences.

We may consider a situation to be unproductive where two persons facing random travel times schedule trips to a joint meeting and where the time one person spends waiting for the other. Using the scheduling models above would lead us to consider each person independently, accounting just for the cost of travel time variability for each person, assuming that his/her scheduling preferences were constant and exogenous. Fosgerau, Engelson and Franklin (2014) analyse such a situation and find that it is not only the variability of travel times that matters but also their correlation: If travel times are positively correlated then random delays are less costly since the participants in the meeting will tend to experience similar delays and then not waste that much time waiting for the other. The setup also has consequences for how departure times depend on the distribution of travel times, since now the travel-time distribution faced by one person, as well as the departure time decision made by that person, affects the outcomes for the other person. This reasoning applies to situations where more participants travel to a joint meeting, but not to situations where some participants do not have to travel to reach the meeting.

We can think of many situations where one person arriving late implies costs for others. Another central feature that seems to be commonly present is that some people are able to adapt to the travel-time variability faced by others. So far, these issues seem not to have been investigated at all, and research to gauge the impact on the costs of travel-time variability would be welcome. A realistic ambition may be to form an opinion about whether using the simple single agent scheduling models discussed above would lead us to overestimate or underestimate the "true" costs of travel-time variability.

Expectations and information

Travel times vary for many different reasons. There is systematic variation over the week and during the day due to systematic variations in demand. On top of that, there is demand variation that is more or less predictable in principle due to holidays, special events, weather and road works. Then there are incidents and accidents, which by nature are highly unpredictable.

To somebody having no information, just observing travel times, all travel-time variation would be random. At the other extreme, a completely informed traveller might be able to predict travel time perfectly and for that traveller, travel time would be not random at all. Actual experienced travellers are somewhere in between. They might have a good feeling for systematic variation by day of the week and by time of day, they might know about special events and road works, and they could be more or less sophisticated in their ability to predict travel times.

In order to predict the cost of travel time-variability, it is necessary to form an opinion about what information travellers use to form their expectations. A straightforward and practical solution is to assume that travellers know the traffic model used in a given application; they are just as well informed as the model. This solution has the attraction of being simple but it is not completely innocuous. Traffic models may predict travel times that depend on the time of day for different days of the week. In that case, travellers may be better informed than the traffic models since they may have information about special events and special days, that is not used in the traffic models. This may imply some overstatement of the cost of travel-time variability.

Non-rational behaviour

The models presented above are based on classical economic rationality assumptions. These assumptions are useful in constraining models to deliver predictions that make basic sense. In short, the assumptions amount to the following. Travellers have preferences over outcomes, meaning that they evaluate a potential trip in terms of the characteristics of that trip and nothing else. These preferences are such that they can be expressed in terms of a utility function. When faced with choices with uncertain outcomes, such as the arrival time when travel time is random, travellers make the choice with the maximum expected utility. It is clear that this is not an exact description of human behaviour, and certainly utility maximisation does not have an exact counterpart in the brain. But that is not really a problem: we are content if the description is about right on average.

A body of evidence is emerging, showing that human behaviour, especially in certain experimental settings, deviates systematically from rational behaviour. Two phenomena of particular relevance for valuing travel-time variability are described by prospect theory (Kahneman and Tversky, 1979). One is loss aversion: that preferences depend on the size and direction of change from a reference point. In the case of travel, the reference point may be a recent trip, or just some kind of "normal" trip. What may be considered as a reference point depends a lot on the context. In any case, the presence of loss aversion contradicts the assumption that preferences should depend on outcomes only. The other phenomenon is probability weighting of utility, which leads to small probability events having a larger influence on behaviour than they would under the plain mathematical expectation.

The stated preference experiments that are often used to measure the value of travel time or the value of travel-time variability may induce non-rational behaviour. It is well documented that we can produce loss aversion in experiments designed to measure the value of travel time (De Borger and Fosgerau, 2008). Furthermore, there is much evidence of probability weighting in experiments involving gambles (e.g. Wu, 1996), and this would presumably extend to stated preference experiments presenting random travel-time variability.[2]

It is clear that the issue of systematic deviations non-rational behaviour must be confronted. A view is emerging that actual behaviour may be seen as approximating rational behaviour, where the approximation relies on heuristics that work well on average but which may lead to clear deviations of behaviour from rationality in certain settings (Steiner and Stewart, 2015). Then the essential question from the point of view of this chapter is whether the presence of travel-time variability induces significant deviations from rational behaviour in the non-experimental settings that travellers encounter in their daily life. We will discuss this issue further below, after presenting some empirical findings specifically related to the value of travel-time variability.

Estimating the parameters

Carrion and Levinson (2012) review a large number of studies of the cost of travel-time variability. They face a number of difficult issues in comparing studies and are not able to provide firm conclusions, but their collection of estimates indicates that a reliability ratio (with respect to the standard deviation of travel time) of around 1 is plausible. The ballpark figure of 0.75 mentioned above is also plausible in the light of this evidence. Most of the empirical evidence is from stated preference studies. Revealed preference studies are emerging, primarily based on data from tolled lanes in the US (Small et al., 2005).

We shall discuss some recent stated preference studies in detail in order to assess the appropriateness of the scheduling model for such data. Hjorth et al. (2015) compare the step and slope models using stated preference data collected from commuters in Stockholm (Börjesson, 2008). Commuters with fixed work times were best described by the step model; this is a plausible result since that model incorporates a preferred arrival time. Commuters with flexible work times were better described by the slope model. The preferred slope model has constant utility rate at home and constant slope on the utility rate at work, which implies a constant value of travel time and a constant value of travel-time variance.[3]

It is possible to estimate the step and slope scheduling models in their structural form, estimating scheduling preference parameters using observations of choices between trips with different departure times and either different travel times or different travel-time distributions. It is alternatively possible to estimate reduced forms of the same models, using observations of choices between trips with different distributions of travel time. This provides an opportunity for testing the scheduling models. Börjesson, Eliasson and Franklin (2012) carry out such a test using stated preference data concerning public transport trips.

Their “structural” stated preference design lets respondents choose between trips that are specified in terms of departure time, a deterministic travel time and a cost. They estimate scheduling preference parameters from these data. Their “reduced form” stated preference design is somewhat different: respondents choose between trips that are specified in terms of an undelayed travel time, a probability of delay, a delayed travel time as well as a cost. The departure time is not specified in this design. Börjesson et al. compute the mean and the variance of travel time for each trip and use this to estimate reduced form models. They estimate both step and slope models in both structural and reduced form. The estimates show that the value of travel time is roughly the same in the structural and the reduced form models. The value of travel-time variability, however, differs a lot: it is five to ten times higher in the reduced form models than in the structural models.

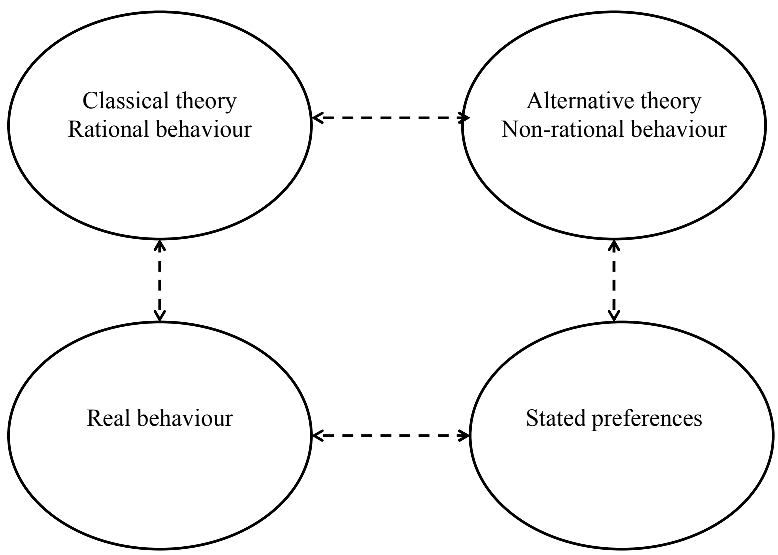
A drawback of the Börjesson et al. study is that their two stated preference designs are somewhat different. The differences between the structural and reduced form estimates could be due to design differences. This is remedied by Abegaz, Hjorth and Rich (2015) who use a stated preference design with two very similar choice experiments. Both experiments comprise two alternative trips described by a cost and two potential travel times with corresponding probabilities. In addition, one of the experiments specifies the departure time for each alternative trip. The two experiments are then as similar as they can be; the only difference is whether the departure time is given to respondents.

Abegaz et al. estimate structural models on data from the experiment including the departure time attribute and reduced form models on data from the experiment not including the departure time attribute.[4] They reproduce the findings of Börjesson et al. that the value of travel time is about the same in the structural and reduced form models, but that the value of travel-time variability is much higher in the reduced form models than in the structural models.

One objection that can be made against these experiments is the following. The models that are estimated assume the scheduling preferences that underlie the stated preferences are stable and do not change during the experiments. It is however quite conceivable that respondents take into account that they are able to reschedule their activity at the end of the trip. When they are asked to choose between alternatives that specify departure times, then they could be conscious of their ability to change their schedule if they were somehow forced to depart later or earlier than their preferred departure time.

In any case, the finding of large discrepancies between the two kinds of experiments poses a serious problem, and some resolution needs to be found. There are essentially two possibilities: either we believe the theory or we believe the stated preference data at face value. It is not possible to do both. We ultimately care about actual behaviour, but we require theory in order to develop models and to carry out cost-benefit analyses. Classical cost-benefit analysis is carried out under classical rationality assumptions as discussed above. For this analysis to be meaningful, rationality assumptions must provide an approximation to reality that is acceptable (Figure 2.4).

Figure 2.4. **Theory and data**



We may eventually be forced to adopt theory that incorporates non-rational behaviour. This will require fundamental changes to the way we think about and carry out cost-benefit analysis. An intermediate possibility is to introduce a distinction between classical (or hedonic) preferences and choice (or decision) preferences (e.g. Köszegi and Rabin, 2006; Steiner and Stewart, 2015). Classical preferences would capture underlying rational preferences and would enter cost-benefit analysis. Choice preferences would describe actual behaviour and would be explicitly linked to classical preferences through some account of systematic deviations from rationality. The link would be such that classical preferences could be backed out from observed choice preferences. Such a theory is conceivable but is not currently available for the case of travel-time variability.

Finally, we have the behaviour that we observe in stated choice experiments. In that context we know that behaviour may deviate significantly from rational behaviour and it is by no means a given that the stated choice experiments provide valid information about real behaviour. Several conclusions are possible.

One possible conclusion is that the behaviour observed in stated preference experiments is simply not valid as a proxy for real behaviour. It is easy to argue in favour of this conclusion, especially in the context of travel-time variability. The choice situations are hypothetical and unavoidably comprise a large amount of information to be digested and processed by respondents. We have little reason to believe they are actually able to do that. In that case, we should simply abandon stated preference experiments as a way of obtaining the value of travel-time variability and look for ways to use revealed preference data. We may still think that people have heuristics that allow them to make reasonable choices in real context involving random outcomes, where they will have repeated experience and lots of time to learn and revise decisions.[5]

A second possible conclusion is that we should find a way to introduce the distinction between classical preferences and choice preferences that govern behaviour in stated preference experiments. Then we would use stated preference data to measure choice preferences, and then back out classical preferences to obtain values that can be used in cost-benefit analysis. The question of the validity of stated preferences regarding actual preferences would remain, however.

The third possible conclusion is to accept the behaviour we observe in stated preference experiments at face value. If evidence of non-rational behaviour is present in the data at hand, then we must account for that; we cannot just ignore this evidence. Then we will have an inconsistency between our estimates and the context in which we will use them. This is not a very attractive option.

Conclusion

Travel-time variability is clearly quantitatively important. Including it in cost-benefit analysis will influence the ranking of projects and will therefore have significant real implications.

We have a simple and firm theoretical foundation in scheduling models for including travel-time variability in traffic models and in cost-benefit analysis. Different models lead to different predictions that can be held against empirical evidence. We have research that provides some broader perspectives that allows us to form opinions regarding how well we might do in capturing the full costs of travel-time variability.

Our ability to put numbers on the value of travel-time variability can certainly be improved. We have evidence of serious problems with our stated preference experiments that are most often used to derive values. This does not mean, however, that zero is the best estimate of the value of travel-time variability: not including travel-time variability in analysis is not the neutral option, but will imply a bias towards projects that do not improve reliability.

Stated preference experiments, at least in the context of travel-time variability, entail some fundamental problems that may be insurmountable. The reason why stated preference experiments are still popular might be that such data are comparatively cheap to collect and easy to analyse. It seems clear that practice should move towards using revealed preference data as far as that is feasible. The additional cost of acquiring data and estimating models should be held against the possibilities for improved selection of transport projects and policies. Given the very high stakes involved, the investment in data and analysis should easily pay off: after all, infrastructure is much more expensive than research and data.

We are entering an era of big data. Research is now increasingly getting access to new large datasets that describe traffic conditions and actual travel behaviour. We have high-frequency measurements of speed and density at dense sets of locations in large road networks. We have datasets that track the location and time of every single train in national rail networks. We have datasets that track large numbers of GPS enabled cars, trucks and mobile devices through transport networks. We have travel card data that track travellers through large public transport networks. When the coverage of data is sufficient, we can infer travel times and travel-time distributions across large networks. We can observe trip origins and destinations as well as route choices.

It is clear that it is not easy to use all this data. A lot of processing is required. But the gain, on the other hand, is clear: it is real data, not hypothetical choices. The revealed preference estimates of the value of travel-time variability of Small, Winston and Yan (2005) relied on data describing a choice between just one tolled and one untolled route. To go to other countries and to other contexts, a number of challenges will have to be overcome. One challenge that will probably be common is that we prefer to have price variation in order to assign monetary values to travel time and travel-time variation. Otherwise we have to infer monetary values in some indirect way, perhaps translating distance into an equivalent monetary cost, which introduces additional uncertainty. Another challenge is that we need to estimate models that describe route choice in large networks. A new generation of models is emerging (Fosgerau et al., 2013) that may be useful in this context. These new models resolve the issue that the number of possible routes in large networks is extremely large. Current practice route choice models use ad hoc devices to circumvent this problem but at the cost of systematically biasing estimates.

Making use of the emerging large datasets has many other benefits. They can be much more comprehensive than surveys and can therefore provide a much more accurate picture of traffic conditions and demand patterns than was feasible in the past. The opportunity for using these data for valuing travel-time variability may come about just as a by-product of the many improvements to traffic modelling that these data make possible.

Notes

1. It is feasible to check empirically whether the standardised travel time distribution can in fact be considered independent of departure time and how the mean and standard deviation evolves over congested peaks. For any given model, it is feasible to evaluate the numerical consequences of assuming that travel time is independent of the departure time in the case that it is not.

2. Hjorth and Ramjerdi (2011) estimate a cumulative prospect theory (Tversky and Kahneman, 1992) model on stated preference data presenting random travel times and find that extreme outcomes are over-weighted.

3. A more technical conclusion from their work is that models with multiplicative error terms were better than the standard models with additive error terms (Fosgerau and Bierlaire, 2009). There were empirical identification problems for the models with exponential utility rates, which points to a need for developing stated preference designs to strengthen identification of the slope model.

4. They use a multiplicative model (Fosgerau and Bierlaire, 2009) as that provides the best fit.

5. Stated preference experiments are routinely used to measure the value of travel time in cases where travel time is not random (Small, 2012). These choice settings do not involve random travel time outcomes and hence involve much less information, they do not require respondents to digest and process probabilities. Respondents should then be better able to make choices in these experiments than in experiments involving random travel times.

References

Abegaz, D., K. Hjorth and J. Rich (2015), “Testing the slope model of scheduling preferences on stated preference data”. Working Paper.

Bates, J., J. Polak, P. Jones and A. Cook (2001), "The valuation of reliability for personal travel". *Transportation Research Part E*, Vol. 37, pp. 191–229.

Benezech, V. and N. Coulombel (2013), "The value of service reliability". *Transportation Research Part B*, Vol. 58, pp. 1–15.

Börjesson, M. (2008), "Joint RP–SP data in a mixed logit analysis of trip timing decisions". *Transportation Research Part E*, Vol. 44, pp. 1025–1038.

Börjesson, M., J. Eliasson and J.P. Franklin (2012), "Valuations of travel time variability in scheduling versus mean–-variance models". *Transportation Research Part B*, Vol. 46, pp. 855–873.

Carrion, C. and D. Levinson (2012), "Value of travel time reliability: A review of current evidence". *Transportation Research Part A*, Vol. 46, pp. 720–741.

De Borger, B. and M. Fosgerau (2008), "The trade-off between money and time: a test of the theory of reference-dependent preferences". *Journal of Urban Economics*, Vol. 64, pp. 101–115.

Engelson, L. and M. Fosgerau (2011), "Additive measures of travel time variability". *Transportation Research Part B,* Vol. 45, pp. 1560–1571.

Feo-Valero, M., L. Garcia-Menendez and R. Garrido-Hidalgo (2011), "Valuing freight transport time using transport demand modelling: a bibliographical review". *Transp. Rev*, Vol. 201, pp. 1–27.

Fosgerau, M. and M. Bierlaire (2009), "Discrete choice models with multiplicative error terms". *Transportation Research Part B*, Vol. 43, pp. 494–505.

Fosgerau, M. and L. Engelson (2011), "The value of travel time variance". *Transportation Research Part*, Vol. 45, pp. 1–8.

Fosgerau, M., L. Engelson and J.P. Franklin (2014), "Commuting for meetings". *Journal of Urban Economics*, Vol. 81, pp. 104–113.

Fosgerau, M., E. Frejinger and A. Karlstrom (2013), "A link based network route choice model with unrestricted choice set". *Transportation Research Part B*, Vol. 56, pp. 70–80.

Fosgerau, M. and D. Fukuda (2012), "Valuing travel time variability: Characteristics of the travel time distribution on an urban road". *Transportation Research Part C*, Vol. 24, pp. 83–101.

Fosgerau, M. and A. Karlstrom (2010), "The value of reliability". *Transportation Research Part B*, Vol. 44, pp. 38–49.

Hjorth, K., M. Börjesson, L. Engelson and M. Fosgerau (2015) "Estimating exponential scheduling preferences". *Transportation Research Part B*. Vol. 81.

Hjorth, K. and F. Ramjerdi (2011), A prospect theory approach to travel time variability. Presented at the International Choice Modelling Conference.

Kahneman, D. and A. Tversky (1979), "Prospect Theory: An Analysis of Decision Under Risk". *Econometrica*, Vol. 47, pp. 263–292.

Köszegi, B. and M. Rabin (2006), "A model of reference-dependent preferences". *The Quarterly Journal of Economics*, Vol. 121, pp. 1133–1166.

Li, Z., D.A. Hensher and J.M. Rose (2010), "Willingness to pay for travel time reliability in passenger transport: A review and some new empirical evidence". *Transportation Research Part E*, Vol. 46, pp. 384–403.

Noland, R.B. and K.A. Small (1995), "Travel-Time Uncertainty, Departure Time Choice, and the Cost of Morning Commutes". *Transportation Research Record*, Vol. 1493, pp. 150–158.

Oort, O. (1969), "The evaluation of travelling time". *Journal of Transport Economics and Policy*, Vol. 3, pp. 279–286.

Polak, J. (1987), "A more general model of individual departure time choice", in Proceedings of Seminar C. Presented at the PTRC Summer Annual Meeting.

Schrank, D., B. Eisele and T. Lomax (2012), TTI's 2012 Urban Mobility Report. Texas Transportation Institute, Texas A&M University, College Station.

Schrank, D., T. Lomax and B. Eisele (2011), TTI's 2011 Urban Mobility Report. Texas Transportation Institute, Texas A&M University, College Station.

Small, K. (1982), "The scheduling of Consumer Activities: Work Trips". *American Economic Review*, Vol. 72, pp. 467–479.

Small, K.A. (2012), "Valuation of travel time". Economics of Transportation, Vol. 1, pp. 2–14.

Small, K.A., C. Winston and J. Yan (2005), "Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability". *Econometrica*, Vol. 73, pp. 1367–1382.

Steiner, J. and C. Stewart (2015), "Perceiving Prospects Properly". Working Paper.

Texas Transportation Institute, Cambridge Systems, Inc. (2006), "Travel time reliability: Making It There On Time, All The Time" (No. FHWA-HOP-06-070). Federal Highway Administration.

Tversky, A. and D. Kahneman (1992), "Advances in prospect theory: Cumulative representation of uncertainty". *J Risk Uncertainty*, Vol. 5, pp. 297–323.

Vickrey, W. (1973), "Pricing, metering, and efficiently using urban transportation facilities". *Highway Research Record*, Vol. 476, pp. 36–48.

Vickrey, W.S. (1969), "Congestion theory and transport investment". *American Economic Review*, Vol. 59, pp. 251–260.

Winston, C. (2013), "On the Performance of the U.S. Transportation System: Caution Ahead". *Journal of Economic Literature*, Vol. 51, pp. 773–824.

Wu, G. (1996), "Curvature of the probability weighting function". *Management Science*, Vol. 42, pp. 1676–1690.

Chapter 3

Forecasting travel-time reliability in road transport: A new model for the Netherlands

Marco Kouwenhoven[1]

Pim Warffemius[2]

In this chapter we describe how we included travel-time variability in the national Dutch transport forecasting model and the policy impacts of this new forecasting tool. Until now, travel-time reliability improvements for road projects were included in Dutch cost-benefit analysis (CBA) by multiplying the travel-time benefits from reduced congestion by a factor 1.25. This proportionality is based on the linkage between congestion reduction and reliability improvements. However, this treatment of reliability is not useful to evaluate policies that especially affect travel-time variability. From the start, this method was provisional and meant to be replaced by a better method capturing travel-time variability. For this, we derived an empirical relation between the standard deviation of travel time, mean delay of travel time and length of route. This has been implemented in the national Dutch model as a post processing module. The new travel-time reliability forecasting model will be incorporated in the Dutch guidelines for CBA.

1 Significance, The Hague, Netherlands

2 KiM Netherlands Institute for Transport Policy Analysis, Ministry of Infrastructure and the Environment, The Netherlands

Introduction

Absence of travel-time reliability is what the travellers notice: "that frustrating characteristic of the transportation system that prompts motorists to allow an hour to make a trip that normally takes 30 minutes because the actual trip time is so unpredictable" (TRB, 2000: p. 4-1). The ITF (2010: p. 31) defines reliability as "the ability of the transport system to provide the expected level of service quality upon which users have organised their activity". The key aspect of this definition is the assumption that network users have an expectation of a particular level of service and that reliability is a measure of the extent to which the traveller's experience matches their expectation (Hellinga, 2011). In other words, reliability is equivalent to the predictability of travel times and, from the perspective of a traveller, associated with the statistical concept of variability. If one takes the perspective of a system or infrastructure network manager, then reliability measures are focused on the network and its performance, i.e. the fraction of time during which the system performs below a certain quality standard. In this chapter we take the perspective of the road user (passenger as well as freight transport) whereby reliability is focused on trip characteristics.

Travellers and firms may account for the variability in their trips and transport of goods by building in time-buffers as insurance against late arrival. This implies that the consequences of late arrivals can be costly. Not only does the efficiency and productivity lost in these buffers represent a cost that travellers and firms absorb due to unreliability, but also stress, late arrivals, missed connections, missed appointments and early arrivals can be costly. Reliable travel times are intrinsically valuable and network users place a significant value on reliability. Therefore, reliability can be formulated in terms of societal costs. This has the advantage that the investment costs of an infrastructure project to improve reliability can be traded-off against its benefits for society.

In the Netherlands, transport infrastructure projects and other transport policies are *ex-ante* evaluated using cost-benefit analysis (CBA). To incorporate reliability improvements in project and policy evaluation, the Dutch Ministry of Infrastructure and the Environment included the societal benefits of reliable and predictable travel times into CBA. Since 2004, in Dutch CBA practice, an extra benefit of 25% of the travel-time benefits due to reduced road congestion is added to account for reliability benefits (Besseling et al., 2004). This approach is only used for road projects and is based on the linkage between congestion reduction and improved reliability. However, it does not evaluate consequences of policies that especially affect travel-time variability.

From the start, this method was provisional and meant to be replaced by a better method capturing travel-time variability. To include travel-time reliability in CBA, three types of information are needed, namely (De Jong and Bliemer, 2015):

- monetary values to convert reliability benefits into money units
- a model to predict how much an infrastructure improvement project will change travel-time variability
- a model to predict whether network users will change their route choice, mode choice or departure time choice due to changes in travel-time variability.

The monetary value of changes in average travel time has been long included in CBA by making use of the so called value of travel-time savings (VTTS). The VTTS refers to the monetary value travellers place on reducing their average travel time by one hour. In contrast, the value of travel-time reliability savings (VTTRS) to convert changes in travel-time variability in monetary units is relatively new. A recent study for the Dutch Ministry of Infrastructure and the Environment, delivered updated VTTS and VTTRS based on primary data (KiM, 2013; De Jong et. al, 2014; Kouwenhoven et. al, 2014). Based on earlier work (Hamer et. al, 2005; HEATCO, 2006) it was decided that the variability of travel time should be measured by standard deviation of the travel-time distribution. The main reason behind this choice was the assessment that including travel-time variability in transport forecasting models would be quite difficult and that using the standard deviation would be the easiest option. Any formulation that would go beyond the standard deviation or the variance of travel time[1] would be asking too much of the Dutch national and regional transport models (LMS and NRM) that are regularly used in CBA in the Netherlands. The standard deviation is also used as a reliability indicator in the US, UK and Scandinavian countries. The disadvantage of this definition is that it does not capture skewness of the travel-time distribution. It is well known that travel-time distributions are skewed and their long tails towards extreme travel times are an important aspect of travel-time reliability (see also Hellinga, 2011). Studies on network robustness and vulnerability focus on these extreme travel times and their causes.

The inclusion of travel-time variability in transport forecasting models is challenging because transportation planning models that are used to evaluate and prioritise transport policies have been developed to capture average travel time and not travel-time variability. To adapt the Dutch national and regional models to capture reliability in terms of standard deviation, a project was started in 2013. The objective was to find a (new) empirical relation between the standard deviation of car travel time and other variables available in LMS and NRM. This chapter reports the main results of this study. The improved modelling to forecast travel-time variability will be implemented in Dutch policy making. Incorporating the consequences of policies affecting travel-time variability into infrastructure CBA encourages proper consideration of options.

The next section of this chapter discusses the methodology. The database which was used to derive the empirical relation is then described in the following section. Subsequent sections show how the best empirical relation was fitted. The policy impacts of the new reliability forecasting tool are then discussed before the chapter ends with our conclusions and future steps.

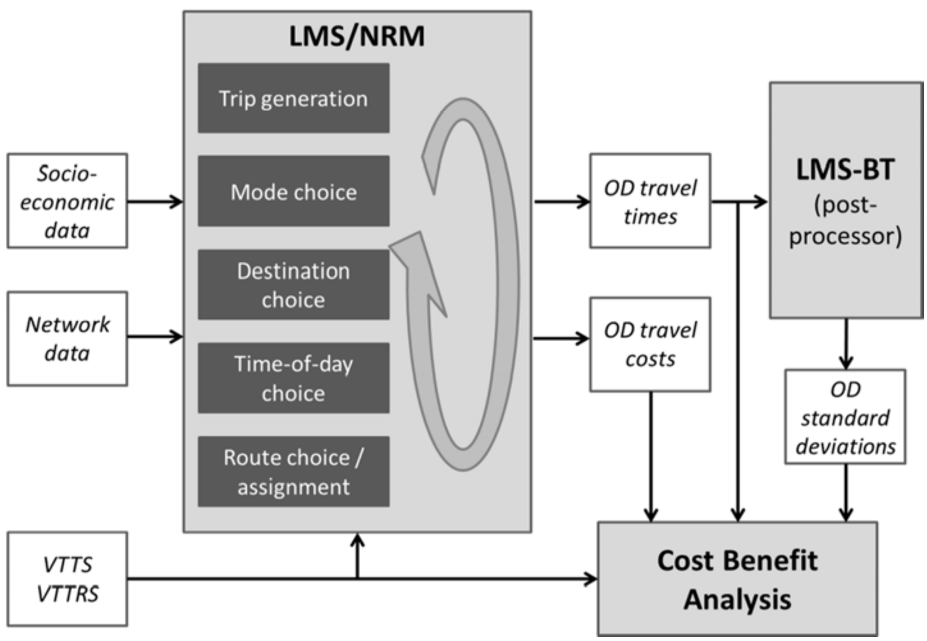
Methodology

In the Netherlands, traffic forecasts for CBAs are usually made using the LMS (the national transport model) or one of the NRMs (the regional transport models). These are similar tour-based models containing the four steps of a classical transport model (tour generation, destination choice, mode choice and route choice) plus a departure time choice module (Willigers and de Bok 2009; Significance, 2011). The tour generation, destination, mode and departure time choices in the LMS and NRM are based on disaggregate models, i.e. these choice models are estimated at the level of an individual traveller. Route choice (i.e. the assignment) for road transport is modelled at the level of origin-destination flows. In an iterative process, the resulting travel times are fed back into the earlier modules so that the travellers can adapt their choices as a result of possible congestion.

The assignment module generates the mean travel time on an average working day for each origin-destination pair and for each time-of-day period. In this process, flows are assigned to routes through the network and for each link in the route a travel time is calculated based on speed-flow curves and on a queuing model. The speed-flow curves are determined using empirical data for each road type. Note that the calculated mean travel time includes delays due to congestion.

Next to travel time and cost, reliability can be an important driver for mode choice, route choice and time-of-day choice (de Jong and Bliemer, 2015). Ideally, this variable would be included in these choice modules as an explanatory variable. However, this would require extensive data collection, modelling and adaptation of the current models. As a second best and much quicker solution, we have developed a post-processing module that calculates the reliability of road travel times for each origin-destination pair and for each time-of-day period. This allows us to calculate travel-time reliability levels for any future scenario and to include the costs or benefits of possible changes in reliability in the CBA (Figure 3.). Each policy measure that can be simulated with the LMS/NRM (adding road capacity, road pricing, etc.) can also be studied for its effects on travel-time reliability.

Figure 3.1. **The role of the transport models LMS/NRM and the post-processor LMS-BT in CBA**



This post-processing module requires an empirical relation between travel-time reliability and any of the output variables that are available in the LMS and NRM such as travel time, congestion, flow, etc. but also the road characteristics that are available in the model (maximum allowed speed, number of lanes, etc.). To derive such a relation, we need to have a database with observed levels of reliability and of all the other variables. This database is described in the next section.

When compiling such a database a number of decisions need to be made. These decisions can have a profound impact on the results. To prevent any inconsistencies, it is paramount that consistency with both the LMS/NRM and CBA procedure drives these decisions:

- As explained in the introduction, we use "the standard deviation of the travel-time distribution" as our reliability indicator. However, we still need to define *which* travel-time distribution. It is common to compile a travel-time distribution by measuring the mean travel times on a number of days within a certain period (e.g. one month or one year) when departing at the same time (e.g. between 08:00 and 08:15). In this way, reliability is interpreted as day-to-day travel-time variability. In this chapter we follow this approach. We have measured mean travel times over a period of one year (2012) for vehicles departing in the same 15-minute interval. This means that for each 15-minute period, we have a different travel-time distribution and hence a different value of the reliability indicator.

- Note that by taking the mean travel time when departing at the same time, we already exclude the vehicle-to-vehicle variation from the reliability indicator. To a certain extent, this makes sense. Some vehicle-to-vehicle variation is caused by driver characteristics; some prefer to drive more slowly than others. These drivers may have other mean travel times, but still have the same travel-time reliability. However, some of the vehicle-to-vehicle variation is caused by infrastructure: departing a fraction of a second later may result in just having to stop for a red traffic light and having a delay of one minute. Ideally, this is the type of variation that is included in the reliability indicator. However, in our project it is excluded because of the method we use to measure reliability.
- The LMS/NRM forecasts mean travel times and vehicle volumes for an average hour during the morning peak (lasting from 07:00 until 09:00), during the evening peak (from 16:00 to 18:00) and for a typical hour during the rest of the day (defined as an average hour between the mid-day period, i.e. between 10:00 and 15:00). In order to find the average reliability indicator for these three periods, we average the reliability indicators over all the 15-minute periods during these periods (eight 15-minute periods for each peak and twenty-eight 15-minute periods during the mid-day period). The reliability indicator for each 15-minute period is weighted with the average volume during that period.
- Note that this is different from first averaging the travel time over the full morning peak (for example) for every day of the year and then producing the travel-time distribution and determining the standard deviation. Travel times will not be constant over a two-hour period and this variability must be included. In order to calculate the pure reliability, the day-to-day variability must be determined with a small departure time resolution (15 minutes or less, depending on the time period over which the travel time on a day can be considered more or less constant), then the reliability indicator should be determined for each (small) time interval. Only in the final step, is the reliability indicator averaged over a longer period (e.g. over the whole peak).
- In Dutch policy, reliability and predictability are considered from the viewpoint of the traveller, (Ministry of Infrastructure and the Environment, 2012). Therefore, reliability should be defined in terms of the deviation of the real travel time from the predicted travel time. As a consequence, we should not look at the mere day-to-day variability of travel times, but correct for the variation in *expected* travel time. As an example: suppose that the travel time on a certain route is always 70 minutes on Mondays, and always 65 minutes on other days, then the travel time is perfectly predictable, and hence, perfectly reliable – given that the travellers know this.
- Extremely long travel times have a severe impact on the computed standard deviation. The long travel times can be caused by malfunctions in the detectors, but they can also be real events (e.g. a breakdown of the traffic system because of a severe accident or due to extreme weather conditions. Should these be included in the travel-time distribution? To decide on this, it is crucial to understand how results will be used. It is important to be consistent with the method

used to determine the mean travel times in the demand model: did they include these long travel times? Also, consistency is needed with the method used to determine the VTTRS.

- In the Dutch situation, the speed-flow curves that are the base for the travel time calculations in the national model do not include extreme events. Additionally, the value of travel-time reliability savings was determined using a stated preference experiment in which a travel-time distribution was shown without any extreme events (Significance et al., 2007). Finally, we believe that these extreme travel times have different causes than normal day-to-day variation. From a policy point-of-view it is better to treat them separately.[2] Hence, in this project we excluded these data points from the analysis. In the next two sections of this chapter, we analyse the impact of this on the outcomes.

Data

We compiled two databases with travel times: one for highway trips and one for trips on other roads. Since the Dutch national model produces forecasts for an average working day, we selected data from all 251 days in 2012 that fell within this definition.

Most Dutch highways are equipped with detector loops that measure average vehicle speed and volume at one-minute intervals. Fifteen-minute averages of these variables were available for this project. We defined 250 routes on the Dutch highway network. Each realistic and logical route started at (or near) a highway entrance and ended at a highway exit, meaning that each route can comprise multiple links. The routes covered the network as completely as possible and overlapped each other as little as possible. Characteristics can be found in Table 3.1.

Table 3.1. **Characteristics of selected routes**

	Highways (250 routes)			**Other roads (40 routes)**		
	Average	Minimum -	Maximum	Average	Minimum -	Maximum
Length (km)	41.5	1.9 -	224.8	5.3	1.7 -	13.8
Average speed (km/h)	93.6	28.3 -	116.7	48.6	13.8 -	97.8
Av. max. speed (km/h)	112.7	92.4 -	123.6	66.0	23.5 -	109.0
Number of links	34.1	1 -	200	*n.a*	*n.a.*	

Major urban and regional roads are equipped with video cameras. With the use of Automatic Number Plate Recognition (ANPR) techniques, passage times of individual vehicles are recorded. Combining information from multiple cameras average travel times over fifteen minutes can be obtained. Since this project focussed on highway travel, only 40 urban and regional routes on non-highways roads were defined to be used in this project. The full data set was compiled in four steps, as described below.

Determining raw travel times

The travel-time data from video cameras for other roads are already for the full route. However, the highway detectors only provided local speeds and these needed to be converted to mean travel times between two detector loops.

Vehicle speeds at each detector were estimated by adding these mean travel times, the total travel time for each route when departing within a certain 15-minute interval was determined. We took into account the fact that a vehicle on a long route does not pass all detectors in the same 15-minute interval. Since it turned out to be technically complex to combine data from consecutive days, we only looked at the ninety-two 15-minute intervals between 0:00 and 23:00. The average volume over a route was calculated by averaging the volumes at the detector loops (weighted with the length between the loops).

In this way, a database with mean travel times and volumes for 250 highway routes, on each of the 251 working days, for a departure time in each of the ninety-two 15-minute intervals was compiled. These data were enriched with the free-flow travel times and with route characteristics such as the length and the capacity (i.e. for each route, the maximum volume over 251 working days and ninety-two 15-minute periods).

Travel-time delays will be correlated between adjacent links, so the standard deviation of the total route travel time cannot be derived directly from the standard deviations of link travel times. Therefore, only the total route travel time is stored and the standard deviation is calculated only in the final step.

Excluding extreme events

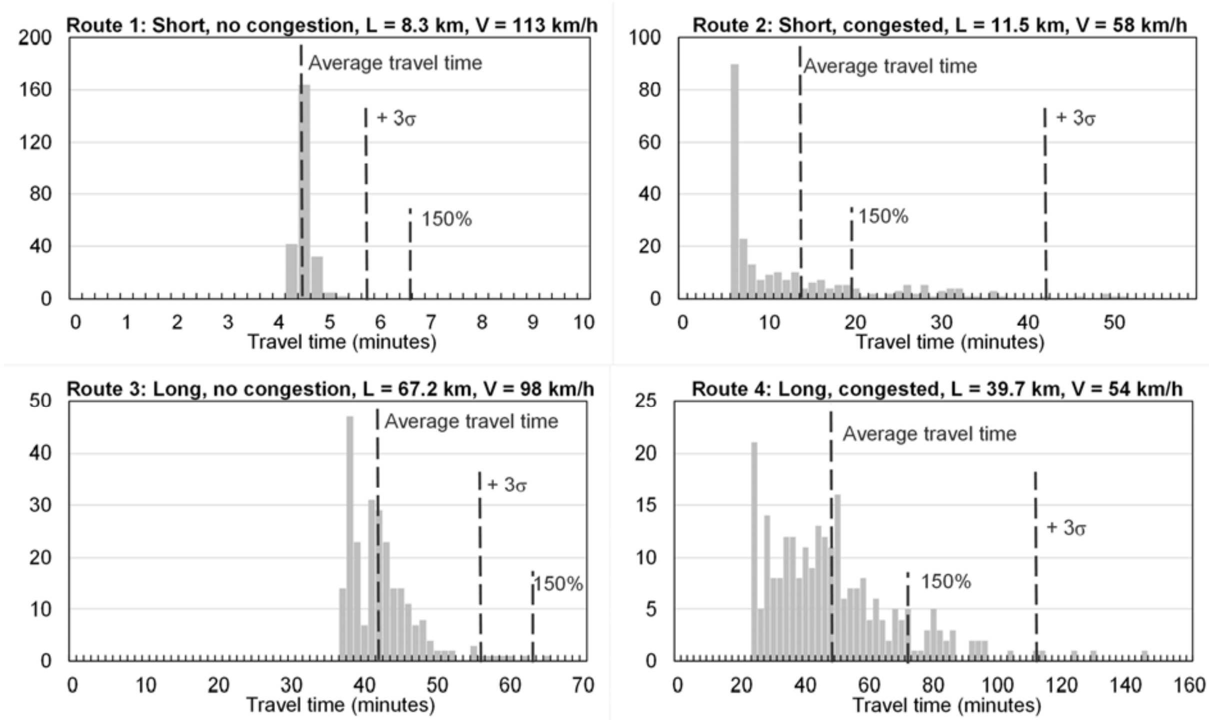
As explained in the previous section, we exclude extreme events from the travel-time distributions, in part because they have a strong impact on the standard deviation, but also for consistency in their application. However, it is not clear where to put the boundary for exclusion of an extreme travel time.

To determine this boundary, we visually inspected the raw travel-time distributions of all 250 highway routes (see Figure 3.2 for four examples). A boundary of three times the (raw) standard deviation above the mean travel time produced a good match with a visual classification of outliers for most routes. However, especially for routes with low congestion an additional criterion turned out to be necessary: the travel time of an outlier should be at least above 150% of the mean travel time. So, travel times are excluded with:

$$TT_{i,j,k} > \max\left(\overline{TT_{j,k}} + 3\sigma_{j,k},\ 150\% \cdot \overline{TT_{j,k}}\right) \tag{1}$$

in which $TT_{i,k}$ is the travel time for day i, route j and 15-minute departure time period k, $\overline{TT_{j,k}}$ is the mean travel time for departure time period k, and $\sigma_{j,k}$ is the standard deviation of the travel-time distribution for time period k (before exclusion of extreme events). For each 15-minute period, equation (1) results in the exclusion of on average 4 out of the 251 working days. As a result, the average standard deviation over all 250 routes is reduced by 29 %. In other words, 1.6% of the days contribute to almost one-third of the standard deviation.

Figure 3.2. **Travel-time distributions for four routes with different characteristics**



Correcting for variations in travel-time expectation

As explained in in the methodology section, we want to determine the deviation of the real travel time from the travel time as expected by the traveller. Unfortunately, no information is available about these expectations. Therefore, for each day, we approximate the expected travel time by taking the mean of the travel times on the same day-of-the-week in the four weeks before and after this day. For instance, our approximation of the expected travel time for Wednesday, 20 June, is the mean of the travel times on the Wednesdays 23 and 30 May; 6, 13 and 27 June; 4, 11 and 18 July. This running average reflects day-of-the-week and seasonal fluctuations in the mean travel time, but no incidental variations. No travel times that are marked as an outlier, are included in the calculation of the expected travel time. For holiday periods, when almost no congestion exists, the mean travel time of the four days directly before and after each day are taken.

As a result of this correction, the average standard deviation over all 250 routes is reduced by another 12% on top of 29% reduction in the previous step.

Calculation of the reliability indicator

For each route, for each day of the year, for each 15-minute period, we calculate the deviation of the real travel time from the expected travel time. Next, for each route and each 15-minute period, we determine the standard deviation of the deviations over all days (excluding the extreme events). For each period of the day (morning peak, evening peak, mid-day period), the mean travel time and the mean standard deviation is calculated by taking the average over the relevant 15-minute periods weighted by their mean flows.

Testing alternative empirical relations for travel-time reliability

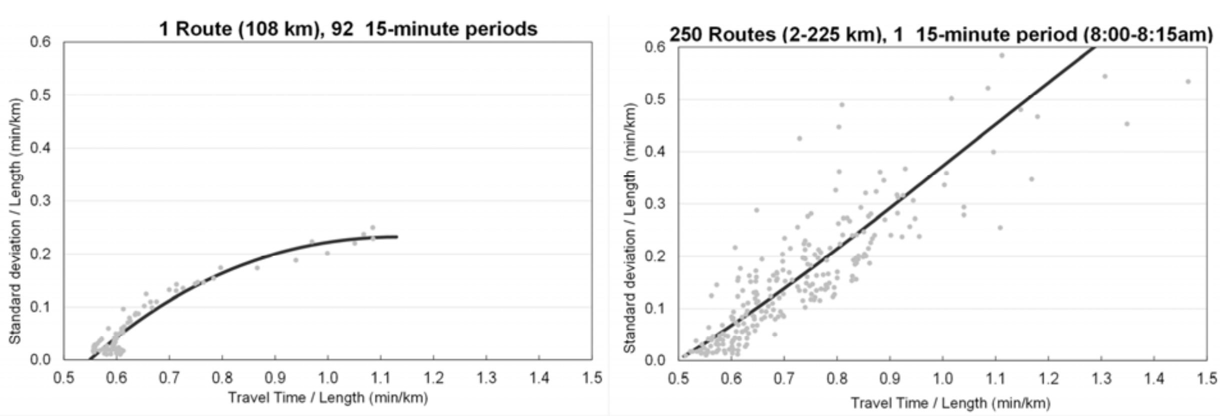
Our final databases contain 750 highway observations (250 routes times, 3 time-of-day periods) and 120 observations non-highway observations (40 routes times, 3 time-of-day periods). Each observation consists of an average travel time (averaged over all workdays in 2012), a standard deviation (of the distribution of deviations from the expected travel time), and other characteristics of the route (length, number of lanes, etc.). The non-highway routes are very diverse in nature (highway feeder routes in urban areas versus regional routes connecting two highways etc.) and are relatively short (see Table 3.1). This limited number of observations did not allow for extensive modelling efforts. However, the highway database was extensive and we were able to fit several functional forms described in the literature such that the results could be compared.

This analysis aims at finding the best functional forms of the empirical relations for our highway database. We do not try to compare the estimated coefficients with other studies as these values strongly depend on the outlier criterion used, on the way the variation of expected travel time is taken into account and on the period over which the data is averaged (also discussed in the next section). The conclusions on the best functional form are only valid for our own data set. It is well conceivable that other functional forms describe datasets in other countries better.

Many other researchers have used data sets that combined several routes and a range of 15-minute periods. The variation of standard deviation by routes may follow a different relation than the variation by 15-minute periods as can be seen from Figure 3.3. For the purpose of this project, we are only interested in the variation between routes for the morning peak, evening peak and mid-day period. In the evaluation of the results, we concentrate on the best functional form for the variation of the standard deviation by routes for the morning peak. The best function is one that describes the data best, and does not show any curvature that is not supported by the data (neither inside nor outside the data range).

In this analysis we tested functional forms that were suggested in earlier studies covering a wide range of possible functional forms that use different independent variables (travel time, travel time per kilometre, congestion index, mean delay). However, this list is not comprehensive. For a more complete overview of functional forms, see De Jong and Bliemer (2015).

Figure 3.3. **Travel time per km versus standard deviation per km**



Note: Values for a single route (left) and for 250 routes for the morning peak (right). Both datasets are fitted with cubic polynomials (dark lines) which are clearly different.

Linear model (the Netherlands – Hellinga)

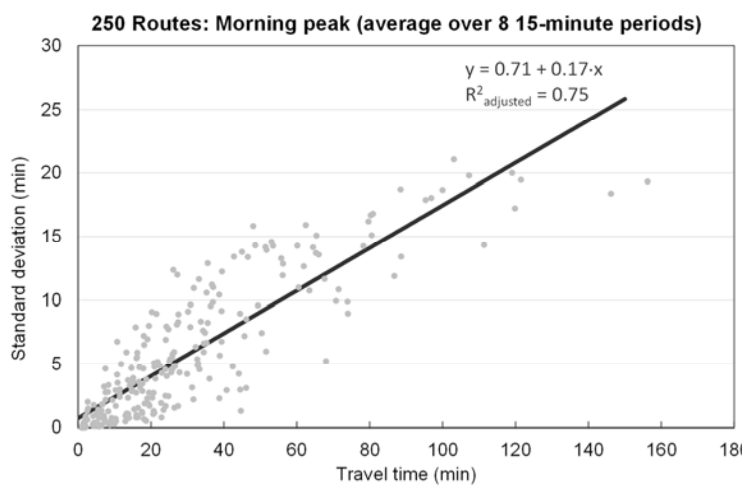
The first functional form under consideration is a simple linear relation between the standard deviation σ and the mean travel time *TT*:

$$\sigma = a[0] + a[1] \cdot TT \tag{2}$$

in which $a[0]$ and $a[1]$ are coefficients to be estimated. This functional form was used by Hellinga (2011) in his study of the variation of travel time on a single route of about 25 kilometres on the A12 highway in the Netherlands. Travel times were derived from detection loop data averaged over 15-minute periods. Each 15-minute period provided one data point, so his final data set consisted of 92 points, since he excluded trips after 23:00.

Hellinga only considered the variation between 15-minute periods and concluded that a linear relation was sufficient. If we analyse the variation between 15-minute periods for each of our 250 routes, we see that for short routes a linear relation is sufficient, though for longer routes with congestion, a decreasing slope can be observed (as is illustrated by the single route data in Figure 3.3-left). If we fit the linear relation (2) to our morning peak data for 250 routes (Figure 3.4) we see that our data can be nicely described by this relation, though a reasonable amount of spread around this relation remains. The adjusted R^2 is 0.75.[3]

Figure 3.4. **Travel time versus standard deviation fitted with a linear function**



Length-standardised linear model (US – SHRP2)

Several US researchers within the second Strategic Highway Research Program (SHRP2, see Mahmassani et al., 2014) use the standard deviation of standardised travel time. This network approach has the advantage that it can also be applied if multiple routes are used between A and B with different lengths. It is especially suitable for dense urban networks. The SHRP2 researchers have shown that there is an almost linear relation between the travel time per unit length and the standard deviation per unit length

$$\frac{\sigma}{L} = a[0] + a[1] \cdot \frac{TT}{L} \tag{3}$$

in which *L* is the length of the route and $a[0]$ and $a[1]$ are coefficients to be estimated.

If the observations in our database are converted to this metric, a linear function fits very well (Figure 3.5, dark line). However, some variation remains unexplained: the adjusted R^2 is 0.78, which is slightly better than for the linear relation above.

Length-standardised cubic model (United Kingdom – Mott MacDonald)

In the UK, Mott MacDonald estimated relations for the day-to-day variability which they describe as "what remains after accounting for all predictable variations (time of day effects, day type effects and seasonal effects) and variability due to incidents" (Sirivadidurage et al., 2009). They used data from inductive loop sensors, automatic number plate recognition and matching and GPS tracking, averaged over 15-minute periods on several highway routes. Journey times that are more than 2 standard deviations above the mean are flagged as incidents and were excluded. Predictable variations were accounted for by allocating each day of the year to one of 21 day types and by determining average journey times and standard deviations for each of these day types.

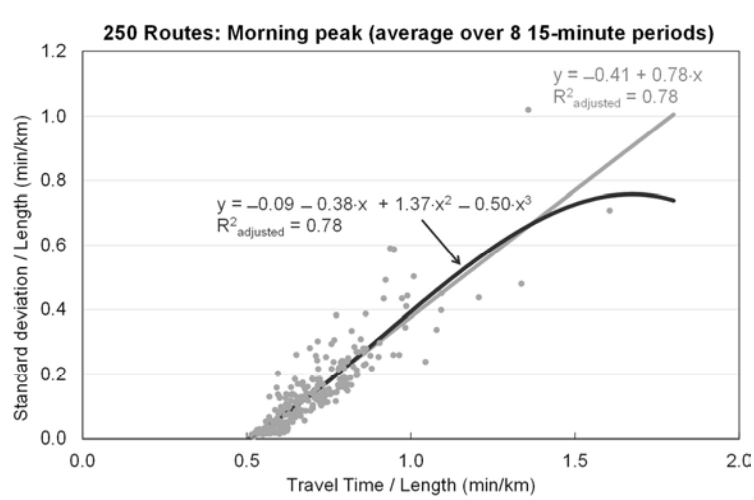
They present graphs with mean journey time per kilometre versus standard deviation per kilometre for several motorway types. The graphs for motorways with mandatory variable speed limits and with dynamic hard-shoulder running show indications for a relation that is slowly increasing for low congestion levels, increasing quickly for medium congestion levels and flattening for high congestion levels. They tried several functional forms and they obtained the best results when describing the standard deviation of travel time per kilometre as a cubic polynomial of the mean travel time per kilometre:

$$\frac{\sigma}{L} = a[0] + a[1] \cdot \frac{TT}{L} + a[2] \cdot \left(\frac{TT}{L}\right)^2 + a[3] \cdot \left(\frac{TT}{L}\right)^3 \qquad (4)$$

They observed that mean *delay* per kilometre instead of mean *travel time* per kilometre gave a marginally better fit, however, the computation of free flow travel times had some difficulties.

The fit of function (4) on our morning peak data (Figure 3.5, dark grey line) flattens above around 1.3 minutes/km which is not supported by the data. The adjusted R^2 is the same as for the fit of the linear function (3). We conclude that for the standard deviation per kilometre as a function of the travel time per kilometre a linear function is sufficient and that applying a cubic polygon does not improve the fit.

Figure 3.5. **Travel time per km versus standard deviation per km fitted with a linear relation and a cubic polynomial**



Power-law relation between coefficient of variation and congestion index (UK – ARUP/WebTAG)

Arup et al. (2003) analysed travel-time variability on urban roads by estimating a model to travel times from a few probe vehicles in London and Leeds gathered over a period of about one month. They observed that variability is likely to be greater as flows reach capacity. Based on some theoretical considerations, they estimated a power-law relation between the coefficient of variation (CV, i.e. the ratio of standard deviation to the mean travel time), the congestion index (CI, i.e. ratio of the mean travel time to the free-flow travel time) and the length of the route L:

$$\frac{\sigma}{TT} = a[0] \cdot \left(\frac{TT}{TT_{ff}} \right)^{a[1]} \cdot L^{a[2]} \tag{5}$$

in which TT_{ff} is the travel time under free-flow conditions.

A consortium led by Hyder Consulting (Hyder Consulting et al., 2008a, 2008b; Gilliam et al., 2008) collected new data from GPS equipped vehicles on 34 routes (up to 12 km long) within the 10 largest urban areas in England for a period of three years. They estimated the same function on their data and found similar coefficients as Arup et al. Today, this functional form is recommended in the WebTAG guidance of the UK Ministry of Transport (2014).

Unfortunately, this functional form does not fit our data well (adjusted R^2 is only 0.57, which is much lower than for the functions above), as can be seen from Figure 3.6 (light grey line). This is probably due to the fact that our data is for highways and this functional form was derived for urban roads. Most notably, our data strongly supports a functional form that goes (closely) through the point (CI,CV) = (1,0) whereas this functional form does not. For highways routes it is understandable that in the absence of any congestion, very little travel-time variability is observed, whereas for urban roads variability will remain through differences in signalling or pedestrian crossings.

Exponential function between coefficient of variation and congestion index (Sweden – Eliasson)

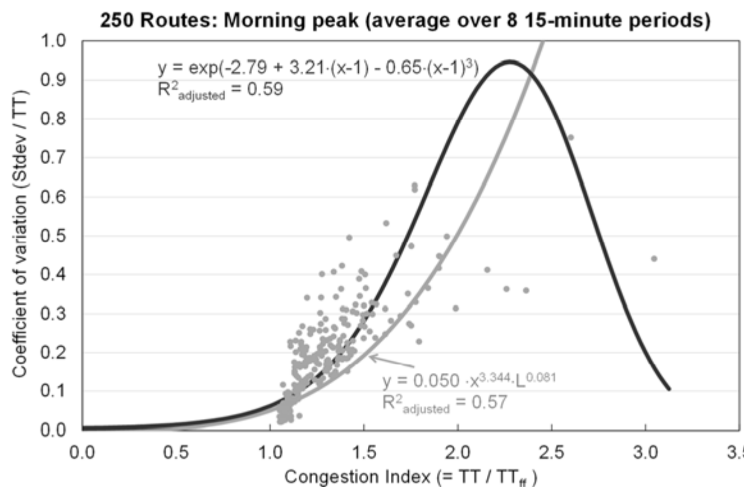
Eliasson (2006) fitted an exponential function to the coefficient of variation for 20 roads and for ninety-six 15-minute periods in Stockholm, Sweden. Data was collected from automatic camera systems taking pictures of licence plates. These roads were characterised as “urban”, i.e. neither highways, nor small local streets. Lengths varied between 300 metres and 5 kilometres.

When inspecting the relation between the congestion index and the coefficient of variation for all 15-minute periods for each road, Eliasson noticed that the coefficient of variation remained roughly constant for low levels of congestion and increased for slightly higher levels. For high levels of congestion the coefficient of variation decreased again. Therefore, he used a cubic polynomial (excluding the second-order term) of the congestion index minus 1:

$$\frac{\sigma}{TT} = \exp\left(a[0] + a[1] \cdot \left(\frac{TT}{TT_{ff}} - 1 \right) + a[2] \cdot \left(\frac{TT}{TT_{ff}} - 1 \right)^3 \right) \tag{6}$$

Again, we tried to fit this functional form to our data, but this did not lead to a satisfactory result (Figure 3.6, dark grey line), though the adjusted R^2 is slightly better than for the fit of function (5). In our data we observed neither a roughly constant coefficient of variation for low congestion levels, nor a decreasing coefficient of variation for high congestion levels. This different behaviour is likely due to the fact that our database consists of longer highway routes rather than short urban routes. Furthermore, we investigate the variation between routes whereas Eliasson also included the variation between 15-minute periods.

Figure 3.6. **Congestion index versus coefficient of variation fitted with a power law and with an exponential function**



Power-law relation between standard deviation and mean delay (Germany – Geistefeldt et al.)

Recently, the German Federal Ministry of Transport (BMVBI) funded a research project on the reliability of travel time on their highways. Geistefeldt et al. (2014) suggested using a power-law function between the standard deviation and the mean delay (i.e. the difference between the mean travel time and the free flow travel time):

$$\sigma = a[0] \cdot MD^{a[1]} \tag{7}$$

where *MD* is the mean delay. They estimated their coefficients on simulated data from a macroscopic traffic simulation model.

This functional form seems to describe our data well (Figure 3.7, light grey line). Note that the spread of the data points in Figure 3.7 is small compared to the spread when relating the standard deviation to the travel time (Figure 3.4) or when relating the travel time per kilometre to the standard deviation per kilometre (Figure 3.5). The adjusted R^2 of 0.82 is the better than for the fits previously discussed. So, using the mean delay as the explanatory variable seems to be a good idea.

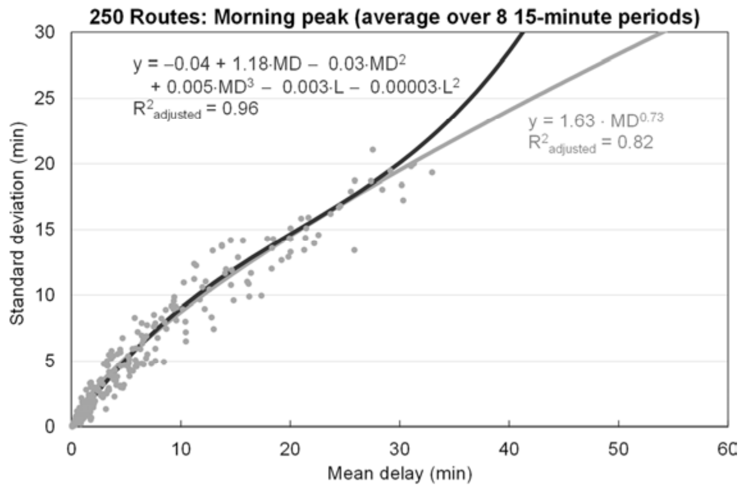
Polynomial of mean delay and length (the Netherlands – Peer et al.)

Peer et al. (2012) estimated a relation between the standard deviation and the mean delay. In her PhD research she tried multiple functions on data from 145 highway routes and fifty-seven 15-minute periods. The best function contained (among other terms) a cubic polynomial in the mean delay and a quadratic polynomial in the length:

$$\sigma = a[0] + a[1] \cdot MD + a[2] \cdot MD^2 + a[3] \cdot MD^3 + a[4] \cdot L + a[5] \cdot L^2 + other\ terms \tag{8}$$

This function fits our data very well (note the adjusted R^2 of 0.96 in Figure 3.7, dark grey line). However, the slope of the fitted function for the morning peak seems to steepen above a mean delay of about 30 minutes which is not supported by the data. So, the cubic polynomial may lead to unwanted behaviour outside the range on which it was fitted.

Figure 3.7. **Travel time versus standard deviation fitted with a power law and a cubic polynomial**



A new empirical relation for the Netherlands

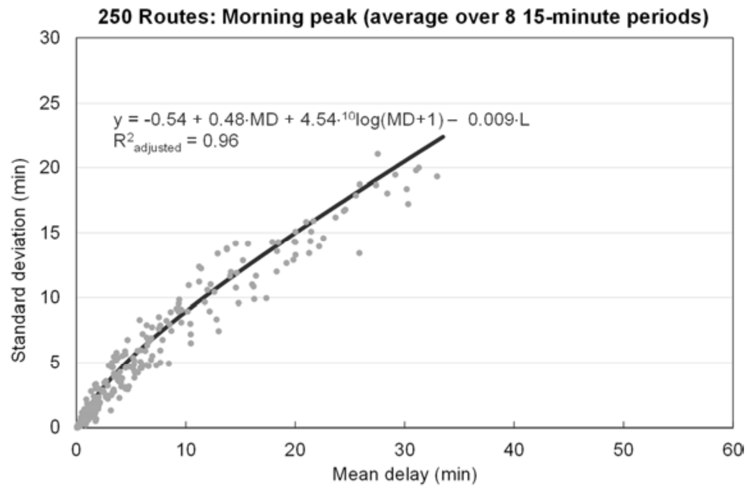
Best functional form

From Chapter 4 we conclude that the best results are obtained when relating the standard deviation to the mean delay. However, using a cubic polynomial may not be optimal. Therefore, we decided to test a combination of a linear and logarithmic function for the mean delay, and added a linear term in the length. Higher order terms and terms proportional with other parameters such as density, number of lanes, average weather conditions, and frequency of incidents, were not found significant.

$$\sigma = a[0] + a[1]\cdot MD + a[2]\cdot \log(MD+1) + a[3]\cdot L \qquad (9)$$

For the morning peak data, this fits the data very well (Figure 3.8) and does not lead to unwanted behaviour outside the range on which it was fitted. This function is selected as our final relation to forecast standard deviation based on the elements available from the traffic model.

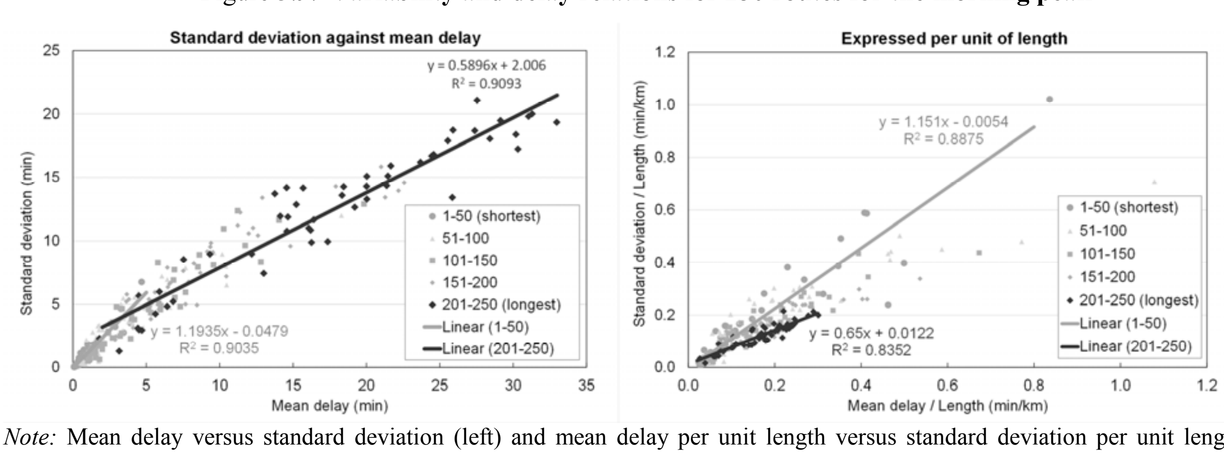
Figure 3.8. **Mean delay versus standard deviation fitted with a combination of a linear and a logarithmic function**



Dependence on trip length

When we look at the estimation result for our final function, we note that the coefficient on length is very small (only 0.009 as can be seen from Figure 3.8). One might conclude that the length is not important. However, length is also related to mean delay: the longer the route, the more likely it is that some congestion occurs. This becomes clear when we divide all data points based on their length. In Figure 3.9-left, the 50 shortest routes are displayed as light grey circles, while the longest routes are shown as dark grey diamonds. We see that all the light grey circles are on the left of the diagram, while the dark grey diamonds are on the right. We have made a linear regression on longest and shortest routes. We see that the slope decreases with length as well. This property is also clearly visible when we plot the standard deviation per kilometre as a function of mean delay per kilometre (Figure 3.9-right): the slope of their linear relation is correlated with length. This can be intuitively understood: if for a long route on a certain day, the congestion is worse than normal, traffic might flow better downstream, so any delay can be (somewhat) compensated later along the route. This will reduce the variation of the day-to-day travel time.

Figure 3.9. **Variability and delay relations for 250 routes for the morning peak**



Note: Mean delay versus standard deviation (left) and mean delay per unit length versus standard deviation per unit length (right). Light grey dots indicate the 50 shortest routes (less than 12.6 km) and the light grey line is the linear regression through these points. Similarly the dark grey diamonds indicate the 50 longest routes (above 63 km) and the dark grey line is the linear regression through these points.

Differences between time-of-day periods

We used the same functional form to analyse the evening peak (16:00 – 18:00) data and the mid-day data (10:00 – 15:00), see Table 3.2 for the estimates of the coefficients. Even though the three periods have significantly different coefficients (based on an F-test), the functional form fits each data set well.

For the mid-day period, we did not find *a*[2]- and *a*[3]- coefficients that were significantly different from zero. Therefore, we tried a fit with these coefficient constrained to zero, effectively turning equation (9) into a linear equation. This is understandable since the maximum mean delay for the mid-day period is only 10 minutes. Even in the morning peak the observations in Figure 3.10 below a mean delay of 10 minutes almost follow a straight line. Note that we kept the *a*[0] constant though it is not significantly different from zero, since we did not want to force the function to go through the point with $(MD,\sigma) = (0,0)$.

Table 3.2. **Best fit coefficients for the empirical relation between the standard deviation and the mean delay (equation 9) for highway routes**

	Morning peak			Mid-day period			Evening peak			Units
	Coefficient		(t-ratio)	Coefficient		(t-ratio)	Coefficient		(t-ratio)	
a[0]	-0.540	± 0.186	(-2.9)	-0.066	± 0.051	(-1.3)	-0.901	± 0.172	(-5.3)	*min.*
a[1]	0.476	± 0.026	(18.2)	1.034	± 0.019	(53.1)	0.268	± 0.017	(16.1)	
a[2]	4.538	± 0.415	(10.9)	-			5.555	± 0.351	(15.8)	*min.*
a[3]	-0.009	± 0.003	(-2.7)	-			0.011	± 0.003	(4.0)	*min./km*
Adj. R^2	0.956			0.919			0.960			

Note: for the mid-day period no significant value for the a[2] and a[3] coefficients was found.

Impact of excluding outliers

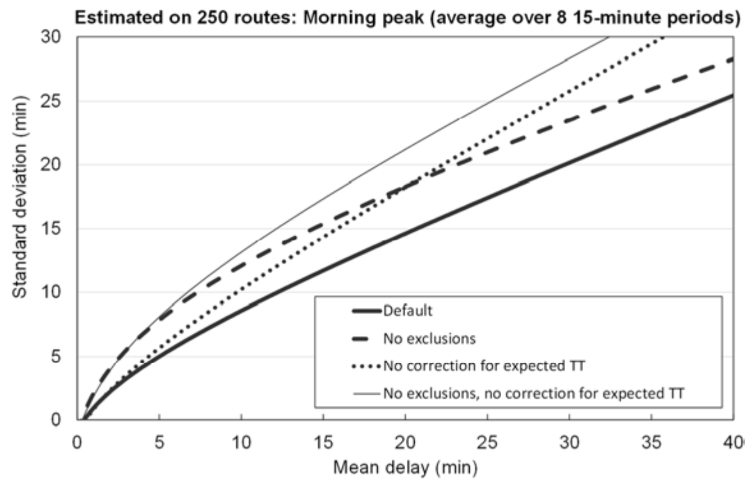
We noted above that the exclusion of the outliers caused a decrease of the average standard deviation of 29%. However, excluding outliers also has an impact on the mean delay. So, it is theoretically possible that the data before and after exclusion fall on the same line: the exclusions may only cause a shift along the line. To test this, we fitted a function on the full data set (no exclusions), but with the correction for expected travel time as discussed. The resulting best fit is the dashed line in Figure 3.10, which lies roughly 3 minutes above the default (solid) line. From this, we conclude that outliers have more impact on the standard deviation than on the mean delay and excluding the outliers has a strong impact on the coefficients of the relationship.

Impact of correcting for travel-time expectation

The correction for the expected travel time (see the third step in the Data section) influences the standard deviation but not the mean delay. So, we expect that without this correction, the empirical relation between mean delay and standard deviation would shift upwards. This can indeed be seen from Figure 3.10. If no correction is made for expected travel time (dotted line), the curve is located up to 30% above the default line.

We conclude that both the outlier criterion and the travel-time expectation correction both have a clear impact on the coefficients (though our best functional form still describes the data well). As such, comparisons of coefficient values between different studies are not very useful unless these studies use exactly the same outlier criterion and travel-time expectation correction.

Figure 3.10. **Results of fits for mean delay versus standard deviation under several choices of the data analysis**



Results for other roads

We also fitted the same functional form to our database of 40 routes on other (non-highway) roads. Since these routes are small compared to the highway routes (see Table 3.1), we also have relatively small mean delays and standard deviations. As a result, only the linear term in equation (9) was found to be significant. Figure 3.11 shows the data and the fit for the morning peak. Table 3.3 shows the best fit coefficients for all time-of-day periods. Note that insignificant constants were kept in the models. The coefficients for the evening peak are significantly different from those for the morning peak. The coefficients for the mid-day period are significantly different to those of the evening peak, but not from those of the morning peak. Also note that the slope for the morning peak (0.468) is much lower than the slope for the same period for short highway routes (1.1935, see Figure 3.9-right). So, the reliability relation for other roads is clearly different from that for highways.

Figure 3.11. **Mean delay versus standard deviation fitted with a linear function**

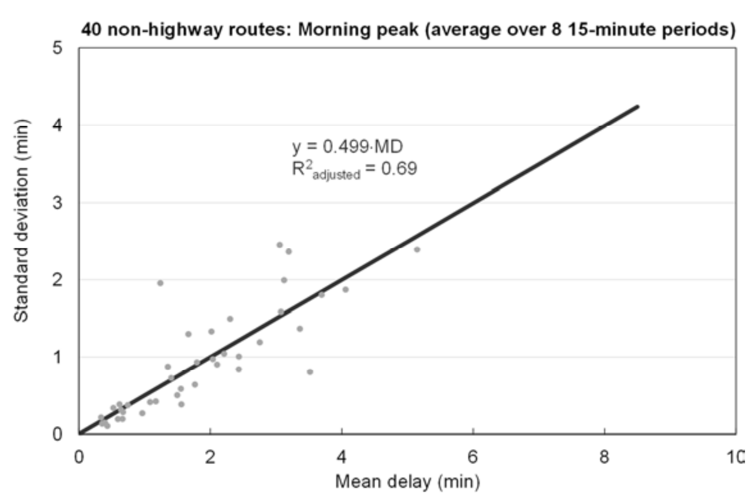


Table 3.3. **Best fit coefficients for the empirical relation between the standard deviation and the mean delay (equation 9) for other routes**

	Morning peak			Mid-day period			Evening peak			Units
	Coefficient		(t-ratio)	Coefficient		(t-ratio)	Coefficient		(t-ratio)	
a[0]	0.049	± 0.120	(0.4)	-0.074	± 0.049	(-1.5)	-0.079	± 0.106	(-0.7)	*min.*
a[1]	0.468	± 0.054	(8.7)	0.534	± 0.030	(17.6)	0.637	± 0.044	(14.6)	
a[2]	-			-			-			*min.*
a[3]	-			-			-			*min. / km*
Adj. R^2	0.662			0.891			0.848			

Note: "-" indicates that no significant value was found.

Policy implications

Current treatment of reliability in CBA

Until now, reliability is included in Dutch CBAs using the practical and provisional way as developed by Besseling et al. (2004). That means that reliability benefits are included by multiplying the travel-time benefits from reduced congestion by a factor of 1.25. This proportionality is based on the linkage between congestion reduction and reliability improvements.

However, this current treatment of reliability in CBA is not useful to evaluate policies that especially affect travel-time variability. An approach to reflect the effects of policies that affect travel-time variability in CBA will encourage proper considerations of options. Project appraisal will then not only offer incentives for policies that reduce the average travel time but also for policies that improve travel-time variability.

Better capturing the effects of policies that affect travel-time variability

The new travel-time reliability forecasting model does not require any adjustments with respect to the transport model. It is a separate module that uses outputs from the transport model to forecast the impact of infrastructure projects on travel-time variability. It is a so called post-processing module. Its outputs will not feed back into the transport model. That means that the reactions of network users to changes in reliability are not incorporated in the predicted levels of reliability.

The empirical relations presented in the previous section were built into this post-processing module. Based on a LMS/NRM scenario, this module calculates the value of the reliability indicator for each origin-destination pair. However, due to the iterative assignment process in the LMS/NRM, multiple routes can be assigned to people travelling between an origin and a destination. Our post-processing module repeats this route assignment and stores all routes in each iteration step. Once the final link travel times have been calculated, our module loops back to all these routes and calculates the reliability for each of them using equation (9) and the coefficients from Tables 3.2 and 3.3. The final value of the reliability indicator for an origin-destination pair is an average of the reliability indicators in each iteration step weighted by the flow assigned in that step.

If a route travels over both highways and other roads, the reliability indicator is calculated for both road types separately. The total reliability for this route is the root of the squared sums of these two reliability indicators. Implicitly, we have assumed here that travel-time delays on highways are not correlated with those on other routes. A (limited) analysis of our data has shown that this correlation is indeed small, so this is a reasonable assumption.

The module also calculates a value for the national (or regional) reliability indicator by adding the standard deviations of all origin-destination pairs weighted by their traffic flows. These totals are calculated for each time-of-day period and can be added to get a reliability indicator for a whole day, using a weight of 2 for both peak values. We analysed 24-hour data to derive that the mid-day period should get a weight of 9.5 to get a correct daily total.

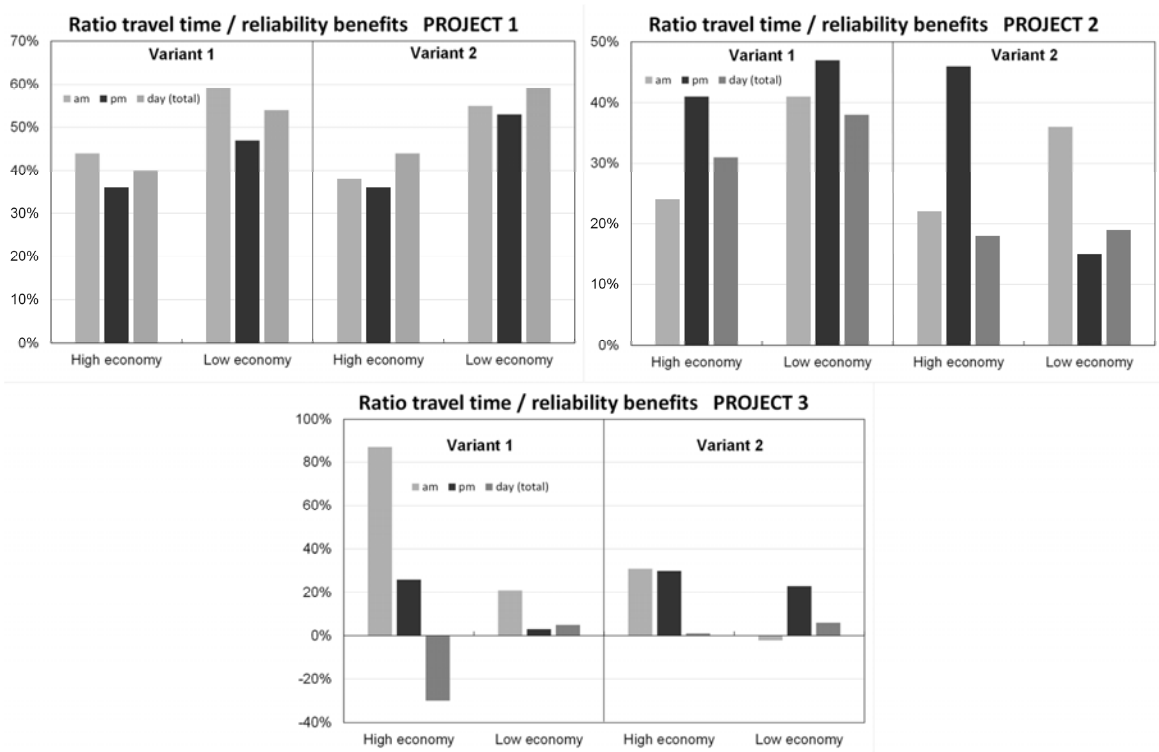
A test run with this new module revealed that in 2004 the reliability indicator (i.e. the summed standard deviations over all origin-destination pairs) was 48 400 hours for one hour in the morning peak; 60% of this originated from highways and 40% from other roads. The corresponding LMS-run showed that the total delay for all travellers in one hour in the morning peak was 77 000 hours, so the ratio for the (national) reliability to the travel-time delays is 63%.

The new module clearly makes a distinction between travel-time gains due to shorter routes (e.g. a bypass) or due to reduction of congestion. The former does not lead to reliability gains, whereas the latter does. If some mild congestion occurs on the bypass, reliability may even deteriorate, while overall travel times go down. The new module also takes the exchange of traffic between highways and other roads into account. When congestion is reduced on a highway, this may cause diversion of traffic from secondary roads which can lead to non-standard amounts of reliability gains and losses due to the different relations to mean delay on both types of roads. Also, a lower maximum speed will lead to a reduction of mean delays and hence better reliability which is indicative for the more uniform traffic flow (though we have not yet tested the size of the effect of this policy with observed data).

Impact on CBA results

As a test, a research team from 4Cast simulated the reliability effects of several future infrastructure projects with the new reliability forecasting model described in this chapter. The reliability benefits (in terms of euro, i.e. the changes in reliability and multiplied by the VTTRS) appeared to be between 15% and 60% of the travel-time benefits, though higher and lower values also occurred (depending on the project, the time-of-day period and the economic scenario, see Figure 3.12). These numbers are of the same order-of-magnitude as the initial rule-of-thumb of adding 25% to the travel-time benefits (Besseling, et al. 2004). The range between the projects is mainly caused by the differences in trip length and the amount of traffic.

Figure 3.12. **Ratio of travel-time benefits and reliability benefits for three projects, each for two variants and two economic scenarios**



Source: Authors adapted from 4Cast.

Conclusions and future steps

The most important findings of this chapter are:

- When fitting functions between reliability and parameters that are available in demand models, distinction should be made between variation between 15-minute periods and variation between routes as each may be described by a different function.
- For our data set consisting of 250 routes on Dutch highways for three periods of the day (morning peak, evening peak, mid-day), the best empirical relation to describe reliability was an expression of the standard deviation as a function of the mean delay, the logarithm of the mean delay and the length of the route. Other functional forms that have been described in the literature either had much lower adjusted R^2 or showed behaviour that was not supported by the data.
- If observed travel times over multiple days are used to compile a travel-time distribution, a decision needs to be made whether to exclude outliers and whether to correct for variation in travel-time expectation. Consistency with the congestion functions in the demand model and

with the method used to determine the valuation of travel-time reliability should be leading in this decision.

- Excluding outliers can have a profound impact on the standard deviation and on the coefficients of the empirical relation. In our project, a criterion of three times the standard deviation above the mean travel time with a minimum of 150% of the mean separated the visually clear outliers from the tail of the standard travel-time distribution. On average, the travel time on 4 out of 251 working days exceeded this criterion. Excluding them reduced the standard deviation by 29%.
- Correcting for variation in travel-time expectation also reduced the standard deviation by another 12%. This depends on the method used to calculate the expected travel times. Very little research is available to support the selection of a method for this.
- This unreliability model is built into a post-processing module for the national and regional transport models. These transport models have not been altered but their outputs are used to calculate the changes in the standard deviation of the travel-time distribution due to an infrastructure project.
- The post-processing module calculates the reliability benefits (i.e. the changes in the standard deviation expressed in hours and multiplied by the VTTRS) which can be used in a CBA.
- Better capturing the effects of policies that affect travel-time variability in CBA will encourage proper consideration of options. Project appraisal will then not only offer incentives for policies that reduce the average travel time but also for policies that disproportionately improve travel-time variability.

Future steps

Reliability will be better embedded in the transport policy-making process by the following concrete policy actions. First, the new travel-time reliability forecasting model will be incorporated in the Dutch guidelines for CBA. Second, the consequences of policies that especially affect travel-time variability will be part of CBA. Attribution of an economic value to travel-time variability recognises that transport projects can create more value than they have traditionally realised when they invest to reduce congestion if an improvement in reliability is produced independent of a reduction in travel time. Third, in order to properly consider such investments in the resource allocation decision process they will be included in the investment tradeoff analysis to prioritise, rank and select infrastructure improvement projects. And finally, a guideline on also including the consequences of extreme travel times, network robustness and vulnerability into the decision making process will be developed. However, a special VTTRS to value extreme changes of travel-time variability in the CBA does not exist.

Better integration of reliability into transport policy making is synthesised into a short- and mid-term strategy as discussed below to improve the post-processing reliability model. However, it is recommended that these future steps are embedded in a long term strategy (10+ years) to be developed for the national and regional models to assess unreliability in CBA. The basis for such a strategy can be to identify the set of policy measures for which evaluations are or likely will be required. These policies should be matched against the capabilities of the set of modelling tools available.

Short-term improvements of the post-processing reliability model

- The reliability model only deals with road transport. However, the national Dutch transport model is also capable of forecasting the effects of changes in the average travel times for public transport (train and bus/tram/metro). It should be possible to estimate equations explaining the standard deviation of travel time for public transport from explanatory variables available in

LMS or NRM. At the moment of writing this chapter, KiM works on a project to measure how different policies affect travel-time reliability in public transport chains.

- Dutch highways are well equipped with detector loops providing inputs for the transport model. However, network users make trips on other roads as well. The regression line is fitted on 250 highway routes and 40 routes on other roads. Collecting extra data and expanding the database can improve the regression analysis for non-highway routes.

Long-term improvements of the post-processing reliability model

- Build a specific database for policies which will increase the travel time but may decrease unreliability. These are policies such as changing the maximum speed or ramp metering. Based on this database a specific regression line can be fitted.
- In reality, mode choice, departure time choice and route choice are sensitive to reliability. The post-processing reliability model can be extended with a feed-back loop into the transport model so that the decisions of the network users are impacted explicitly by changes in reliability.
- The standard deviation contains several sources of unreliability, namely due to recurrent congestion, road works, accidents, unexpected weather conditions, and a random component of day-to-day variation in travel times. Extreme events are removed from the data before fitting the function. Therefore the model predicts reliability changes without considering extremes. Analysing the extreme events, will provide insight in the robustness and vulnerability of the network. However, a special VTTRS to value extreme changes of travel-time variability in the CBA would need to be developed through additional primary research.

Acknowledgements

This research was sponsored by the Dutch Ministry of Infrastructure and the Environment. We would like to thank Jasper Willigers (Significance) and Nick Bel (currently ProRail) for their help with the data analysis. We thank 4Cast for their tests with the newly developed tool.

We thank Arjen ’t Hoen, Jan van der Waard (both KiM Netherlands Institute for Transport Policy Analysis), Henk van Mourik (Ministry of Infrastructure and the Environment), Marcel Mulder (Rijkswaterstaat – Water Traffic and Environment, Ministry of Infrastructure and the Environment) and Gerard de Jong (Significance) for their comments and suggestions.

Notes

1 Fosgerau (Chapter 2) prefers to use the variance of travel time rather than the standard deviation. In his paper, he shows that the variance is theoretically more appropriate for commuters with flexible work times. Furthermore, using the variance has the advantage that it is additive over links (provided that the travel times on the links are independent). However, for this study we prefer to use the standard deviation since (a) it is consistent with the Dutch valuation study, (b) most travellers in the peaks are commuters with inflexible work times and (c) the typical link lengths in our study are so short that the travel times on adjacent links are certainly correlated.

2 Extreme travel times are removed from the data before fitting the function. Therefore, our model predicts reliability changes without considering these extremes. Analysing extreme events (separately) will provide insight in the robustness and vulnerability of the network.

Policies may affect the reliability, but also may affect the robustness and the vulnerability of the network. Both should be included in a cost-benefit analysis. The tool that we describe in this paper only looks at the reliability component without the extreme events. Including the extreme events in this or in a separate tool is one of the longer-term improvements of the post-processing reliability model.

3 All estimations in this paper were made using the LFIT algorithm of Numerical Recipes (Press et al., 1992).

References

Besseling, P., W. Groot and A. Verrips (2004), “Economische toets op de Nota Mobiliteit”, CPB Document 65, Centraal Planbureau, The Hague.

de Jong, G.C. and M.C.J. Bliemer (2015), “On including travel time reliability of road traffic in appraisal”, *Transportation Research A*, Vol. 73, pp. 80-95.

de Jong, G.C., M. Kouwenhoven, J. Bates, P. Koster, E.T. Verhoef, L. Tavasszy and P.M.J. Warffemius (2014), “New SP-values of time and reliability in freight transport in the Netherlands”, *Transport Research Part E*, Vol. 64, pp. 71-87.

Eliasson, J. (2006), “Forecasting travel time variability”, paper presented at European Transport Conference 2006, Strasbourg.

Geistefeldt, J., S. Hohmann and N. Wu (2014), “Ermittlung des Zusammenhangs von Infrastruktur und Zuverlässigkeit des Verkehrsablaufs für den Verkehrsträger Straße”, Schlussbericht für Bundesministerium für Verkehr und digitale Infrastruktur, March 2014.

Gilliam, C., T. Kean Chin, I. Black J. and Fearon (2008), “Forecasting and appraising travel time variability in urban areas”, paper presented at ETC 2008, Noordwijkerhout.

Hamer, R., G.C. de Jong and E.P. Kroes (2005), “The value of reliability in Transport – Provisional values for the Netherlands based on expert opinion”, RAND Technical Report Series, TR-240-AVV.

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

Hellinga, B. (2011), “Defining, measuring, and modeling transportation network reliability”, report written for the Dutch Ministry of Transport, Rijkswaterstaat, Delft.

Hyder Consulting, I. Black and J. Fearon (2008a), “Forecasting Travel Time Variability in Urban Areas”, Deliverable 1: Data Analysis and Model Development, Department for Transport, United Kingdom.

Hyder Consulting, I. Black and J. Fearon (2008b), “Forecasting Travel Time Variability in Urban Areas”, Deliverable 2: Model Application, Department for Transport, United Kingdom.

ITF (2010), *Improving reliability on surface transport networks*, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/9789282102428-en

KiM Netherlands Institute for Transport Policy Analysis (2013), “The social value of shorter and more reliable travel times”, https://english.kimnet.nl/publications/reports/2013/11/18/the-social-value-of-shorter-and-more-reliable-travel-times

Kouwenhoven, M., G.C. de Jong, P. Koster, V.A.C. van den Berg, E.T. Verhoef, J. Bates and P.M.J. Warffemius (2014), “New values of time and reliability in passenger transport in the Netherlands”, *Research in Transportation Economics*, Vol. 47, pp. 37-49.

Mahmassani, H.S., J. Kim, Y. Chen, Y. Stogios, A. Brijmohan and P. Vovsha (2014), “Incorporating reliability performance measures in planning and operations modelling tools”, SHRP2 report S2-L04-RR-1, Transportation Research Board, Washington DC.

Ministry of Infrastructure and the Environment (2012), “Structuurvisie Infrastructuur en Ruimte (SVIR)”, The Hague, http://www.rijksoverheid.nl/documenten-en-publicaties/rapporten/2012/03/13/structuurvisie-infrastructuur-en-ruimte.html (accessed 8 December 2015).

Ministry of Transport (2014), “Transport analysis guidance: WebTAG”, https://www.gov.uk/guidance/transport-analysis-guidance-webtag (accessed 8 December 2015).

Ove Arup and Partners Ltd, J. Bates, J. Fearon and I. Black (2003), “Frameworks for Modelling the Variability of Journey Times on the Highway Network”, Arup, London.

Peer, S., C. Koopmans and E.T. Verhoef (2012), “Prediction of travel time variability for cost-benefit analysis”, *Transportation Research Part A*, pp. 79-90.

Press, W.H., S.A. Teukolsky, W.T. Vetterling and B.P. Flannery (2007), *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., Cambridge University Press, New York

Significance, VU University Amsterdam and J. Bates (2007), “The value of travel time and travel time reliability, survey design”, Final Report prepared for The Netherlands Ministry of Transport, Public Works and Water Management, Significance, Leiden.

Significance (2011), “Schattingen van keuzemodellen voor het LMS 2010”, Report prepared for RWS, Significance, The Hague.

Sirivadidurage, S., A. Gordon, C. White and D. Watling (2009), “Forecasting day to day variability on the UK motorway network”, presented at ETC 2009, Noordwijkerhout.

TRB (2003), “Interim Planning for a Future Strategic Highway Research Program”, NCHRP national Cooperative Highway Research Program, Report 510, Washington DC.

Willigers, J. and M. de Bok (2009), “Updating and extending the disaggregate choice models in the Dutch national model”, paper presented at the ETC 2009, Noordwijkerhout.

Chapter 4

Estimating wider economic impacts in transport project prioritisation using ex-post analysis

Glen Weisbrod[1]

Transport project prioritisation and selection processes require consideration of many aspects of costs, intended benefits and other impacts. Economic analysis methods can measure many of those factors, though the analysis methods must be specified in ways that meet the information needs of decision-makers. This chapter examines how cost-benefit analysis, economic impact analysis and multi-criteria analysis approaches have evolved and been applied to address the specific form of governmental decision processes that exist in the U.S. and some other countries. It discusses how "ex-post" case studies and associated statistical studies of have been promoted and utilised to both inform and refine "ex-ante" evaluation methods. It concludes by discussing the advantages, limitations and trade-offs involved in the use of this approach for transport project decision making.

1 Economic Development Research Group, Boston, MA, USA.

Introduction

Decisions regarding transport projects, programs and policies are made at several stages – including planning, prioritising, funding and implementation. These decisions are best made when they are informed by complete and accurate information regarding requirements and consequences of following through on those decisions. Various forms of economic impact and cost-benefit analysis (CBA)are frequently applied to inform these decisions, though the specific approaches have varied across countries and governments over time. Overall, there has been a common progression (evolution) towards more complete and accurate information, and efforts to better align analysis methods with decision needs.

This chapter examines one aspect of this evolution, which pertains to the definition, measurement and use of non-user impact measures in transport decision making. It further focuses on non-user economic impacts and their wider economic development consequences. It discusses how and why the definition and measurement of these effects differ among (national and state or regional) governments, and specifically how "case-based evaluation" methods and "*ex-post* case studies" are being increasingly used in the US to inform decision making.

Much of the chapter focuses on the evolution of evaluation and decision processes in the US, with some parallel examples drawn from Canada and Australia. The presentation is intended to provide an instructive example of how these issues are being addressed in the context of a very specific federal system of government. There is no intent to suggest that the approaches discussed in this report are necessarily applicable to other forms of governmental decision making. Rather, the discussion is intended to highlight a core issue – how can cost-benefit (efficiency) analysis exist alongside broader methods that consider distributional equity and strategic policy factors in decision making. These broader methods can include other forms of economic analysis, including economic development and financial impact evaluation methods.

The rest of this chapter is organised into four parts. First, this chapter defines benefits, and identifies available methods to assess them. It considers the motivations and intended benefits of transport projects and how those goals can or should affect the evaluation (appraisal) process. It then reviews the evolution of economic evaluation methods over time to encompass cost-benefit analysis (CBA), economic impact analysis (EIA) and financial impact analysis (FIA) – and the intent of these methods to facilitate better accounting of goal achievement and impact effects.

Second, the chapter reviews the development and growth of *ex-post* case study analysis, *ex-ante* multi-criteria analysis and evidence-based planning methods in the US. It discusses the motivation for these approaches, and specifically their ability to cover threshold effects, interaction effects, and distributional effects. It also examines the consequences of these approaches for decision making at various planning stages.

Third, the chapter critically examines the strengths and weaknesses of *ex-post* analysis and evidence based planning methods, and extracts key impact factors that need to be better considered in decision frameworks. This includes needs to recognise how the creation of new spatial, temporal and distributional linkages may affect technology adoption and activity patterns in ways far broader than efficiencies for existing transport system beneficiaries.

Finally, the chapter discusses how the study findings can be incorporated into cost, benefit and impact accounting systems, and how that information can be used to better communicate to public and private stakeholders and decision-makers.

Defining and measuring wider economic benefits and impacts

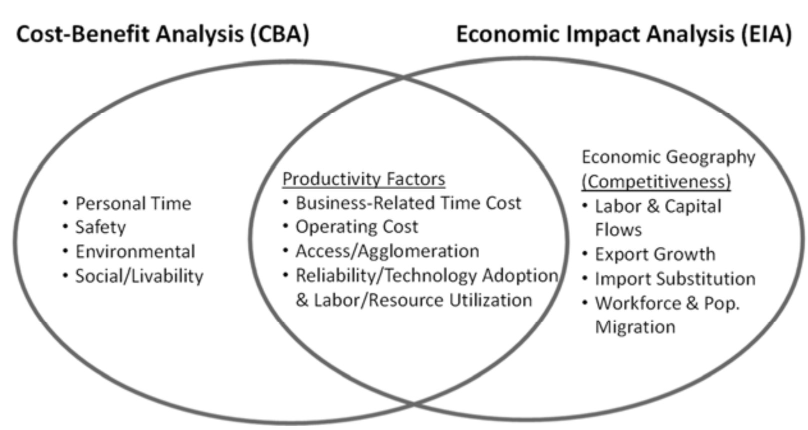
Benefit and impact definitions

To understand the evolution of methods for evaluating wider economic impacts (and their use for prioritising and selecting projects), it is necessary to first establish the criteria used to define the concepts of economic benefits and impacts. The definitions of these concepts, in turn, trace back to three separate lines of economic analysis that each reflect a different perspective and purpose in transport evaluation processes. The three lines of analysis are: CBA, EIA and FIA. Their inter-relationships, differences and uses have been discussed in a number of prior papers and reports (for instance Thompson, 2008; Weisbrod et al., 2015). While they can complement each other and can be used together to inform decision making, they also yield very different ways of viewing economic benefits and impacts.

The Venn diagram in Figure 4.1 shows the overlap in coverage between CBA and EIA – as practiced in the US, Canada and sometimes also Australia. The core overlap area comprises business-related money benefits, the left side comprises non-money benefits to society, and the right side comprises wider consequences for the development of a given region's economy. Elements considered in CBA are referred to as "benefits" and elements considered in EIA are referred to as "economic impacts." The latter are usually calculated using a regional economic simulation model.

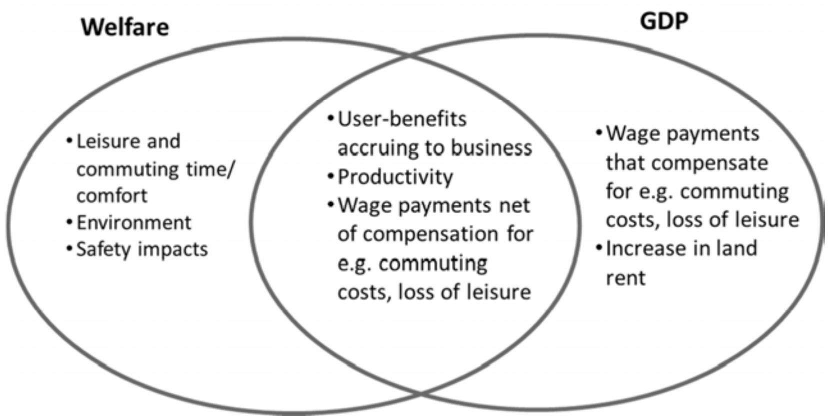
The Venn diagram in Figure 4.2 shows the overlap between welfare benefits and GDP impacts in UK Dept. for Transport documents. While it bears a superficial resemblance to the diagram above it, there are fundamental differences. Most critically, the second diagram focuses only on CBA measurement and it illustrates the distinction between elements of GDP that can and cannot be included as welfare gains in CBA. This second diagram takes a partial equilibrium perspective and hence does not attempt to represent the full range of GDP effects (associated with labour, capital and trade flows) that could be represented in a separate EIA study.

Figure 4.1. **Distinctions between CBA and EIA (US)**



Note: This also draws from earlier studies: see Wang, 2015; Weisbrod et al., 2015, 2006.
Source: Weisbrod, Mulley and Hensher, 2015.

Figure 4.2. **Distinctions between CBA and EIA: Coverage of welfare and GDP effects (UK)**



Note: This also draws from earlier studies: see Department for Transport (UK), 2005.
Source: Venables, Laird and Overman, 2014.

As will be discussed, an underlying reason for the difference in these two approaches is that project prioritisation and selection decisions in the US are largely made at the state level. At that level of government, EIA is widely conducted and the full set of economic impact factors shown in Figure 4.1. are widely considered as elements of a *strategic policy goal* for economic development. These impact elements are considered alongside user benefits in prioritisation and selection processes. This theme of differences in the context for measuring economic development impacts, and in how those measures are applied, is explored in more detail in the remainder of this chapter.

The title of this chapter uses the phrase "wider economic impacts" to emphasise that it is examining the full set of economic development impact elements that are considered in EIA studies. In the US, these are referred to as "economic impacts" but there is a need to differentiate this broad category from use of the same phrase in the UK's transport guidance, where the term "economic impact" has now been assigned to the narrower category of what used to be called "wider economic benefits (WEB)." That latter concept (WEBs) is defined as the middle portion of Figure 4.2, excluding user benefits accruing to business.

Criteria for evaluating proposed projects

It is important at the outset to establish both the process and criteria used for *ex-ante* evaluation of proposed projects because they will dictate the kind of information that is needed. In the federal system of the US, central government does not have responsibility for prioritisation and selection decisions for surface transport projects. Rather, state departments of transport (DOTs) usually make those decisions with metropolitan planning organisations (MPOs) leading the process for urban areas. There is a formal process of steps that starts with a declaration of policy goals (guiding principles) and an evaluation of alternatives for long range vision plans that address the policy goals. There is then a screening and prioritisation of proposed projects that are consistent with the long range plan, which are then listed on a five- or six-year TIP (transportation improvement plan). Selected major projects go through an "alternatives analysis" that evaluates alternative project location and design options, and the selected plan is then moved forward for final funding and implementation decisions. (This process is described in ICF, 2009.) This progression of steps has a direct parallel in many other countries.

A recent study examined the formally stated goals of the long-range visions or plans of US State DOTs. It found that two-thirds of the states cited the same four key strategic goals for their residents: safety, mobility/accessibility, environmental stewardship and economic development. In other words, the ultimate outcomes desired by the long range transport vision plans were to improve the lives of people and their living/working environment. The efficiency of movement was far less frequently cited, meaning that it was most often seen as an intermediate consideration rather than as an ultimate outcome (Volpe, 2012).

The long-range goals are important because they provide a basis for subsequent project planning, prioritisation, design and funding steps. At each step, there is a need for consideration of the costs as well as benefits or impacts of proposed alternatives, with benefit metrics defined in line with stated goals. While the decision processes at each of these steps is different, the methods used for evaluating alternative proposals (scenarios, schemes and projects) generally require three attributes:

- Relevance: To be relevant, an evaluation should consider project motivations and objectives, and assess the extent to which the project achieves those intended consequences as well possible unintended consequences. It should also consider the project "requirements" – money investment and non-money actions that need to occur in order for the project to go forward. Any misalignment of evaluation and objectives can represent a gap that undermines the usefulness of the evaluation for decision makers.
- Practicality: To be practical, an evaluation should be capable of discerning differences among competing projects in different settings. It should recognise cases where local project settings and contexts shift the upside and downside likelihood of impacts among alternatives. Anything less may fail to distinguish among competing projects that have similar passenger volumes and traveller savings but vastly different settings and contexts, and hence differences in the nature of their wider impacts. This distinction can be critical for prioritising and selecting projects within a given budget.
- Accuracy: To be accurate, an evaluation should discern needs and impacts relative to threshold factors concerning: (a) minimal acceptable conditions and (b) reasonable ranges for travel times and transport costs. Anything less may fail to distinguish projects that are addressing critical local needs and deficiencies, from other projects where needs are less severe. They may also fail to distinguish projects where impacts are too small or widely dispersed to have any real impact on behaviour or economic outcomes, from projects that have impacts sufficiently large and concentrated to have very observable and desired impacts.

Consideration of these three attributes, in combination with state-level decision making, can help explain the relative use of cost-benefit analysis and multi-criteria analysis techniques in the US. They also help explain the interest in *ex-post* analysis and evidence-based planning. This chapter examines these connections. It reviews the evolution of project evaluation methods in the US by considering how they have been evolving to better address the three criteria. We first look at the issue of relevance, in order to better understand how benefit evaluation methods are evolving to better account for project motivations. The practicality and accuracy issues are picked up later in this chapter, insofar as *ex-post* case studies are helping to illuminate the needs and opportunities to improve on these matters.

Historical background: The role of wider benefits in investment objectives

The first matter is to establish the range of project motivations that may be *relevant* considerations in the definition of project benefits. It can be useful to start with a wide view of how project motivations have been viewed, and how project goal achievement has been assessed over time.

Concern over project benefits and costs is not new. Processes of planning, funding and implementing transport projects, programs and policies have been going on for several thousand years – back to ancient times. This includes caravan services in ancient Mesopotamia, tourism cruises on the Nile in Egypt, passenger horse cart services in villages of ancient Greece, and land/sea intermodal trade centres across the Roman Empire (Casson, 1994; Bernstein, 2008). More recently (in the last four centuries), there have been major investments in canals and waterways, urban transport, intercity rail and highways, as well as air and marine ports. In each case, investment and implementation decisions had to be made with at least implicit or tacit consideration of the technological and financial viability of the project, as well as the existence of adequate benefit or payback for (public and private) investors.

In the examples cited above, the intended motivations of parties who built the transport facilities and initiated the transport services span a wide range. They include: enabling or strengthening national defence, new forms of trade among markets, access to jobs, access to recreation opportunities, adoption of new products and distribution technologies, shifts in urban land development patterns, time and cost savings, reliability, quality of life (liveability) and inward flows of investment and wealth, as well as reducing noise and air pollution, and improving health and safety. These goals encompass both direct benefits to users and wider effects on environments and economic development. In other words, interest in non-user benefits – now referred to as "wider" benefits – is longstanding and not just an attempt by countries with mature transport systems to justify projects that fail user benefit cost tests.

Formal consideration of economic development benefits in the US can be traced back to massive public sector investments in water (canal, dam and irrigation) projects in the 1800s. The Erie Canal (completed 1825) is considered a quintessential example of a transport infrastructure investment that was designed as a strategic investment intended to generate wider economic benefits. It had the direct effect of enabling larger canal barges to replace smaller horse-drawn carts as a means of transporting grain from the nation's interior to coastal population centres. As a result, the cost of wheat in New York City dropped twentyfold. The direct effect on cost savings was due to reduced travel time and transport cost per vehicle, and increased scale economies from larger vehicles (North, 1961).

However, the direct cost savings for pre-existing urban markets was dwarfed by a far larger secondary effect, which was to open up the broad Ohio River Valley to population and business growth that was previously not economically feasible or sustainable. The result was a large shift in investment and economic growth into the region, also enabling further increases in exports and national economic growth. It fostered development of new agricultural production technologies that could serve the broader markets. Thus, the Erie Canal had important direct effects on travel efficiency, but also wider secondary effects on productivity, technology adoption and spatial and economic sector growth patterns. These broader consequences are similar to economic development and growth effects that occurred in ancient times as new trade routes developed, and in later times as rail, highway and air routes developed.

A recent study looked at the motivations for US highway capacity projects in the last three decades, and found that these same motivations are still claimed today (EDR Group et al., 2012). Of particular note is the inclusion of factors representing connectivity and accessibility to markets as a factor enabling future economic development opportunities, as opposed to cost savings that affect existing travellers. These findings are presented in a more complete manner in a discussion of *ex-post* case studies later in this chapter.

Evolution of project evaluation (appraisal) methods

There has been a significant evolution over time in the way that economic development impacts have been considered in the *ex-ante* evaluation of infrastructure projects in the US. Following the Erie Canal, claims of wider secondary benefits on regional economic development were used to justify a continuing set of publicly funded waterway and dam projects in the US over the next century. However, when the

Great Depression hit in the 1930s, there was a clamouring for more accountability in public spending. The application of benefit-cost concepts was first required in the US in 1936, in a law that specified that federally funded water infrastructure projects should be undertaken if "the benefits to whomever they may accrue are in excess of the estimated costs" (US Flood Control Act of 1936, Section I). Initially, both direct savings to users and wider, secondary benefits that were induced by the project were counted (Hufschmidt, 2000).

By the 1960s, the field of welfare economics had become more developed. Consequently, narrower standards that focused on welfare economic principles were adopted and extended to transport investment. This perspective is reflected in the original US Red Book on road user benefits (AASHO, 1960, later updated several times with the latest being AASHTO, 2010). It is also reflected in UK applications dating back to the 1960s.

While those early guides ignored the wider, secondary effects, there have since been continuing efforts to make CBA more comprehensive by incorporating more explicit consideration of secondary environmental, economic and social benefits. By the 1990s, emission rates and unit valuation factors had been sufficiently established to enable environmental benefits to be included in transport CBA.

Efforts to also add wider economic development impacts emerged in the 1990s, driven by concerns among rural states that the traditional CBA methods favoured investment in speeding up high volume urban roads but did not provide a way to value rural market connectivity investments. The Wisconsin Department Of Transportation (DOT) case (Weisbrod and Beckwith, 1992) received significant attention as a first effort to take on this issue. The DOT wanted to justify a 293-km highway linking Green Bay to Minneapolis. The project was seen as capable of spurring regional economic growth in a northern region, particularly in economic sectors related to food product packaging and tourism, but it failed traditional CBA tests. The resulting study developed a benefit-cost ratio that substituted an estimate of GDP income growth (which accounted for enhanced market access) in place of the traditional business time and cost savings, enabling the project to pass the test and move forward. Indiana, Montana and other state DOTs followed by adopting revised CBA methods that used the same basic concept, but incorporated more sophisticated modelling of regional GDP impacts (Kaliski et al., 1999; Wornum, 2005). This same approach had also been applied in Scotland, in a study for the M74 motorway extension (Oscar Faber/TPA, 1993). Later, more refined approaches for incorporating GDP effects into CBA were developed in the UK (Dept. for Transport, 2005).

Subsequently, concerted efforts were made to extend CBA methods to include wider productivity effects. In the UK, these efforts focused largely on incorporating urban agglomeration and labour force effects. In the US, there was more focus on freight logistics and supply chain connectivity effects that affected technology adoption as well as scale economies. (Shirley and Winston, 2004; ICF and HLB, 2004).

An even more important change occurring over the 2005–15 period has been a move by many State DOTs to adopt formal scoring systems based on multi-criteria analysis (MCA). That approach has been implemented in ways that make it possible to combine traditional user benefit measures with macroeconomic impacts and a wider set a strategic and social goal achievement measures. Typical MCA factors considered in State DOT prioritisation process are shown (for five example states) in Table 4.1. These factors generally fall into four categories: (1) travel-related benefits, which are typically estimated by transport models, (2) strategic goal related measures, (3) public policy (social goal) related measures, and (4) regional economic impacts, which are typically estimated with economic models. In most states, findings on user benefits and wider benefits became considered in a broader MCA scoring framework for prioritisation decisions.

Looking more closely at the set of strategic factors, it can be seen that many of them directly relate to supply chain or market accessibility, or connectivity to broader opportunities including major economic corridors, supply chains, international gateways and/or intermodal terminals. Reducing “bottlenecks” and increasing “reliability” are often also distinguished because they can affect supply chain productivity and technology adoption (related to loading and stocking inventory) – which are effects beyond what is counted in generalised congestion and reliability effects on travel-time savings. Some states count these factors apart from user benefit and regional economic impact factors, though other states measure and apply them as inputs to a regional economic impact model.

The MCA calculations are carried out by staff of the State DOTs, who typically apply both travel demand and regional economic models to generate many of the factor metrics that then go into the scoring calculation. The weights assigned to individual factors vary from state to state, but in general they tend to be derived from a formal public input process, survey process or expert panel, and are then approved by the state legislature. The number of projects now being rated in these ways ranges from hundreds at a time (in the case of Kansas and Ohio) to several thousand at a time (in the case of North Carolina).

Table 4.1. **Multi-criteria rating factors used for prioritisation**

Rating factor	CO	OH	NC	MO	WI	KS	UK
Traveller benefit and environment (quantitative)							
Efficiency: Travel time, cost, level of service	X	X	X	X	X	X	X
Safety (accident rate)	X	X	X	-	X	X	X
Pollution: emissions/greenhouse gases	X	X	-	X	X	-	X
Strategic (system productivity) benefit							
Intermodal facilities, access and interchange	-	X	*(a)*	X	*(a)*	*(a)*	X
Reduce localised congestion bottlenecks	X	X	X	X	X	X	*(b)*
Connectivity to key corridors, global gateways	-	-	*(a)*	X	X	*(a)*	-
Reliability of travel times	X	X	*(a)*	-	*(a)*	*(a)*	X
Truck freight route, supply chain impact	-	-	X	X	*(a)*	X	-
Social goal achievement (qualitative)							
Location: area revitalisation/regeneration	-	X	-	X	-	-	X
Land use: supports cluster or in-fill development	X	X	-	X	-	-	X
Economic policy: support target industry growth	X	-	-	X	-	-	-
Leveraging private investment	-	X	-	-	-	-	-
Local public support	X	X	X	-	X	X	-
Macroeconomic outcomes (modelled)							
Economic productivity calculation	X	*(a)*	*(a)*	-	*(a)*	*(a)*	X
Job growth, reduced unemployment	X	X	X	-	X	-	-
Gross regional product	X	X	X	-	-	X	*(a)*

Notes: (UK Appraisal Table is also included as “UK” for comparison)

X = explicitly included as an element of the rating system
(a) = implicitly allowed via calculation of additional productivity benefit in CBA or macroeconomic impact
(b) = included in travel efficiency benefit shown above
" - " = not part of formal rating system, but may still be considered through other elements of the decision process
CO=Colorado, OH=Ohio, NC=North Carolina, MO=Missouri, WI = Wisconsin, KS=Kansas

Source: Weisbrod and Simmonds, 2011, with further update by the author.

Since the state DOTs have fixed annual budgets for transport investment, they apply these scoring systems to both rank projects and select projects to be implemented. (However, subsequent steps of alternatives analysis and funding processes still depend to various degrees on consideration of CBA, financial analysis and funding program eligibility.) It is notable that the UK appraisal guidance also incorporates an "Appraisal Summary Table" (AST) that covers a similarly wide range of strategic policy factors that fall outside of CBA, and it provides a means for them to be considered in decision making as part of a larger "Business Case" (see last column of Table 4.1.). However, the AST does not incorporate that same kind of prescriptive measurement definitions and assigned weights that exist in the MCA scoring systems adopted by many US states.

At this juncture, the point to be drawn is that many of the US states have adopted MCA rating systems as a way to combine strategic economic development and social policy considerations alongside travel efficiency (user benefit) considerations in their transport infrastructure decision making. While there are advantages and disadvantages of relying on MCA rating systems in this way, and there clearly are alternative ways of informing decision making that are used in other countries, this chapter does not pursue that topic. (Readers are referred to Worsley and Mackie, 2015, for a discussion of issues regarding the balancing of CBA with strategic and financial considerations in the UK context). Instead, this chapter focuses on the observation that reliance on multi-criteria ratings and consideration of economic development impacts expands the number of transport project factors that need to be measured. That, in turn, increases the need for observational data and research regarding: (a) the measurement of accessibility, connectivity and related productivity factors, and (b) our understanding of how they lead to wider impacts on job and income creation. That has been a major impetus for *ex-post* case studies and decision processes that incorporate local factors, which are discussed later in this chapter.

Why the move toward MCA in the US

There appear to be several plausible explanations for the broad adoption of MCA-based ratings by the State DOTs in the US. The decision process of State DOTs focuses on selecting projects based on the use of prioritisation processes, which are applied to fixed capital investment budgets. The use of scoring systems for prioritisation enables the State DOTs to consider efficiency alongside many of the very same non-efficiency factors that they are required to consider in their environmental impact reports. This includes not only economic development and environmental impacts, but also equity and cumulative impacts on achievement of public policy goals (ICF, 2009).

Another likely reason for the movement to MCA-based ratings is that they provide a way for states to give priority to projects that are of strategic, long-term importance for providing sustainability and future economic development for states. Specifically, many of the states have long-range plans that recognise "strategic economic corridors," emerging "technology clusters," "labour market access" and "export gateways" and environmental sensitivity as issues deserving of special attention. The MCA ratings provide a means of giving weight to investments that support those goals and promote inward investment. Many states also make use of regional economic models to aid in evaluating the economic impacts of projects, and use the results to drive elements of their multi-criteria ratings.

In contrast, there tends to be greater reliance on formula-driven CBA appraisal processes in the UK and Scandinavian nations – where there are more centralised transport budgeting and selection processes. It can be postulated that this is a natural consequence of differences in the level of government that is responsible for decisions. After all, central governments typically face limitations on the ability to modify their decision processes to recognise regional differences in the values and priorities of residents. This makes it more logical for them to rely on CBA formulas that feature fixed elasticity and mark-up factors, and allow for local factors to be considered separately.

Further support for this interpretation about central governments comes from the experience of the US Dept. of Transportation (US DOT) and its "TIGER" program. Historically, the US DOT did not make project selection decisions. However, when Congress responded to the Great Recession by setting up a grant program to increase spending in the economy, US DOT established a process for judging applications. The resulting process relied on traditional CBA as the major screening criterion, with a separate and more qualitative process for considering local factors in a way that generally paralleled the UK process (US DOT, 2014).

Development of evidence-based planning and analysis methods in the US

Evolution of *ex-post* case studies of economic impacts

"Evidence-based" planning and prioritisation processes rely on evidence from previously observed cases, together with consideration of specific aspects of the proposed project, its setting and local values. The prioritisation ratings of the State DOTs generally encompass these very types of considerations, and thus depend on having a base of supporting evidence concerning their importance and impact, which comes from *ex-post* analysis studies.

Ex-post analysis studies have been developed in the US for two main reasons: (1) to document how government programs have led to goal achievement, thus demonstrating the value of program funding for politicians, and (2) to draw lessons from past cases that enable local and state transport planning staff to make better planning and benefit estimation for future projects, and then communicate them more effectively to broader public audiences. Those two motivations created a focus on economic development outcomes rather than transport outcomes.

In 1991, the US General Accounting Office (GAO) issued guidance on *ex-post* program evaluation of federally funded programs. It called for establishing economic impact metrics tied to program goals, use of *ex-post* comparison with matched pairs or statistical controls to account for exogenous changes over time, and suitable effort to attribute credit for observed changes (US GAO, 1991). This guidance was then used when the Appalachian Regional Commission (ARC) – a collaboration of 13 States and the federal government – funded case studies of the economic impacts of 300 local public works (roadway and water/sewer) projects. These projects were completed between 2000 and 2010 to support economic growth in economically depressed communities. The studies measured impacts on job and income growth. Over the 1990s, *ex-post* studies of local economic impact were also conducted by seven different State DOTs covering over 50 community bypass roads (Fitzroy and Weisbrod, 2014).

In 2001, US DOT issued a guidebook that set forth standards for documenting the actual *ex-post* economic effects of highway investments (EDR Group and Cambridge Systematics, 2001). It offered prototype designs for studies of the economic development impacts of highways at regional, corridor and local levels. In the next four years, US DOT sponsored a series of *ex-post* case studies of the economic impacts of major new rural highway projects around the country (e.g., FHWA, 2005).

In 2008, effort was started by the Strategic Highway Research Program, operating under the auspices of the Transportation Research Board, to assemble a national database of *ex-post* case studies concerning the economic development impacts of transport projects. With USD 2.5 million of funding, the TPICS (Transportation Project Impact Case Studies) database was developed covering highway and intermodal terminal case studies (www.tpics.us). Expansion of the system to cover public transport projects is also underway.

Cases in this database are required to include (a) project context and objectives, (b) both pre- and post-project economic measures, (c) inclusion of a counter-factual reference (such as surrounding area or state-wide average changes during the same time period), and (d) attribution of relative credit for observed changes that can be assigned to the transport project. Training materials were also developed for conducting new case studies and an analysis study was conducted to evaluate results of the first 100 case studies. The database has since continued to grow and it has been turned over AASHTO, the association of State DOTs, for further development. It is now being rebranded as the *EconWorks* Case Study database (https://planningtools.transportation.org/13/econworks.html).

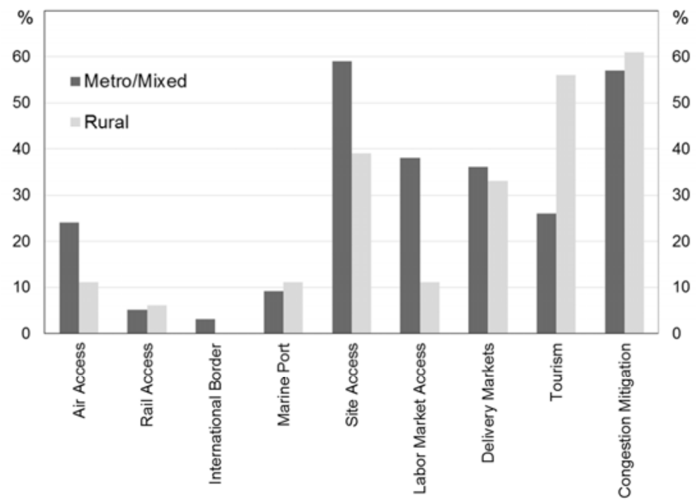
Under the AASHTO umbrella, State DOTs are now making use of the *ex-post* database to assess its usefulness for extracting planning process lessons and insights; its transferability for identifying the range of likely impacts of proposed new projects in very early stage evaluation of proposed new projects; and its applicability to further improve economic impact forecasting models and methods.

Findings from *ex-post* case studies

The second of our three evaluation attributes is *practicality.* For evaluation or appraisal processes to be of practical use in project prioritisation and selection, they must be capable of discerning differences in potential impacts and benefits of competing alternatives that are due to variations in type of projects, type of settings and resulting classes of benefit. Such differences can indeed be observed from *ex-post* case studies. In general, the following types of findings have emerged concerning the incidence and rate of project impacts on inward investment, employment and income growth. The findings summarised below were all drawn from empirical analysis of the TPICS database (EDR Group, 2012), with some additional examples and illustrations drawn from other studies cited below.

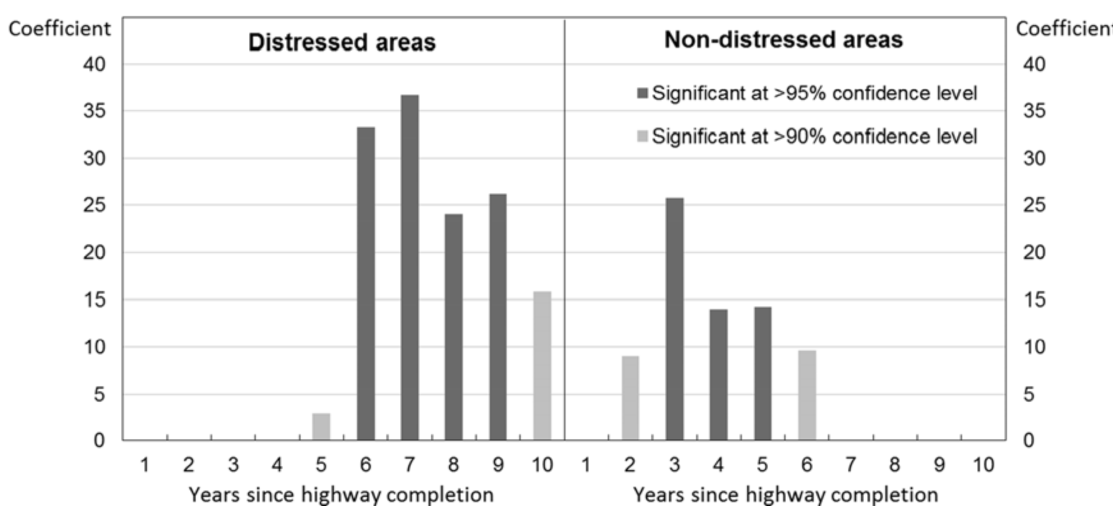
- Type of project matters. Some projects are built to enable and generate wider economic development impacts, while others are constructed to address safety deficiencies, environmental concerns, functionally obsolete designs, or facility maintenance and rehabilitation. In general, only projects that are intended to enhance improve user costs, market access or locational connectivity can be expected to enable wider economic growth (in terms of jobs and income). There is no point in wasting resources to look for ex-post evidence of wider economic development impacts from the other types of transport projects, nor should they be expected in *ex-ante* benefit forecasts for those projects.
- Benefits are not necessarily just for existing travellers. Over half of all highway system capacity projects have a goal of enhancing future accessibility to labour and buyer/supplier markets, or connectivity to intermodal terminals. These projects are effectively supporting the growth of future economic activities (job and income growth opportunities) rather than just generating savings for current travel activities. For that reason, *ex-ante* appraisal of project benefits must recognise the role of expanding market access (as well improving intermodal connectivity and reliability) as factors enabling productivity and inward investment gains (Figure 4.3).
- Time periods of impact can vary substantially. Wider economic development impacts can take a decade or longer to occur. The pace of impact occurrence also depends on the local setting; it often takes longer in economically distressed areas (Figure 4.4). Yet it is in those areas that the impacts may be most desired and needed. *Ex-ante* prediction of agglomeration benefits and wider GDP impacts should recognise this delay aspect of impacts.

Figure 4.3. **Motivations for highway investments**



Source: Economic Development Research Group, 2012.

Figure 4.4. **Time lag in economic growth effects following highway investments**



Note: Study covered highway investments in the 13-state Appalachian region. The vertical scale represents standardised regression coefficient values.

Source: Cambridge Systematics et al., 2008 (pp. B-17 and B-18).

- Settings and local conditions also affect observed economic development impacts. While only projects expected to have economic development impacts were studied, about 15% had no net impact or a small negative impact on the area economy. Lack of benefit was often related to a deficient business climate, as represented by a lack of supportive local regulation (zoning), utility infrastructure and financial support policies. This occurred more often in rural areas. Ideally, *ex-ante* forecasts of expected economic development impacts should also be capable of adjustment to allow for local support factors.

- Concentration of beneficiaries matters. It is particularly difficult to observe wider economic development impacts for projects that reduce bottlenecks (choke points). While their time and cost savings may be particularly distinct, their market access benefits will tend to be dispersed and diluted. For that reason, care should be taken in ascribing economic benefits associated with enhanced market access unless there is an identifiable business location area for which access is clearly expanded up by such improvements.

- Incidence of local business cluster effects, as well as the conditions enabling them, can be observed. Case studies have documented impacts of new highway and transit projects on several distinct types of clusters: (1) supply chain clusters that extend along highway corridors, (2) centralised logistics clusters that locate where major long distance routes intersect, (3) software and emerging technology clusters that locate at transit-served areas of major cities that have research-focused universities, (4) industrial clusters that locate near major intermodal freight (air, rail) terminals, and (5) banking, finance and corporate headquarter clusters that locate in large markets with good international air services and usually good transit service. They share a common feature which is a dependence on both market scale and system connectivity factors. Examples are shown in Box 1 below. It is notable that these examples highlight the same specialised access, connectivity and reliability elements that were listed in Table 4.1 within the group of strategic, productivity-related rating factors.

These *ex-post* case studies of transport-driven business clusters do have some common features: the clusters are spatially distinct, they are all highly specialised in terms of the type of business located at them, and they all feature strong connectivity to wider markets (labour markets, freight delivery markets, or intermodal transport facilities depending on the type of cluster). In each case, it may also be observed that the cluster developed for a specific sector of the economy because transport improvements enabled the use of newly emerging technologies, such as just-in-time delivery, centralised warehousing and collaborative processes for software development. In other words, they built on improvements in market scale and location connectivity, but these types of clusters were also highly specialised, industry specific, and not widely seen elsewhere. The implication is that localisation benefits can be observed and measured, but they should not be assumed to necessarily be widely applicable for other industries or areas. The body of *ex-post* case studies provides a basis for economic development impact models that estimate major changes in income growth only when certain combinations of factors come together.

There are no examples of retail or commercial clusters listed here. Since these projects represent shifts in where local residents spend their money rather than sources of new income coming into the state, the regional economic impact models generally show that they do not generate any net income growth at a state level. They do not affect any of the strategic factors nor the economic impact factors used in the MCA rating systems of State DOTs (Table 4.1), so they get priority ratings only if they support public policy goals of revitalising economically depressed areas or generating in-fill development.

A conclusion to be drawn from the *ex-post* case study literature is that market access impacts are about much more than just scale economies. More fundamentally, they are about enabling new forms of economic activity to occur, new technologies to be implemented, and strategic policy goals to be achieved.

Box 4.1. **Examples of clusters profiled in *ex-post* case studies**

An *automobile supply chain cluster* that developed along a 150-mile (200-km) stretch of I-65 and I-75 in rural Kentucky and Tennessee. The cluster's spatial boundaries and corridor location pattern reflect the limits of same-day truck deliveries, "just in time" supply chain technology, and "lean production" processes. The rural location optimises access to multiple urban markets and minimises locally congested traffic sites. The spatially dispersed and largely rural siting of industries within the cluster reflects reliance on low labour cost areas (Rosenfeld, 2000; Vadali and Chandra, 2014).

A *centralised distribution warehousing cluster* located at two highway intersections along I-81 in the Scranton and Harrisburg areas of Pennsylvania. This highway developed as a north-south truck route that avoids the far more congested I-95 corridor. The two intersection sites provide east-west access for more reliable same-day delivery from very large, centralised warehouses to simultaneously serve four large metro areas: Baltimore, Philadelphia, New York and Washington DC. Centralised warehousing is a relatively new process that uses information technology and scale economies to enable high turnover stocking strategies to maximise productivity (FHWA, 2012; TPICS, 2013).

Computer software clusters located along rail transit routes in former warehouse districts in San Francisco and Boston. Both clusters are in cities that have very large, diverse and highly educated labour markets. They are located near transit stations, feature access to urban amenities, have direct transit access to major universities with R&D technology programs, and also have transit access to an international airport. (EDR Group, 2013b)

An *intermodal "logistics park" cluster* in Will County, Illinois. This industrial park at the south of the Chicago metropolitan area provides Midwest manufacturers with highway access (via I-80 and I-55, and Chicago's outer beltway I-355). It is adjacent to an intermodal rail terminal where products can be transferred from trucks to freight trains going directly to the west coast ports of Los Angeles, Long Beach, and Seattle-Tacoma, serving Asian trade routes. It is used to import electronic and retail goods into the region from Asia, and export regional agricultural products to Asia (TPICS, 2013).

Emergence of evidence on threshold factors

The third of our three evaluation attributes is *accuracy*. To maximise accuracy, an *ex-ante* evaluation should be capable of distinguishing projects that will lead to observable impacts and address critical local needs and deficiencies, from other projects that will have less dramatic and more diluted benefits. In that respect, *evidence-based analysis* can have significant advantages over theoretically driven *ex-ante* models when it comes to planning and decision processes that require estimates of anticipated project benefits and impacts. The reason is that the evidenced-based case study analysis can reflect both interaction effects and catalytic effects which are tied to benefit thresholds. This is in contrast to theoretical models that most often apply constant elasticity, coefficient and mark-up factors, and assume constant trade-offs among independent cost and benefit elements.

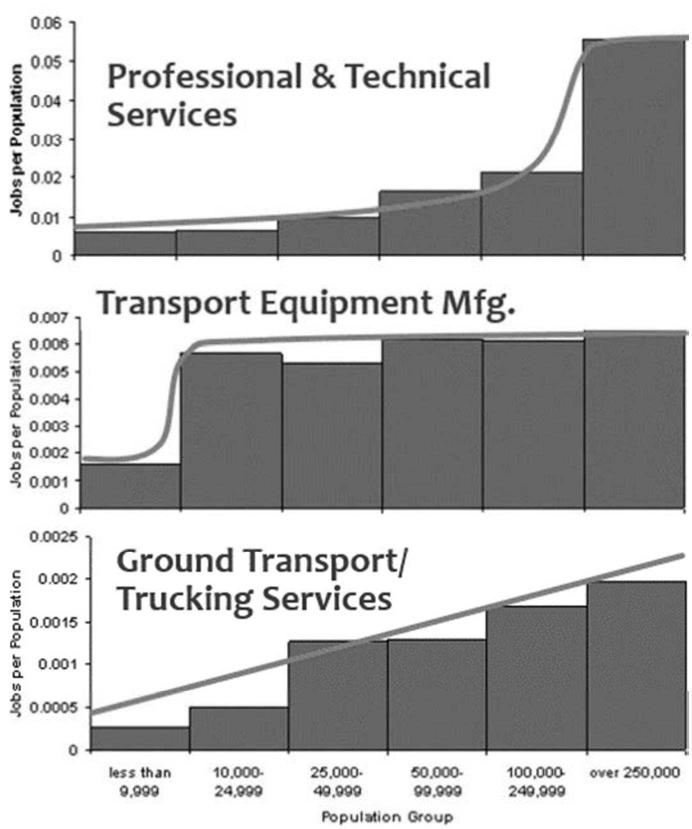
In fact, the use of MCA ratings increases the need for efforts to distinguish the components of productivity, which can include effects of improving reliability along with access to wider labour markets, to wider customer markets, and to intermodal terminals that are windows to even larger markets. This has helped fuel the need for studies that examine these individual elements of productivity. Many of them were examined in a recent literature review conducted for the NCHRP study of productivity impacts (EDR Group et al., 2013a).

Five types of threshold effects have been identified and measured as a result of *ex-post* evaluations and other statistical studies. They are: (a) labour market scale effects, (b) commuting time thresholds, (c) labour force participation rate thresholds, (d) regional truck delivery thresholds and (e) intermodal access time thresholds. Findings on these five types of thresholds are summarised below.

Labour market size thresholds

One of the findings from a series of studies sponsored by the Appalachian Regional Commission is that the concentration of specific industries in a labour market area will differ depending on the size of that overall labour market (EDR Group et al, 2007). This labour market size effect is independent of transport conditions within the area. As shown in Figure 4.5, some industries (such as transport equipment manufacturing) that have relatively modest worker training and education needs can exist in all but the smallest labour market (i.e., those with at least 10 000 workers), and show no further gain from larger labour markets. At the other extreme, professional and technical services, which require more specialised skills, tend to gain productivity and concentration in very large labour markets (i.e., over 250 000 workers) where there is a sufficient size of customer base and a higher likelihood of finding workers with the required matching skills. Others such as transport services continue to increase in concentration as labour market size grows, which suggests continuing scale economies of operation but no specialised worker skill or customer requirements that would require a minimum size labour market (Figure 4.5).

Figure 4.5. **Relative concentration of industries by size of labour market**



Source: EDR Group et al., 2007.

Table 4.2. **Sensitivity of industries to access measures**

		Sensitivity to Access Measure (1-10 scale)			
NAICS	**Sector Description**	40-min Market	3-hr Delivery Market	Commercial Airport	Rail Intermodal
111	Crop Production	3	5	0	3
112	Animal Production	0	5	0	3
113	Forestry and Logging	5	0	2	0
114	Fishing, Hunting and Trapping	0	3	0	3
115	Support for Agriculture and Forestry	3	0	0	0
211	Oil and Gas Extraction	0	0	0	0
212-213	Mining and Support Activities	3	0	4	5
221	Utilities	5	0	3	5
230	Construction	8	5	7	8
311	Food Products	3	0	0	0
312	Beverage and Tobacco Products	10	0	0	3
313	Textile Mills	5	5	2	3
314	Textile Product Mills	5	10	0	0
315	Apparel Manufacturing	5	5	0	0
316	Leather and Allied Products	5	3	2	5
321	Wood Products	0	5	0	5
322	Paper Manufacturing	0	5	0	5
323	Printing and Related Activities	10	10	7	0
324	Petroleum and Coal Products	6	0	0	0
325	Chemical Manufacturing	5	3	4	3
326	Plastics and Rubber Products	8	10	0	3
327	Non-metallic Mineral Products	5	5	2	0
331	Primary Metal Manufacturing	3	5	4	0
332	Fabricated Metal Products	10	5	2	0
333	Machinery Manufacturing	0	5	2	0
334	Computer and Electronic Products	3	5	2	3
335	Elec Equipment, Appliances	0	10	3	0
336	Transportation Equipment	5	5	3	3
337	Furniture and Related Products	5	10	3	0
339	Miscellaneous Manufacturing	5	5	3	0
420	Wholesale Trade	10	0	3	0
441-454	Retail Trade	8	3	3	5
481-487	Transportation	5	0	3	0
491-493	Mail, package delivery and warehousing	10	0	2	3
511	Publishing Industries (except Internet)	10	0	10	0
512	Motion Picture and Sound Recording	10	3	9	0
513	Broadcasting	10	0	5	0
514	Internet and data process svcs	8	3	5	0
521-523	Monetary, Financial, and Credit Activity	10	0	3	0
524	Insurance Carriers	10	3	5	0
525	Funds, Trusts, Financial Vehicles	5	5	5	0
531	Real Estate	10	0	7	0
532	Rental and Leasing Services	10	0	5	0
541-551	Prof. Scientific, Technical, Services	10	3	10	0
561	Admin and Support Services	5	0	10	0
562	Waste Mgmt and Remediation	3	5	3	0
611	Educational Services	10	5	3	0
621-624	Health Care and Social Services	8	0	0	0
711-713	Recreation and Amusements	5	0	10	0
721-722	Accommodations, Eating and Drinking	5	0	7	0
811-812	Repair, Maint. and Personal Services	5	0	7	0

Source: Alstadt et al., 2012.

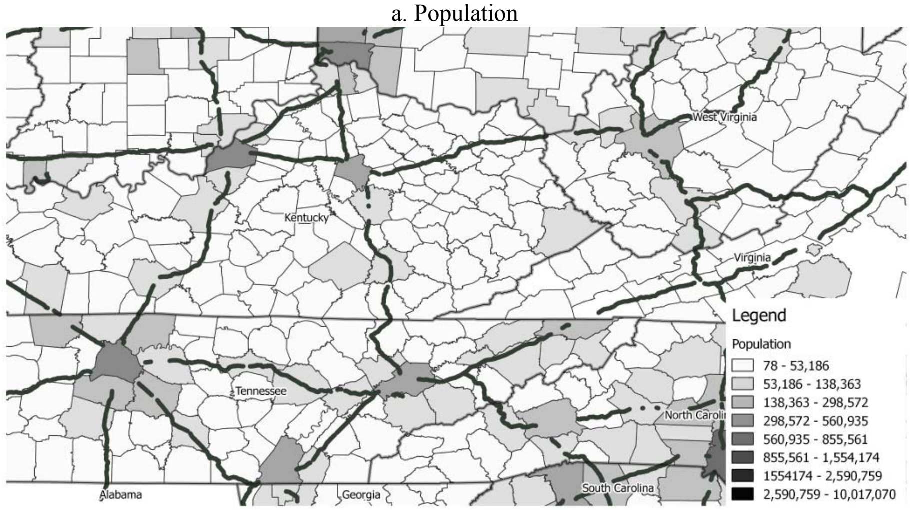
Regional truck delivery market size thresholds.

Truck deliveries are naturally subject to threshold effects related to both regulations on daily driving hours and business hours of operation. The area that can be served with same day product and service deliveries is roughly three hours from the place of origin. This is based on a window of eight hours of operation, with an allowance of three hours for each of the outbound and inbound trips, plus one hour at each end for pickup and delivery. Statistical studies have confirmed that manufacturing industries tend to locate where they can maximise the size of same-day truck delivery markets rather than locating where labour market access is maximised. This reflects an optimisation of product buyer-supplier supply chains. The result is that both worker compensation and business concentration levels for manufacturing firms tend to rise with greater three-hour truck delivery markets – a clear indication of a threshold and a productivity effect.

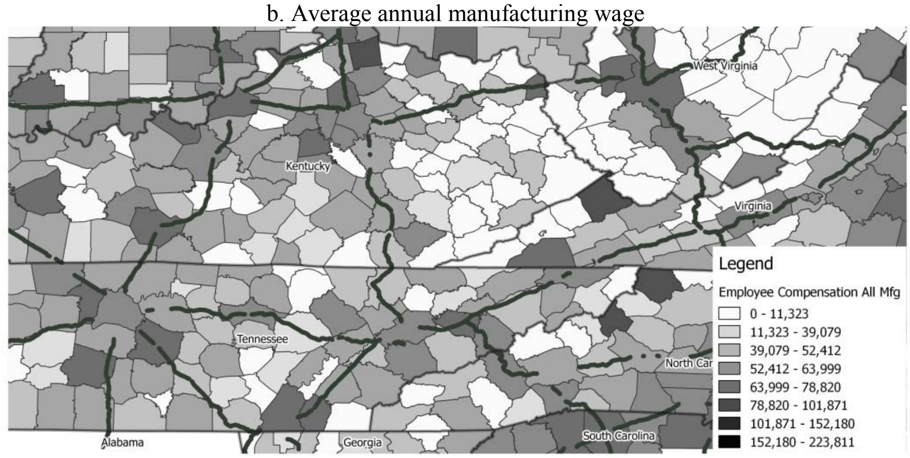
This finding is demonstrated by statistical analysis of the relationship between concentration of industries in a county and various measures of market access and intermodal connectivity from the population centre of that county (Table 4.2). Higher numbers and darker shading denote a stronger relationship; the three-hour delivery market is shown to be important primarily for manufacturing industries (Alstadt et al., 2012). Figure 4.6 illustrates the wage effect by showing the area of where automobile parts suppliers are clustered, and it can be seen that manufacturing wages are high not just where the population centres are located, but also along major highway corridors between those centres.

Figure 4.6. **Population concentration and manufacturing wage rates among counties in central Appalachia**

a. Population



b. Average annual manufacturing wage

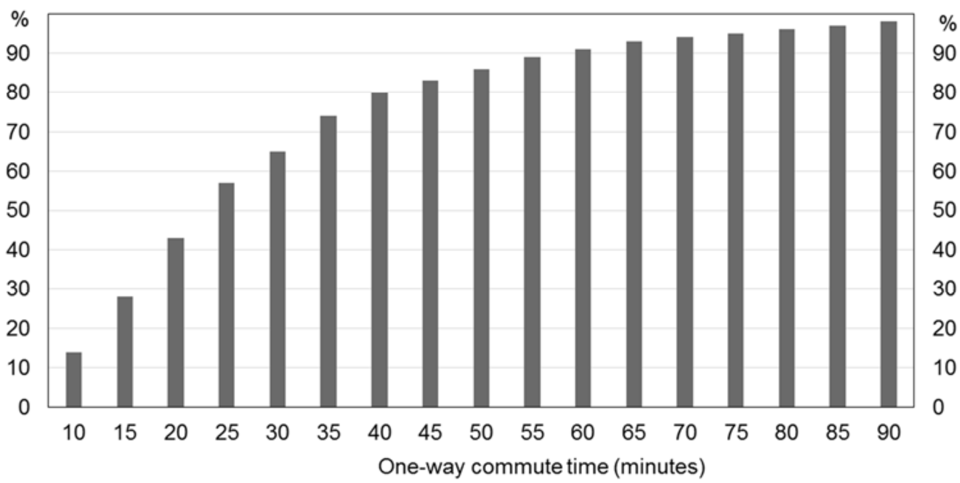


Source: Analysis by the author, based US Census and Bureau of Labour Statistics, 2010-2013.

Commuting time thresholds

While travel time from home to work can vary widely within a labour market, there are travel-time thresholds that reflect worker preferences to avoid very long commutes. Most planning studies assume that this threshold value is a somewhere between 30 and 45 minutes. This interpretation is supported by the American Community Survey, which indicates that two-thirds of all commutes in the US are less than 30 minutes, 80% are less than 40 minutes and 90% are less than 55 minutes (Figure 4.7). However, in small- to medium-size communities in the US, the entire metropolitan labour market area will tend to have driving times within 40 minutes, so the economic impact of reducing commuting times becomes of significant importance only for the larger metropolitan areas.

Figure 4.7. **Distribution of commuting time** (cumulative per cent)



Source: American Community Survey, 2009.

Labour force participation rate thresholds

It has been theorised that transport improvements can enhance labour market participation by enticing more workers into the labour market. However, there is little observational evidence in the US of these effects actually occurring in ex-post studies. The only exception is in rural areas where there is high unemployment, and in those areas there is some evidence that labour force participation goes up when additional employment growth occurs (Bradley, 2000; EDR Group, 2007).

Intermodal terminal access thresholds

For industries that depend on worker or customer travel to/from broader external markets, there is a premium value in having access to major airports. For other industries that depend on freight deliveries for incoming delivery of parts and outgoing shipments of finished products, there is a premium value in having access to major seaports, intermodal rail terminals and/or air cargo terminals. For instance, a statistical study of the relative concentration of industries among US counties showed that low travel time to a major airport is a major determinant of business location for two sets of industries: (a) tourism and conference serving sectors (including recreation, lodging and restaurant), and (b) finance, professional and technical services, which have high rates of worker business travel (see Alstadt et al., 2012; and Table 4.2 above).

Use of case study and empirical analysis findings

Adaptation of *ex-post* empirical findings to inform economic impact forecasting models

The findings from case studies and other empirical studies of productivity elements indicates that actual impacts can sometimes be larger and sometimes smaller than would be predicted by applying constant model factors based on national averages. Project characteristics – such as the type of access, connectivity or reliability change – will matter to specific types of industries. Local factors – particularly those relating to industry mix, and project setting – will help determine which industries are affected. Together, project characteristics and local factors interact, causing some projects to have significantly higher or lower impacts than would otherwise be expected. Regional economic analysis methods can incorporate these considerations in predictions of impacts on investment, job and income growth.

The two regional economic model systems that are widely used by State DOTs for prioritisation and major project evaluation have been updated in recent years to incorporate these very factors. Both REMI TranSight and TREDIS now incorporate "economic geography" concepts by featuring separate inputs for changes in local *labour/commute* market access and regional *freight delivery* market access, as well as traditional travel time and travel cost savings effects. (The latter model also includes inputs for changes in access to intermodal terminals.) *Logistics* impacts are a major concern to many State DOTs, and these models also consider how delivery reliability also affects business competitiveness. The inputs to both systems include project-induced changes to both travel characteristics and accessibility characteristics, and both systems then consider how different industries value and respond to those changes - Table 4.3 shows the transport input variables for one such system. As a consequence, characteristics of the local economy play a major role in model predictions of economic impact. Since these systems have multi-regional CGE-type formulations, they also estimate changes in domestic and international export flows, inward investment flows, labour supply/demand and wage rates over time. Both also show fiscal (government revenue) impacts.

Table 4.3. **List of transport changes that are economic model inputs**

Generalised cost factors by mode (car, truck, transit, rail, air, marine, bike), trip purpose and time period	**Accessibility characteristics** by mode (car, truck, transit, rail) and time period
• Trips • In-vehicle travel time, vehicle-hours of travel • Wait/schedule delay time • Out-of-vehicle travel time • Vehicle-miles of travel (VMT) • Congestion (percentage congested VMT or buffer time) • Vehicle occupancy • Fare/fee/toll – per person, per vehicle or per mile or km	• Local* market for labour commute to work (car, transit only) • Local* market for goods and services delivery (truck only) • Regional* market for same day passenger trips (car/bus and passenger rail only) • Regional* market for same day freight delivery (truck only) • Long distance access: time to cargo airport (truck only) • Long distance access: time to air passenger terminal (car and transit) • Long distance access: time to intermodal rail freight facility (truck only) • Long distance access: time to passenger train station (car and transit only) • Long distance access: time to marine cargo port (truck and freight rail)

Note: * Local market is typically within a 40-50 minute access time; regional market is for same day return trip (typically within three-hours access each way).

Source: http://www.tredis.com/resources/tech-docs

A decade ago, this class of regional economic impact models (that feature spatial accessibility and business reliability factors) was not available to State DOTs. The development of these models, and their use to inform prioritisation and evaluation by State DOTs, is a direct result of evidence-based analysis studies that focused on understanding behavioural factors affecting business activity growth. Besides providing potentially better sensitivity to distinguish impacts of competing transport projects, another implication of these models is that they have expanded the definition of GDP impacts to include logistics, supply chain, export growth and inward investment impacts.

Application of economic impact forecasting models to inform *ex-ante* project evaluation

In addition to being used for prioritisation by state DOTs, the regional economic impact models are also being applied in *ex-ante* studies to support "alternatives analysis" for major investment projects. In all such cases, they are used as a supplement to cost-benefit analysis (CBA) and financial impact analysis (FIA). Their main use is to help assess strategic economic development goal achievement, and to show how industry growth impacts would be expected to occur across space, time and elements of the economy. (Wang, 2012, provides an overview of this same issue in the Australian context.)

There are examples of such application across the US, Canada and Australia. Three examples of published alternatives analysis studies that have used EIA in combination with CBA and FIA are shown below. In each case, the market access and connectivity aspects of the project and their wider economic development impacts were estimated via an economic model. Each example is listed below, along with a brief description of its strategic economic development goal and reference for further information.

- The North Beaches BRT proposal in Sydney, Australia, (Transport for NSW) intended to establish a link between the emerging Global Arc high tech area, a new medical centre, and downtown Sydney (Weisbrod, Mulley and Hensher, 2015).
- The King-Main LRT proposal in Hamilton, Canada, (Ontario, Metrolinx) intended to improve the link from McMaster University to downtown Hamilton, and support improved feeder service for the commuter rail link to Toronto (Steer Davies Gleave, 2010).
- The new Ohio River bridge proposal linking Indiana and Kentucky (Indiana Finance Authority and Kentucky Transportation Cabinet) intended to link a new Riverport industrial zone to the UPS "WorldPort" freight hub at Louisville International Airport (Weisbrod and Duncan, 2015).

There are three potential advantages associated with the approach of conducting EIA with CBA and FIA, and doing so in a consistent manner. First, it enables economic, strategic and financial considerations to be considered and presented in a more holistic way. They may be considered together via a formal multiple accounts evaluation or via a policy discussion process. Second, it also can be more satisfying for decision makers, as it meets the three criteria set forth at the beginning of this chapter: relevance in terms of aligning evaluation and objectives, practicality in terms of representing the role of project settings and contexts, and accuracy in terms of portraying impacts relative to needs and deficiencies. Finally, it can also enable consideration of ways that projects can address public goals of achieving reasonable levels of efficiency, equity and "strategic policy" achievement.

The idea of requiring these multiple forms of economic analysis to be done together and in a consistent manner is certainly not new. For instance, the US Environmental Protection Agency's *Guidelines for Preparing Economic Analyses* states that "For most practical applications, therefore, a complete economic analysis is comprised of a CBA, an EIA, and an equity assessment." (EPA, 2010: p. 1-5). It further states that "For any regulation, it is essential to ensure consistency between the EIA and the [CBA]. If a [CBA] is conducted, the corresponding EIA must be conducted within the same set of analytical assumptions." (EPA, 2010: p.9-2)

Implications for benefit accounting and decision support systems

Strengths and weaknesses of *ex-post* analysis and evidence-based decision support methods

The assembly of *ex-post* case studies in the US, and their focus on economic outcomes (rather than travel forecasting accuracy) was originally motivated by a desire to more easily show politicians the value of investing in transport. In addition to helping to generate examples of the positive impacts of funding transport projects, the case studies have also fed public interest in ways that transport investments enable new technology processes, industry clusters and commodity flows (trade routes). That topic has driven interest among some State DOTs to better distinguish between projects that have strategic importance for economic development because they connect certain markets, business activity centres and intermodal gateways. It has motivated improvements in regional economic impact models so that they can better distinguish such investments from the many other transport projects that have less clear-cut economic development consequences.

All of the above-mentioned factors help explain the increasing interest of State DOTs in adopting multi-criteria scoring systems that include factors which (directly or indirectly) relate to productivity, connectivity, reliability and accessibility effects, as well as broader economic development. Of course, one can question the trade-off in effort between generating increasingly complex scoring systems that call for calculation of detailed metrics, and simpler systems that are more straightforward and less demanding but perhaps less capable of distinguishing differences among projects.

The presence of *ex-post* case studies also provides an opportunity for direct use. That raises an obvious concern about transferability of results. The TPICS/EconWorks database of *ex-post* case studies in the US has search and interpolation tools that can be misused. It is not hard for proponents of any project proposal to point to success stories and claim that they demonstrate the value of their new proposal. About the only way to minimise that situation is to also make use of more sophisticated economic impact and decision support models that can account for differences in local settings and contexts.

Measurement issues: Additionality and double counting

A continuing issue in economic impact measurement is additionality, i.e. the extent to which observed economic growth impacts are net growth or merely transfers of activity. Most governments consider economic activity shifts within their jurisdiction as a zero net gain, but inflows of money and investment from outside to be a net gain. Of course, that entire situation becomes complex in the US context because there are three levels of transport funding and decision-making – federal, state and local/metropolitan. Each level may view the same economic development in a different way, and that has in fact spurred interest in distinguishing transport projects that are of local significance, state significance and national significance.

The multi-criteria rating systems and economic impact models that have been discussed in this chapter do attempt to address those distinctions. This can be seen in the MCA scoring systems shown in Table 4.1, which variously include rating points for projects that involve “key corridors,” “global gateways,” “freight routes,” “supply chains,” and “intermodal facilities” – all means of differentiating projects that have broader area economic significance. Further differentiation is made in the calculation of state-level economic growth impacts via economic impact models. The economic models used by State DOTs also usually include capabilities to distinguish shifts among regions within the state and between state and national impacts. The latter is done by considering trip ends. For instance, pass-through traffic is a gain at the national level but may generate little to no benefit or income for state residents. There are requirements that projects with federal funding must be assessed from a national viewpoint, but some states switch to a state-level viewpoint for evaluation of projects that are fully state funded.

The topic of double counting is more difficult to unravel. The fundamental issue is whether a project prioritisation or selection process is being unintentionally biased by some projects being assigned more benefit than is rightfully deserved. In CBA, incorrect benefit-cost ratios arise if the total benefit calculation is upwardly biased by the inclusion of two or more benefit elements that overlap in coverage – i.e., at least partially reflect the same effect. For MCA, however, overlap among input elements may not be problematic if the MCA weighting system is adjusted so that there is no skew in the relative ranking.

The problem that arises is that many of the MCA rating factors, and factors affecting wider economic impacts (in EIA), are theoretically distinct from each other but tend to be significantly correlated in their incidence. This is probably an unavoidable consequence of combining measures of direct user cost savings with measures of accessibility effects and secondary impacts on transport-reliant industries. For instance, consider the case of a section of highway that has a high volume/capacity ratio. The most likely outcome will be slower traffic and hence more delay for travellers (which is a cost factor). But this congestion condition may also reduce travel-time reliability, and late shipments may increase loading dock and stocking costs for freight shippers and receivers (which is a logistics cost). If this situation occurs often, then it will cause businesses to add more buffer time to product and service delivery schedules. It may thus shrink the market area from which deliveries are made from a given location, or the effective density of opportunities that are accessible from it (which is an accessibility factor). And the congested road may also reduce access to supply chain routes, intermodal and international gateways, and other factors that are also sometimes part of MCA ratings. Similar examples of compounding impacts may occur if there is congestion at a rail terminal, airport runway or seaport dock.

Now in theory each of these above-cited impact elements is a distinctly different effect, and one can also construct examples where one of these forms of impact occurs without any of the others. But in practice, they often occur together. As a result, one cannot be certain that the coefficients which were statistically derived to reflect their impacts adequately control for those correlations. This same issue was examined in a US guide to measurement of transport impacts on productivity, and the position that was taken in that report is that correlation does not necessarily translate into double counting (Weisbrod et al., 2014). But if research studies derive valuation or elasticity factors separately for each effect without controlling for other correlated effects, some impacts may be under- or over-estimated.

Another question that arises is whether wider economic impacts are merely a way to generate larger numbers than would otherwise emerge from consideration of user benefits alone. In US practice, this is not the case because the relationship between social benefit metrics in CBA and GDP impacts by economic models in EIA are not closely correlated. In the context of State DOT use, a project that has large time savings for travellers may generate little impact on the State's economy if the traffic is largely pass-through movements. On the other hand, a project may have a dramatic impact on the economy if it affects the competitiveness of the state's export shipments.

Completeness and accuracy of impact elements

Ex-post case studies and associated research on micro-level (small area) impacts do help identify ways to improve project ranking and selection. They include the following:

- Local project details and local context matters. Characteristics of project size and type can interact with characteristics of local settings to affect the size and nature of economic development impacts. Thus, location setting factors can and should be considered in the evaluation of proposed projects.

- Economic development impacts do not automatically occur everywhere. Consequently, state DOT staff tends to accept claims of wider economic benefits in cases where there are specific types of access improvements for specific business activity centres. However, there is more reluctance to embrace studies where there are only non-specific claims of agglomeration benefits.
- Freight and intercity connectivity need to be recognised alongside passenger access effects. After all, transport projects can enable GDP growth not only via market scale economies, but also by enabling new technology adoption and spatial activity shifts that increase net exports. This finding increases the importance of recognising freight logistics and intermodal connectivity effects – a point that is also emphasised by Hoel et al. (2011).
- Thresholds factors exist. There are practical travel-time thresholds affecting the size of labour markets, freight supply chains and intercity business travel markets. These thresholds can affect passenger and freight demand, and the ability of businesses to implement new technologies. The implication is that if threshold factors are recognised, some projects would rise and others would fall in ranking lists.
- Time lags exist. Broader (non-user) economic development impacts occur over time, and may take over a decade to occur depending on the type of project and local setting. The implication is that these time lags should be incorporated into CBA and EIA studies, as otherwise the expected GDP impacts may be overstated.

The multi-criteria rating systems used by many US states are an attempt to measure strategic economic development goal achievement along with the more traditional user benefit measures, so that both can be considered together in decision processes. Consequently, these rating systems are usually designed to be sensitive to the explanatory factors and threshold factors that have emerged from case studies and associated statistical studies – i.e., the bullet items listed above.

Conclusion

Ultimately, we can gain insight into more sophisticated planning and realistic modelling if we consider broader impact factors in decision frameworks. Insight can be gained by examining how the creation of new spatial, temporal and distributional linkages may affect technology adoption and activity patterns.

Finally, there is a need to make use of broader analysis metrics to better communicate economic development impacts to public and private stakeholders. Better communication is enabled when agencies can tell more of a “story” regarding who, when and how wider benefits and other impacts are expected to occur. And to provide that story, more *ex-post* case data is needed and more empirical research is needed that actually pools findings across nations, and among academic and consulting communities.

References

Alstadt, B., G. Weisbrod and D. Cutler (2012), "The relationship of transportation access and connectivity to economic outcomes", *Transportation Research Record*, Vol. 2297, pp. 154-162. www.edrgroup.com/pdf/Alstadt-et-al-TRR-2012.pdf

American Association of State Highway Officials (1960), Road user benefit analyses for highway improvements. http://babel.hathitrust.org/cgi/pt?id=mdp.39015006410040;view=1up;seq=9

American Association of State Highway and Transportation Officials (2010), "User and non-user benefit analysis for highways", NCHRP Project 2-23.

Bernstein, W. (2008), "A splendid exchange: How trade shaped the world", *Atlantic Monthly Press.*

Bradley, D., S. Herzenberg and H. Wial (2001), "An assessment of labour force participation rates and underemployment in Appalachia", Keystone Research Center for the Appalachian Regional Commission. www.arc.gov/assets/research_reports/assessmentlaborforceparticipationratesandunderemployment1.pdf

Cambridge Systematics et al. (2008), "Economic impact study of completing the Appalachian Development Highway System", prepared for the Appalachian Regional Commission. www.arc.gov/assets/research_reports/EconomicImpactStudyofCompletingADHS.pdf

Casson, L. (1994), *Travel in the ancient world*, Johns Hopkins University Press.

Department for Transport (UK) (2005), "Transport, wider economic benefits, and impacts on GDP", Discussion paper. http://webarchive.nationalarchives.gov.uk/+/http:/www.dft.gov.uk/pgr/economics/rdg/webia/webmethodology/sportwidereconomicbenefi3137.pdf

Economic Development Research Group (EDR Group) and Cambridge Systematics, (2001), "Using empirical information to measure the economic impact of highway investments, Volume II: guidelines for data collection and analysis", Federal Highway Administration. http://www.edrgroup.com/pdf/fhwa-hwy-impact-vol-2.pdf

Economic Development Research Group et al. (2007), "Sources of growth in non-metro Appalachia, Volume 3, statistical studies of spatial economic relationships", Appalachian Regional Commission. www.arc.gov/research/researchreportdetails.asp?REPORT_ID=84

Economic Development Research Group (2012), "Interactions between transportation capacity, economic systems and land use", Report S2-C03-RR-1, SHRP2 Program, Transportation Research Board. http://www.trb.org/Main/Blurbs/166934.aspx

Economic Development Research Group et al. (2013a), "Economic productivity & transportation investment: task 1 literature review, stakeholder perspectives and framework outline", NCHRP Project 02-24, Transportation Research Board. http://onlinepubs.trb.org/onlinepubs/nchrp/docs/NCHRP02-24_Task1LitReview.pdf

Economic Development Research Group (2013b), "The role of transit in support of high growth business clusters in the U.S.", American Public Transportation Association.

www.apta.com/resources/reportsandpublications/Documents/TransitHighGrowthClustersUS-Final2013-1124.pdf

EPA: US Environmental Protection Agency (2010), *Guidelines for preparing economic analyses*, updated in 2014. http://yosemite.epa.gov/ee/epa/eerm.nsf/vwAN/EE-0568-50.pdf/$file/EE-0568-50.pdf

FHWA: Federal Highway Administration (2012), "Economic development history of I-81 corridor", US DOT. www.fhwa.dot.gov/planning/economic_development/studies/i81pa.cfm

GAO: General Accounting Office (1991), *Designing evaluations.* Program Evaluation and Methodology Division, GAO/PEMD 10.4.1. Updated in 2010. www.gao.gov/assets/590/588146.pdf

Hoel, L, G. Giuliano and M. Meyer (eds.) (2011), "Intermodal transportation: moving freight in a global economy", Eno Foundation for Transportation.

Hufschmidt, M. (2000), "Benefit-cost analysis: 1933-1985", *Journal of Contemporary Water Research and Education*, Vol. 116. http://opensiuc.lib.siu.edu/cgi/viewcontent.cgi?article=1196&context=jcwre

ICF Consulting and HLB Decision Economics (2004), "FHWA freight BCA study: summary of phase II results", Federal Highway Administration, US DOT. www.ops.fhwa.dot.gov/Freight/freight_analysis/bca_study_phase2/index.htm

ICF International (2009), "A framework for collaborative decision making on additions to highway capacity", SHRP2 Project C01, Transportation Research Board. http://onlinepubs.trb.org/onlinepubs/shrp2/SHRP2prepubC01.pdf

Kaliski, J., S. Smith and G. Weisbrod (1999), "Major corridor investment-benefit analysis system," *Transportation Research Record*, Vol. 1732, Transportation Research Board. www.edrgroup.com/pdf/mcibas.pdf

North, D. (1961), *The economic growth of the United States*, *1790-1860*, pp. 157-207, Prentice Hall, NY. https://archive.org/stream/economicgrowthof00nort

Oscar F./TPA (1993), "The relationship between economic development and transport links, stage 2: economic effects of the M74 northern extension", Glasgow Development Agency.

Rosenfeld, S., C. Liston, M. Kingslow and E. Forman (2000), "Clusters in rural areas: auto supply chains in Tennessee", Tennessee Valley Authority. http://rtsinc.org/publications/pdf/auto_supply.pdf

Shirley, C. and C. Winston (2004), "Firm inventory behaviour and the returns from highway infrastructure investments", *Journal of Urban Economics*, Vol. 55, pp. 398–415.

Steer Davies Gleave (2010), "Hamilton King-Main rapid transit benefits case", prepared for Toronto Metrolinx. www.metrolinx.com/en/docs/pdf/board_agenda/20100219/Hamilton_BCA_FNL_DRAFT.pdf

Thompson, E., D. Rosenbaum and A. Hall (2008), "Evaluating roads as investments: A primer on benefit-cost and economic-impact analyses", University of Nebraska, Bureau of Business Research Publications, No. 5. http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1010&context=bbrpub

TPICS (2013), Transportation Project Impact Case Studies, www.tpics.us

US DOT (2014), "Notice of funding availability for the Department of Transportation's national infrastructure investments". https://www.transportation.gov/sites/dot.gov/files/docs/TIGER%202014%20NOFA_FINAL.pdf

Vadali, S. and S. Chandra (2014), "Buyer-supplier transport access measures for industry clusters", *Journal of Applied Research and Technology*, Vol. 12/5, pp. 839–849.

Venables, A., J. Laird and H. Overman (2014), "Transport investment and economic performance: Implications for project appraisal", (UK) Department for Transport. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/386126/TIEP_Report.pdf

Volpe Transportation Systems Center (2012), "Criteria used in state DOT long range plans", US DOT. https://www.planning.dot.gov/documents/State_plans_report_508_A.PDF

Wang, Baojin (2015), "Estimating economic impacts of transport investments using TREDIS: A case study on a National Highway Upgrade Program", *Australasian Transport Research Forum 2015 Proceedings*. http://atrf.info/papers/2015/files/ATRF2015_Resubmission_180.pdf

Weisbrod, G. and J. Beckwith (1991), "Measuring economic development benefits for highway decision-making: the Wisconsin case", *Transportation Research Record*, Vol. 1262. www.edrgroup.com/pdf/wisconsin-hwy-tq.pdf

Weisbrod, G. (2006), "Evolution of methods for assessing economic development impacts of proposed transportation projects", *International Conference on Transportation and Economic Development*. www.edrgroup.com/pdf/weisbrod-evolution-of-methods-ted2006.pdf

Weisbrod, G. and Simmonds, D. (2011), "Defining economic impact and benefit metrics from multiple perspectives: lessons to be learned from both sides of the Atlantic", *European Transport Conference Proceedings*. www.edrgroup.com/pdf/Weisbrod-Simmonds-ETC-Oct2011R.pdf

Weisbrod, G. et al. (2014), "Assessing productivity impacts of transportation investments", National Cooperative Highway Research Program, Report 786, Transportation Research Board. http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_786.pdf

Weisbrod, G., C. Mulley and D. Hensher (2015), "Recognising the complementary contributions of cost benefit analysis and economic impact analysis to an understanding of the worth of public transport investment: A case study of bus rapid transit in Sydney, Australia", *Proceedings of Thredbo 15 Conference, Santiago, Chile.* www.edrgroup.com/pdf/Weisbrod-Mulley-Hensher-Thredbo14-paper.pdf

Worsley, T. and P. Mackie (2015), "Transport Policy, appraisal and decision-making", Institute for Transport Studies, University of Leeds for the RAC Foundation. www.racfoundation.org/assets/rac_foundation/content/downloadables/Transport_policy_appraisal_decision_making_worsley_mackie_May_2015_final_report.pdf

Wornum, C. et al. (2005), "Montana highway reconfiguration study", Montana Department of Transportation. www.mdt.mt.gov/other/research/external/docs/reconfig/final_report.pdf

Chapter 5

Incorporating wider economic impacts within cost-benefit appraisal

Anthony J. Venables[1]

This chapter analyses three main mechanisms through which transport improvements have impacts that deliver real income gain over and above user-benefits. One is economic density and productivity, a second is induced private investment and associated land-use change, and a third is employment effects. There are relatively well-established methodologies for incorporating the first and third of these in cost-benefit appraisal, and these methodologies are reviewed in the chapter. For the second, the chapter outlines how transport induced investments can create consumer surplus, and describes a method for quantifying this in cost-benefit appraisal. Data issues encountered in implementing these methods are discussed.

1 Department of Economics, University of Oxford, UK

Introduction

The case for investment in transport improvements is frequently made in terms of impact on economic performance. There is an expectation that they will act as a catalyst for private sector investment, creating jobs, boosting economic activity and growing (or rebalancing) the local (or national) economy. These "wider economic impacts" typically go beyond a conventional transport cost-benefit appraisal (CBA) which focuses on the user-benefits created by a project, often derived under the assumption of no change in land-use. This is an unsatisfactory situation which creates a disjoint between the strategic arguments put forward in support of a project, and the associated economic analysis and CBA. Even if the value of wider economic impacts turns out to be small, appraisal must engage with the arguments put forward by scheme promoters and local interests or otherwise risk marginalisation, resulting in a policy process in which decisions are based on bad economics.

Incorporating wider economic impacts in CBA is challenging and has its own risks. Broadening the set of mechanisms that are studied creates the risk that bad arguments may appear to be legitimised, and that effects can be exaggerated. Studies tend to concentrate on areas where a transport improvement expands economic activity, and to ignore areas from which this activity may have been displaced. This, together with reporting of gross value added (GVA) effects, makes it possible that fundamental economic principles – above all that drawing resources into an activity has an opportunity cost – can be overlooked. The challenge is to be ambitious in broadening the scope of appraisal while remaining grounded in rigorous analysis of the social value of transport investments and of any private sector responses that they induce.

How should this be done? One answer is a full economic modelling exercise, in which resource constraints are properly imposed, private sector responses are modelled, market imperfections are made explicit, and real income (utility) benefits are accurately calculated. This may be appropriate for some large projects, but is not a general solution. Such models are expensive and it would be disproportionate to use them for the majority of projects. A consequence of their expense is that typically one model is built and then applied to different situations in a somewhat mechanical manner, paying insufficient attention to the characteristics of the scheme and its likely effects. They then fail to capture impacts, which are likely to be quite different for an urban commuting scheme, an urban by-pass, or an inter-city rail line. These projects have different stated objectives and will trigger different private sector responses. It follows that the appraisals must be designed to be context specific. Some should focus on the consequences of getting more people into a city centre, others on better linking remote locations, and so on.

The need, therefore, is to develop a framework of possible channels or mechanisms through which wider economic impacts can occur, and to find the evidence needed to quantify these mechanisms and apply them in appraisal. The application of these mechanisms to particular projects needs to be context specific, informed by the strategic narrative that motivates the project; some mechanisms are applicable to some types of transport projects, others to others. For larger projects the mechanisms can be formulated in a complete economic model. For other projects this has to be done by the analyst's linear approximation to the formal model; this means that component parts will be studied separately and then added up. Of course, the relationship between the components must be consistent (so adding up does not double-count), the components must be exhaustive (so if some activity expands others may contract), and the focus should be on identifying the true social value of effects.

This chapter sets out and discusses the key components of this approach. The next section gives the outline and relationship between components, which are then discussed more fully in the remainder of the chapter. Some of the elements are now well established, and are applied in practise, for example in the UK Department for Transport's appraisal guidance.[1] Others are more challenging and need to be the subject of further research.

The effects of a transport improvement

A transport improvement brings time and cost savings to users of the transport network.[2] The users are individuals and households in their work and leisure activity, and firms which need to move goods, services, and employees. Time and cost savings change traffic flows, leading to increased flows in some parts of the network and possibly less traffic elsewhere. These changes in costs and in flows are the subject of sophisticated modelling efforts and are the core of transport appraisal. They are illustrated in the left hand column of Figure 5.1. We follow practise in the transport literature and refer to the social value of these change as the *user-benefits* of a project.[3] While they constitute the centre of any transport appraisal, they are not the focus of this chapter.

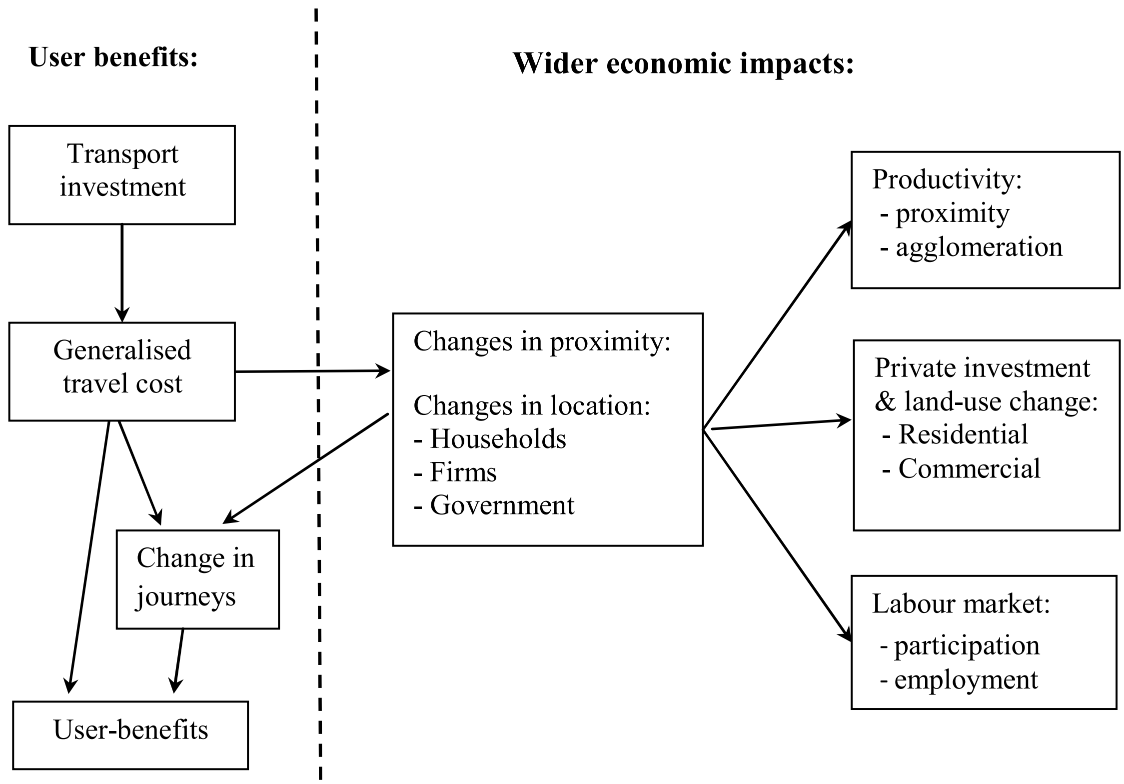
Wider economic impacts are illustrated in the right hand part of Figure 5.1 and arise as a consequence of transport's impact on economic geography. Better transport increases proximity, making economic agents closer together, and may also trigger relocation of economic activity as firms and households respond to new opportunities. Together, these changes create potential sources of "wider economic benefit" through three main mechanisms.

The first is that proximity and relocation shape the effective density of economic activity, and thereby productivity. This is over and above the direct productivity effects of faster journeys, and arises because of the intense economic interaction that occurs in economically large and dense places. This is why cities and other agglomerations exist. This observation is backed-up by a substantial research literature that quantifies the positive relationship between economic density and productivity.

Second, a transport improvement, other things equal, will make affected locations more attractive destinations for investment. User benefits are experienced by residents, workers, and firms, and this may induce investment to occur, changing land use. Investments include residential development of land, the development of office centres or retail parks, or the redevelopment and regeneration of city centres. They may in turn generate agglomeration and productivity effects, and have further value by changing the "attractiveness" of affected places.

Third, there may be impacts in the labour market, on both the supply and demand side. On the supply side, transport may enable labour force participation. On the demand side, jobs will be created in some places and some activities, while possibly lost in others.

Figure 5.1. **The effects of a transport improvement**



To include these impacts in transport appraisal three questions must be addressed. First, is there a sound reason to think that they create a social value, over and above user-benefits? This requires understanding the mechanisms at work and, essentially, identifying a market failure. Absent such failures, (small) quantity changes are of zero social value, as the price system equates the marginal value of changing an activity to its marginal cost. But if transport induces a change that interacts in some way with a market failure then it will create additional benefit (or cost). Notice that these valuations are in terms of social welfare (ultimate household benefit), not of GVA. The distinction between the two is well known, and the focus throughout this chapter is social welfare.

Second, local changes have to be set in the context of the national aggregate. In practise, this means thinking hard about *displacement*. For example, job creation in one region may be at the expense of job losses in another. Each change may be of interest to local stakeholders, and it may therefore be appropriate to report them in project appraisal. However, national appraisal has to report national aggregates and provide the complete view which may be missing from concentration on effects in the neighbourhood of the project.

Third, the feasibility of predicting and quantifying effects has to be considered. This is technically challenging, particularly for large projects which are claimed to have transformative effects. The difficulty is compounded by complementarities between transport projects and other actions, including other policy changes. A single transport project is unlikely to be sufficient to unlock transformative change, its value depending on complementary transport improvements, land-use planning changes, and perhaps even wider demographic changes. Addressing this requires multiple scenarios, rather than the presentation of a single benefit-cost ratio (BCR).

The next three sections of this chapter look, in turn, at the impact of transport improvements on productivity and proximity, induced private investment and land-use change, and labour market effects. The emphasis will be on mechanisms through which transport may create wider benefits – i.e. the way in which the economic impacts interact with market failures to create sources of additional gain. The penultimate section addresses the issue of how, in practice, it might be possible for appraisal to make quantified estimates of effects, particularly at a national level, and including displacement effects.

Proximity and productivity

It is widely recognised that economic density – the clustering of activity in towns and cities – has a positive impact on productivity, and that such clustering is dependent on effective transport systems. Some of the productivity effects come from interactions between different economic agents that are not fully internalised, creating market failure and wider economic benefits, as recognised in the appraisal methodology of the UK Department for Transport.

Mechanisms

Transport improvements enable savings in transport and communication costs for firms, workers, and consumers, enhancing effective proximity. In turn, cheaper, more reliable and faster transport may allow firms to change the way in which they organise their logistics or production (e.g. just-in-time manufacturing technologies). These gains are user-benefits, and are accounted for in calculation of those benefits. They should not be double-counted as a wider economic impact.

Wider economic impacts arise when economic agents cannot capture the entire benefits (or costs) of their actions, i.e. they create externalities that are of value for other agents.[4] These may be technological (such as knowledge spillovers, which are not intermediated through a market) or pecuniary (going through an imperfect market). By supporting thicker markets and more intense economic interaction, proximity creates a number of these effects. Probably the most important mechanism is that scale and density together create an environment where firms and workers can develop highly specialised products, services and skills. These are typically inputs to firms – the specialist components, and engineers, lawyers, finance experts who may be necessary to efficient operation of a firm.[5] A new specialist supplier will set up once the market is big enough and their presence will make the cluster more attractive as a location for other firms that use the product or service. This grows the market for specialist suppliers, encouraging further entry and hence a cumulative causation process. This is the classic process of cluster formation, such as an auto-industry cluster of assemblers and suppliers or a film industry cluster of directors, actors and technicians. There are spillover effects (externalities) in this process. Indivisibilities or increasing returns to scale mean that a service, skill, facility or product will only be supplied if the market is big enough. The supplier is generally unable to capture all of the benefit, so there is a positive net effect created for others in the cluster.

A further mechanism arises as competition is likely to be intense in a large and dense cluster so monopolistic pockets of inefficiency are less likely to survive. Monopsonistic behaviour, occurring where there are few potential purchasers for a product or skill, can deter investment; this too is less likely to be a problem in a large and dense cluster. There may also be direct knowledge spillovers, as “mysteries of the trade become no mysteries; but are as it were in the air” (Marshall, 1890).

The mechanisms may operate within particular sectors or across a wide range of sectors, the former being referred to as localisation (or Marshallian) economies, and the latter as urbanisation (or Jacob) economies. Within-sector productivity effects create a force for sectorally specialised clusters and possibly specialised cities. This varies across sectors; it is important in some manufacturing sectors, for example developed country manufacturing exhibits automotive clusters, and developing country manufacturing contains clusters in labour intensive sectors such as textiles and garments. Clustering is particularly prevalent in business services such as finance, law, and media. Both the creation and diffusion of knowledge work particularly well in clusters and a large body of literature points to the spatial concentration of innovative activities.[6]

Valuation

A reduced form approach to measure these effects has two elements, one to construct a measure of effective density or "access to economic mass" for each place, and the second to link productivity to this measure. The first stage, measurement of access to economic mass, typically takes the form $ATEM_i = \Sigma_j f(d_{ij})Emp_j$. This says that location i's access to economic mass, $ATEM_i$, is the sum of employment in all districts (indexed j), weighted by some decreasing function, f, of their economic distance to i, d_{ij}. Thus, if a place is near many other places with high employment it will have high *ATEM*. The second step links a location's access to economic mass to its productivity through the relationship, $Productivity_i = F(ATEM_i)$.

A substantial econometric literature quantifies these relationships, seeking to find functions F and f by estimating equations of the form:

$$Productivity_i = F(\Sigma_j f(d_{ij})Emp_j). \tag{1}$$

Appendix I reviews key studies and here we simply note that relationships can be investigated at a sectoral level (localisation economies) or an aggregate level (urbanisation economies). Economic "distance" can be measured in different ways (distance, travel time, or generalised travel cost), and economic activity represented by employment or by other activity measures. The unit of observation can be spatial aggregates (e.g. location/sector averages) or can be individual firms or workers. Estimation includes controls for other determinants of productivity; for example, if the unit of observation is a worker, then skills and age will be amongst the controls.

A reasonable consensus has emerged on the magnitude of effects. An authoritative (although quite old) survey of the literature finds that "in sum, doubling city size seems to increase productivity by an amount that ranges from roughly 3-8%" (Rosenthal and Strange 2004). This means that the elasticity of productivity with respect to city size is in the range 0.05-0.11.[7] This is a large effect in the cross-section, suggesting that productivity in a city of 5 million is between 12% and 26% higher than in a city of 500 000. A meta-study (Melo et al., 2009) suggested that the mean estimate of this elasticity across several hundred studies is somewhat lower, at 0.03, although pointed to considerable variation according to sector, country, and technique employed by researchers. Recent work using individual data (and controlling for individual effects) produces estimates of similar magnitude. At the sectoral level, there is evidence of heterogeneity, with business services and high technology sectors exhibiting the largest localisation economies.

A critical issue for transport appraisal is the construction of the measure of access to economic mass $ATEM_i = \Sigma_j f(d_{ij})Emp_j$."Distance", d_{ij}, is typically measured as a composite of generalised travel costs (GTC) of different modes of transport. The composite can be constructed either by assuming weights of different modes in some functional form (e.g. an index using modal shares as weights) or by letting econometrics determine the contribution to productivity of access by different modes. The latter is preferable, but hard to identify precisely as the GTC of different modes are highly correlated across origins and destinations. The spatial scale of effects (captured in the function *f*) is generally found to be quite limited, with effects concentrated within travel to work areas (e.g. driving times of up to 45 minutes) and attenuating quite rapidly thereafter.

The effect of a transport improvement

A transport investment can change access to economic mass in two distinct ways. One is that it changes levels of activity in each place, Emp_j. This is sometimes referred to as "dynamic clustering" and is associated with land-use change; we discuss it further in following sections. The other is a direct proximity effect. Transport changes the matrix of economic distances (GTC) between places, d_{ij}, making places better connected and increasing the effective density of economic activity. This is sometimes referred to as "static clustering". Implementing this source of wider impact does not require estimates of induced investment response or land-use change. A transport appraisal will have estimates of how the project will change the matrix of GTC between places. This can be fed into equation (1) and the ensuing productivity changes for each location can be computed. A productivity increase derived this way is an additional source of welfare gain – a wider benefit, on top of user-benefits.

Investment and changes in land-use

A transport improvement will generally change the pattern of private investment across locations, and this process of encouraging – or even "unlocking" – private development is often put forward as one of the major impacts of transport projects. The investment response is driven by the user-benefits experienced by residents, workers, and firms. This response changes traffic flows, changes which should be included in calculation of user-benefits (Figure 5.1). Are there circumstances in which the induced investment creates wider benefits, additional to the user-benefits? We address this in two different contexts, first looking at residential development that is dependent on transport improvement, and then at relatively large scale commercial developments – such as city centre redevelopment – for which transport improvement is the catalyst.

Dependent residential development

Transport is a necessary part of many new residential developments, sometimes at a large scale. The proposed "Crossrail 2" in London is linked to construction of 200 000 new homes in North London. What are the circumstances under which this leads to benefit over and above those accounted for in user-benefits? The economic principles for valuing any such change in land-use are straightforward, but worth restating.

Suppose that initially the number of houses in a particular area is Q_0, and the area has a transport improvement which gives the residents of each house user-benefit ΔT. In the new situation – after the transport improvement and any other policy changes are made – the number of houses increases to Q_1. Using standard demand and supply analysis (Appendix II), the change in welfare consists of two parts. The first is user benefits (UB) to existing and new residents, approximated by the rule of a half, i.e. $UB = \Delta T\{Q_0 + (Q_1 - Q_0)/2\}$. The second captures any inefficiencies in land use, as measured by the wedge between price and marginal cost (where marginal cost is the value of the land in its alternative use plus house construction costs and any further costs, such as induced congestion externalities). Calling this wedge in the initial and final situations PC_0, PC_1, the extra social value derived is the number of new houses built times the average value of the wedge, $WB = (Q_1 - Q_0)(PC_0 + PC_1)/2$.

The wider-benefit, WB, part is proportional to the average gap between marginal social benefits and marginal social costs. What supports such a gap? One possibility is planning restrictions, although only if they are more restrictive than is efficient. Thus, the planning authority may place a high value on congestion costs or other negative externalities created by the development, this narrowing the gap between marginal benefits and costs. Another possibility is that there is monopoly power in the supply of housing. Developers holding large stocks of land will restrict supply (equating marginal cost with marginal revenue, not price), and thereby creating such a gap. It is possible that this monopoly power is not exercised by developers, but by existing residents who have captured the planning process and are seeking to restrict building in order to maintain high property prices.

It is sometimes claimed that the full gain (UB + WB) is given by the land-value uplift. This is true only if two conditions hold. First, that the increase in supply of housing does not reduce price (see Appendix II). This requires that the price elasticity of demand is infinite (or the extra supply extremely small). Otherwise, if the elasticity is finite (the demand curve is downward sloping), then price is reduced so not all the benefit goes to land owners, some accruing to house occupants.[8] Land value uplift then underestimates the welfare gain. The second condition is that all externalities (such as increased congestion) are fully accounted for and charged in the calculation of costs. If not, then the presence of uncharged negative externalities will mean that land value uplift over-estimates the benefit of development.

Commercial land-use change

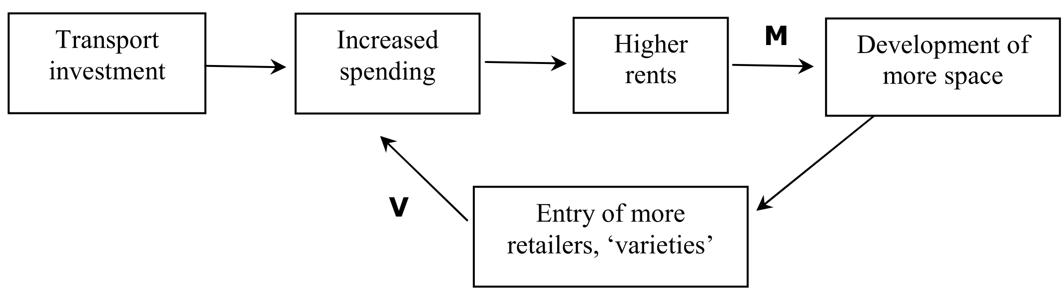
A more complex situation arises if transport acts as the catalyst that induces private investment in a large commercial development – retail, office – and perhaps involving redevelopment of a substantial parcel of city land. It is often suggested that such developments create an additional benefit by making an area “more attractive”.[9] Under what circumstances do these benefits exceed the user-benefits received by travel to and from the area?

A conceptualisation of this is offered in Figure 5.2; the context developed in the figure is that of a retail development, although the arguments put forward are more general.[10] A transport improvement increases spending in a place, as visits respond to lower travel costs. Increased spending raises profitability of shops and hence the landlord is able to charge higher rents. This makes it profitable to develop more space, redeveloping the site – by extension, or perhaps by building taller. This expansion creates more floor space and hence the entry of more shops, in turn making the place a more attractive destination and creating the feedback loop illustrated in the figure.

User-benefits trigger this process, and wider-benefits arise if there are interactions with market failure. There are, arguably, two sources of market failure in this process, labelled M and V in Figure 5.2. The first, M, arises as there may be barriers preventing the level of development reaching an efficient level and hence creating gaps between marginal benefits and costs. The second is at point V, and captures the idea that places become more attractive as they attract more stores. We look first at the attractiveness argument, V, and then turn to barriers to development, M.

The attractiveness argument has foundation if entry of new stores creates some consumer surplus, i.e. consumer utility over and above the value of their spending. This will arise if stores are differentiated from each other, and is formalised in many sub-fields of economics as a variety effect. For example, in international trade it is argued that much of gains from trade (at least, intra-industry trade between similar countries) arises from countries being able to access a wider range of products (for quantification of these effects see Broda and Weinstein, 2006). By analogy, introducing new stores in a retail development creates consumer surplus since it increases the range of choice (number of varieties) available to consumers. The standard methodology for quantifying the gain assumes that demand for the products under study is iso-elastic. Denoting this elasticity σ, the ratio of consumer surplus to expenditure on a new variety is 1/(σ – 1) (Appendix II). Hence, the value of any variety effect, UV, is UV = change in expenditure/(σ – 1). If products are perfect substitutes – the retail development just means more identical stores – then σ is infinite, there is no increase in "attractiveness" and UV = 0.[11] Typical estimates of σ from other contexts suggest values in the range 6 – 10, suggesting a wider benefit mark up of 10-20% of expenditure in the development.

Figure 5.2. **Commercial development**



Three further remarks need to be made about the variety effect. First, following the approach above, it is grounded as a mark-up factor on the change in consumer expenditure in the development. This is project-specific data that is observable *ex post* and likely to be part of development plans at the appraisal or planning stage. Thus, estimates of possible wider-benefit created can be tested against the commercial proposition put forward by developers. This avoids having to resort to ad hoc shifts in demand curves in order to capture these effects.

Second, the discussion has been in terms of retail development. An exactly analogous argument applies to an office development scheme, but with the variety effect restated as an agglomeration effect. In both cases entry of a new firm (shop or office) creates a positive spillover, as the entrant is unable to capture the entire benefit created. This analysis is therefore a restatement of the agglomeration and productivity arguments of the previous section. Of course, only one of the two approaches should be followed for any particular project.

Third, these arguments (and those of the preceding section) have to be placed in the context of product market displacement effects. Would the activity – manufacturing, commercial or residential – take place somewhere else, absent the transport improvement? If so, is it subject to the same market failures? Effects across all geographical areas then have to be combined – some of them positive, and others negative. We return to this issue in the penultimate section below.

We now turn to the other possible source of market failure, the presence of barriers to development, M. Many of the points follow the earlier discussion of residential development. Thus, there may be monopoly power as a developer perceives that building extra space reduces rents. The planning system may over-restrict development, particularly if it only looks at the interests of local residents in the development of a scheme that could bring benefits to a more spatially dispersed group of shoppers or workers. As with residential development, an increase in quantity supplied brings wider benefit proportional to the gap between marginal social benefit and cost.

Additional barriers may be present in large-scale commercial developments as they involve investments by many distinct decision takers – property developers and retailers in the conceptualisation of Figure 5.2, or perhaps multiple developers in a large scheme. If the profitability of the project for one decision taker depends on investment by others (as illustrated by the feedback mechanism of Figure 5.2), then there is potential for co-ordination failure. It is not in the interest of any single investor to invest, but each would invest if they knew that others were doing likewise. This positive interdependence of profitability could arise in starting a new cluster of economic activity (i.e. the productivity arguments of the previous section) or in launching new retail or urban redevelopment schemes. Co-ordination failures thus lead to low level traps and require some policy mechanism to co-ordinate individual actions and break out of the trap. Transport investment can be such a mechanism.

A simple example of this argument is a growing city in which it is clear to all that a secondary centre somewhere on the edge of the city will be successful, but there is no agreement as to exactly where. The expected return to a private investment in any particular place is therefore low or negative, since this may not turn out to be the place that takes-off. This uncertainty creates the low level trap – no-one invests anywhere. There are different ways to resolve this problem. A sufficiently large private developer could move first, being relatively confident of being followed by other investors. The city authorities can produce an urban plan, selecting areas for development. Or transport infrastructure can be built. This now has a dual function; it delivers access and user-benefits and also is a credible signal that a particular place will develop. If this resolves the co-ordination failure then the return to the investment can, potentially, be many times greater than the user-benefits alone.

A different example of co-ordination failure is regeneration of a dilapidated area of a city. It is not worthwhile for one property owner to improve, given other properties remain run-down. But if all do, all are better off. The role of transport as a catalyst to break this trap is less clear-cut than in the previous example (uncertainty is not about where, but about the likelihood of action). However, by increasing the value of properties in the area a transport project may also increase the return to property improvement; if some improvements are initiated it may cascade, raising returns to others.

Evidently, assessment of these effects is context specific and subject to a great deal of uncertainty. Studies of the role of transport in these contexts (e.g. in regeneration schemes) frequently suggest that transport is an important part of a package of measures, but is unlikely to be transformative by itself. More generally, there is considerable inter-dependency between transport and other public projects and policies. Synergies extend not just across transport projects and associated private development, but also across government policies, including land-use policy and wider urban and regional development measures. Transport appraisal needs to recognise potential synergies arising from interaction between policies. If each element of a policy package is necessary for change, and no one of them is independently sufficient, then the scheme has to be evaluated as a whole. Scenarios can be produced of the effects of different combinations of policy and other changes, and each scenario can be valued. However, it is not generally meaningful to attribute returns to each separate part of an integrated policy package.

Employment impacts

Job creation is often held up as a major impact of transport investment, with two distinct mechanisms being suggested. One is on the supply side: better transport may make it easier for people to get to work, and may reduce discouraged worker effects. The other is on the demand side, with induced investment creating new employment opportunities. We discuss each in turn noting that, as usual, the benchmark is a situation where a change in quantities – of jobs or other variables – is of zero social value.

Labour supply: Participation and tax wedges

On the supply side, individuals' labour force participation decisions are based on comparing the costs of working (including commuting costs), against the wages earned from a job. By reducing the cost (in time and money) of getting to work, a transport investment is likely to increase the returns to working; some people, for whom the net returns to entering the labour market were initially not worthwhile, may decide to enter. Such an increase in labour supply and employment raises GVA but, in the simplest circumstances, does not increase welfare. Initially, the individual was not working because the utility from leisure exceeded that from working, once commuting costs are accounted for. If a transport improvement triggers work, the benefit to the individual cannot be greater than the user-benefit received (if it were, the individual would have chosen to work in the first place). However, this conclusion changes if there is an income tax wedge (or loss of state benefits). The individual does not receive the full value of work undertaken because a fraction of it accrues to government. The full gain from entering employment is then the user-benefit plus tax revenue paid (or benefits not received).

This is operationalised in the UK's transport appraisal by calculating the change in the generalised cost of commuting; then estimating how this increased return to working affects the amount of labour supplied (via an elasticity of labour supply with respect to earnings), and hence calculating how much more income is generated and how much of this accrues as income tax (or benefits not paid). By the argument above, only the tax raised (or benefit saved) by the additional employment and output is included as an additional benefit from the scheme.

Similar principles apply if transport triggers a move to more productive jobs. For example, suppose that there is a low paid job nearby and higher paid jobs further away. A reduction in the cost of travel might cause individuals to switch to the higher paid jobs. However, their calculation of the net private gain from switching jobs is based on post-tax income, not the pre-tax wage. The exchequer captures the tax wedge in this decision. This is exactly analogous to the participation decision discussed above, and was part of the Crossrail appraisal (see Box 5.1).

Labour demand and unemployment

Turning to the demand side, if new jobs are created in one place, then the value of output produced by each new job is the wage, and this is set against the value of what workers would have done, absent the new jobs. For workers drawn out of involuntary unemployment the alternative is of low value, so the net benefit is large. This may be an important effect in developing economies or in regions with significant structural un- (or under-) employment. However, for long-run transport projects in reasonably well functioning market economies it seems likely that the labour market will adjust to some "natural rate" of unemployment which is independent of transport investment. If this is the case then an increase in labour demand is met either by increased labour force participation or by drawing workers out of other employment. If demand is met by increased labour force participation then its value is, as above, the tax wedge on income. If it is met by withdrawing labour from other activities, then the value is the alternative wage. There is no net benefit if wages are the same in both jobs. Displacement is 100%, so demand induced employment effects should, from the national perspective, be ignored.

A qualification to this argument is conceptually important, although perhaps not quantitatively large for any single transport project. To draw labour from other activities there may have been an increase in wage rates in the area affected or more broadly. Given the level of productivity, an increase in wages must be financed either by a reduction in profits (or more generally, payments to other inputs), or by an increase in prices. The increase in wages is therefore just a transfer, of no value to aggregate income, unless the people paying for it (consumers and recipients of profits) are, for some reason, people that we do not value. A standard approach would be to suggest that benefit arises to the extent that the increase in price is paid by foreigners, i.e. represents a terms of trade improvement, so the country is able to sell its exports at higher price. This is an additional source of benefit, although one that is unlikely to be quantitatively significant for any single transport project.

Predicting quantity changes

Preceding sections of this chapter have concentrated on the sources of wider benefits and the way in which they can be valued. To apply this in appraisal requires that forecasts can be made of the quantity changes (changes in investment, output and employment, as well as changes in traffic) that are likely to follow from a transport improvement, and which drive the wider impacts. These quantity changes are principally in the neighbourhood of the project, but may also occur elsewhere in the economy, important for establishing displacement effects. There are several – complementary – ways of getting the information required to forecast these quantity changes. One is from the technical details of the project itself. This can be combined with knowledge about the characteristics of the areas and sectors affected. Another is from spatial modelling, computing effects on activities throughout the economy. All approaches need to draw on past experience – both from case studies and from econometric analyses.

Project information

Standard project documentation contains forecasts of levels and changes in generalised transport costs and traffic flows (albeit, often derived with an assumption of fixed land-use). These are necessary to compute the user-benefits of a transport improvement and – with the assumption that all other changes are of zero social value – are also sufficient. To what extent is the information needed to calculate wider benefits contained in this documentation?

First, consider productivity effects. Recall that these operate through two distinct mechanisms, static and dynamic. The former is the change in "distance" (as measured by generalised transport costs), given the location of economic activity; evidently, this information is available from project documentation. The second is the change in economic activity (perhaps as measured by employment, either in aggregate or by sector) in places affected by the project. In some projects information about this is implicit in the project specification. A commuting project contains estimates of the capacity change in the system and hence forecasts of passenger flows. If people are commuting to jobs then the response of employment in each place is implicit, if not explicit, in the passenger forecasts. Furthermore, the characteristics of the place served by the project are known, and forecasts of agglomeration and productivity effects follow from this. The appraisal of London's Crossrail project was based on information of this type (see Box 5.1 and Worsley, 2011).

Similar arguments apply to other projects that lead to land-use change. If a project unlocks residential development or is intended to lead to redevelopment of an urban area, then the planning system has projections of changes in residential and commercial land-use. These should be used in transport appraisal, both in order to get accurate traffic flows, and to evaluate the combined impact of the transport project and other dependent development. As suggested in the section on land use, the wider benefits of changes in "attractiveness" should be based on estimates of expenditure created in dependent developments, estimates that will have been made during design phases of the project. The issue is therefore, not whether the relevant information exists, but ensuring that it is used in transport appraisal.

The third category of wider impacts occurs in the labour market. Labour supply – changes in participation rates or moves to more productive jobs – are inherently local and project specific, and follow from information discussed in the preceding two paragraphs (see the Crossrail example in Box 5.1). By contrast, labour demand is more likely to impact through the national labour market and, as suggested above, is likely to displace workers from other jobs.

Box 5.1. **Commuting to a cluster: Crossrail**

One of the objectives of Crossrail was to increase employment in central London. This, it was suggested, could lead to two sorts of wider benefit. By expanding employment, productivity in the cluster would increase. There would also be a "move to more productive jobs", as workers in central London earn on average 20% more than workers in outer London.

To capture the first of these, studies took an elasticity of productivity with respect to employment of 0.06 and employment growth of 5% giving members of the cluster an average productivity increase of 0.3% (=0.06x0.05), the present value of which equals around two-thirds of estimated user-benefits. (These numbers illustrate the range of cases used in the appraisal). Movement to more productive jobs raises GVA, but only a fraction of this is an increase in real income (welfare). An indifference condition means that the wage differential between central and outer London is matched by the commuting costs of reaching central London. However, individuals' calculation of the net gain from switching jobs is based on post-tax income, not the pre-tax wage. Tax on the incremental income accrues to government, and it is this wedge that constitutes the net social benefit. This revenue effect (a real income gain) adds a further amount to the appraisal, worth somewhat more than user-benefits.

Crossrail appraisal provides a clear-cut example of how to calculate wider impacts in a project specific manner, although it rests on some fine judgements. For example, the quantity effects come directly from the increased capacity of the commuter network. In other cases it might require a fuller modelling exercise, such as use of a LUTI model or similar. The example assumes that additional workers in central London add to the cluster of activity and raise productivity, but the displacement of workers in outer London (or wherever they ultimately come from as equilibrium effects work through the economy) has no offsetting negative productivity effect elsewhere.

Calculation of the value of "moves to more productive jobs" requires information on jobs created and displaced, the wage and productivity levels of each, and a narrative indicating why, in the initial situation, workers were not taking the higher wage and productivity jobs (i.e. why comparable jobs with different wages could co-exist). These elements were all present in the Crossrail appraisal, with commuting costs the barrier to taking the jobs, and fairly clear cut wage differentials. In other contexts these differentials are likely to be much smaller and harder to identify.

Econometrics

Project specific and local information needs to be combined with evidence, derived from econometric analyses and from case studies. Such evidence makes three principal contributions. The first is to provide the elasticities – responses of one variable to changes in another – on which analysis is based. Traffic forecasting is dependent on such elasticities, and so too are the wider benefits mechanisms of productivity (as discussed in Appendix I), surplus derived from consumer demand systems and labour force participation. A second contribution comes from the aggregate studies of the role of transport infrastructure on economic performance. These are in the tradition of Aschauer (1989, 1990), and summarised in the meta-study of Melo et al. (2013), suggesting that elasticities of private output with respect to the stock of transport infrastructure are positive but small, around 0.1 or less. These studies are too highly aggregated to provide precise estimates of effects of particular projects but are useful as a reality check, setting bounds against which estimates of effects of particular projects should be compared.

The third contribution comes from studies of the effects of particular transport improvements. Such studies – be they descriptive case studies or econometric analyses – are fraught with methodological difficulties. Above all, it is difficult to establish the counterfactual, i.e. what would have happened in the absence of the project. New lines of research are making progress on this by comparing areas that are "treated" by a transport improvement with areas that are similar but "untreated".[12] Comparison of treated and untreated areas still poses an identification challenge, as it is hard to establish that transport investments played a causal role in emergent differences between the areas. Did a region boom because a road was built, or was the road built because of an anticipated boom? The literature is getting to grips with these issues by a combination of use of instrumental variables and searching for "natural experiments" (situations where a transport investment was made for reasons uncorrelated with expected economic performance). Significant positive effects of transport infrastructure are found for studies as diverse as Chinese railways and roads in the US and the UK. Problems remain however. Studies typically compare treated and untreated areas, a method that provides no way of splitting effects between positive impacts in treated areas and negative impacts (e.g. due to displacement) in untreated areas. There is also the inherent difficulty of generalising from one case study to the particular circumstances of a new project. But despite these difficulties, progress is being made and it can be expected that accumulation of a large set of high quality studies of past projects will usefully inform the design and likely impacts of new projects.

Spatial modelling

Spatial modelling pulls together project information and data from other sources into a formal structure in which the effects of a project can be simulated. Two frequently used techniques are land-use transport interaction (LUTI) models and spatial computable equilibrium models (SCGE). The advantage of such models is that they are able to give a fuller spatial picture of the impact of a project, and incorporate general equilibrium responses in both product and labour markets. LUTI models generally have much finer spatial structure than SCGE, while SCGE allow for equilibrium response in more variables.

As noted above, these models are expensive, and not therefore appropriate for most relatively small scale projects. They are, in many cases, taken "off-the-shelf", and applied with insufficient attention to all the salient features of the project and situation that they are analysing. There is always a risk that undue weight is placed on estimated effects, even if these are driven by parts of the model that are not widely understood or subject to critical evaluation.

Nevertheless, such models potentially play a valuable role. In particular, they can be employed to develop alternative scenarios of the effects of a project. As we have argued, outcomes vary with the range of complementary policies that accompany a transport improvement, and with private sector responses. There are inherent uncertainties about private sector responses – particularly where these are potentially transformative, and involve some of the positive feedback mechanisms that are associated with wider benefits. These alternatives are best captured by presenting results for a range of scenarios, and formal computable models are well placed to do this.

Table 5.1 offers a summary of circumstances in which formal modelling may be an important part of establishing wider impacts, and circumstances where project information, along with econometric and case study knowledge, is likely to be sufficient.

Table 5.1. **Predicting quantity changes**

	Scope of appraisal Local project information	Spatial modelling needed
User benefit	User benefits	No: (Induced quantity changes are of zero value)
Productivity	Static clustering Dynamic clustering	No: (Changes in "distance" given employment) Yes: If likely displacement of activities with agglomeration potential
Investment and land-use change	Residential Commercial	No: (Quantities set by constraints elsewhere) No: If activity change determined by project design/capacity Yes: If likely displacement of activities with similar market failures
Employment	Participation and better jobs Unemployment	No: (Local effects only) Yes: If regional distribution is of interest. Yes: If displacement < 100%

Displacement

Finally, we return to the question of displacement, occurring if expansion of activity in one area is at the expense of contraction elsewhere. Is formal modelling necessary to capture displacement effects, or can judicious use of information about the project and its context provide sufficient guidance?

The two main channels through which displacement occurs are the labour market and product markets. If a transport project generates investment that creates new jobs, where do the workers come from? Some may come from increased labour force participation, a (local) supply response discussed above. Otherwise, the default position is that they are displaced from jobs elsewhere in the economy. The basis for this judgement is the view that in reasonably well functioning market economies employment – over the long run – is close to some natural rate. Displacement of 100% implies that employment effects are of zero value, so can be ignored – although it may be of interest to report them by sector or by region, and it is possible that policy makers put particular value on additional employment in some sectors or regions.

Turning to product market mechanisms, the issue is whether investment following a transport improvement is additional, or simply a relocation from another place where the activity was of equal value. For activities that are perfectly tradable internationally displacement it likely to be zero; if an internationally mobile vehicle assembly plant is choosing between two different jurisdictions, then the jurisdiction that attracts it is unlikely to see other vehicle assembly activity displaced.

For non-tradables the default position is reversed. A retail development is, to a large extent, drawing custom away from other locations, so displacement is high. Judgement then needs to be exercised as to impacts on the attractiveness of different locations. There may be threshold effects, so that large scale redevelopment of one area may bring the effects outlined in the section on land use while displacement is spread widely, only having marginal effects on other areas and not leading to equal loss of attractiveness. Thus, developing a cluster of activity – e.g. in financial services in the City of London – might draw financial service workers from elsewhere in the UK, but does not undermine another UK financial services cluster.

It is evident from these arguments that displacement effects are highly project and context specific. This reinforces the need to appraisals to be related to the strategic narrative on the project, backed up by knowledge of the sectors and markets that are likely to be affected.

Summary and conclusions

Transport investments can deliver economic benefits over and above conventionally measured user-benefits. They arise as transport fosters intense economic interaction that raises productivity; this can occur in clusters within narrowly defined areas or more widely by linking areas. Transport shapes the level and location of private investment, unlocking residential development and triggering large scale redevelopment of urban and other areas. Transport impacts the labour market, potentially enabling more workers to access jobs. These impacts can yield real income gains, particularly where transport-induced investments interact with market failures associated with increasing returns to scale, obstacles to efficient land use, and labour market imperfections.

Appraisal of transport projects has to combine relevance with rigour. Relevance requires context specificity. There should be a clear narrative of what each project is expected to achieve, and appraisal should capture the causal channels through which the project is expected to have impact. This suggests a modular approach (along the lines followed in this chapter and summarised in Figure 5.1). To maintain rigour, and comparability across projects, modules need to be based on a consistent set of principles. These should be grounded in economics and directed at identifying changes in real income (welfare). This means being careful to identify quantity changes throughout the economy. The value of such changes turns on market failures of some type, and need to be referenced against a benchmark of the "perfect" economy in which small changes are of zero social value.

Some mechanisms and associated appraisal modules are quite well developed and have sound evidence base, notably those to do with proximity and productivity, and with labour force participation and employment. Others, to do with land-use change, dependent development and co-ordination failure, are still in need of further refinement. Such work is relevant not just for appraising transport projects, but for appraisal of micro-economic policy change more broadly.

Notes

1 For fuller discussion of the issues in this paper and their relationship with UK practice see Venables et al. (2015).

2 Throughout we focus on the effects of the completed project. We do not investigate the construction costs of projects, nor include the temporary economic activity created by construction.

3 Of course, they do not necessarily accrue to the user as e.g. they may be shifted to rents and captured in land value appreciation.

4 Duranton and Puga (2004) survey these ideas.

5 The economics literature often models this as the presence of a large "variety" of intermediate inputs. Each variety yields consumer surplus that is not captured by the supplier (i.e. the supplier cannot perfectly price discriminate). See the next section for further development of this idea.

6 See for example, Audretsch and Feldman (2004), Glaeser and Gottlieb (2009).

7 Elasticities are therefore in the range 0.05-0.1 since 2^(0.05)=1.03 and 2^(0.11) = 1.08.

8 Owner-occupiers of existing houses being indifferent about the division.

9 A statement of the issue is given by Simmonds (2012): "if a transport change improves access to a town centre and causes an increase in demand for shopping and services there, this is likely to lead to an improvement in the retail offer of that centre, which will be an externality benefit to residents with easy access to that centre". See also Martinez and Arraya (2000), Geurs et al. (2006; 2010).

10 This is based on Venables (2016).

11 See Mankiw and Whinston (1986) for the possibility of welfare loss when products are perfect substitutes.

12 Redding and Turner (2014) survey some of this literature. Methodologically there is a parallel with drug trials: some areas are "treated" by having investment, others form the control group. However, it is not generally the case that assignment of areas for treatment is random, as it would be with individuals in a drug trial. Instrumental variables are used to address this problem. See Baum-Snow (2007), Donaldson (Forthcoming), Duranton and Turner (2012) for good examples of the approach.

References

Aschauer, D.A. (1989), "Is public expenditure productive?" *Journal of Monetary Economics*, Vol. 23, pp. 177-200.

Aschauer, D.A. (1990), "Highway capacity and economic growth" *Economic Perspectives,* Vol. 14, pp. 14–24.

Audretsch, D.B. and M.P Feldman (2004), "Knowledge spillovers and the geography of innovation" in *Handbook of Urban and Regional Economics*, Vol. 4, ed. Henderson, J.V. and J.F. Thisse, North-Holland, New York.

Baum-Snow, N. (2007), "Did highways cause suburbanization?" *Quarterly Journal of Economics* Vol. 122/2, pp. 775–805.

Broda, C. and D.E. Weinstein (2006), "Globalization And The Gains From Variety", *Quarterly Journal of Economics*, Vol. 121/2, pp. 541-585.

Ciccone, A. and R. Hall (1996), "Productivity and the Density of Economic Activity", *American Economic Review*, 87, 54–70.

Combes, P-P., Duranton, G. and Gobillon, L. (2008), "Spatial wage disparities: Sorting matters!", *Journal of Urban Economics*, Vol. 63/2, pp. 723–742.

De Palma, A, R. Lindsey, E. Quinet and R. Vickerman (eds) (2011), *A Handbook of Transport Economics*. Edward Elgar, Cheltenham.

Donaldson, D. (Forthcoming), "Railroads of the Raj: Estimating the impact of transportation infrastructure", *American Economic Review*.

Duranton, G. and D. Puga, (2004), "Micro-foundations of urban agglomeration economies", in J. V. Henderson and J. F. Thisse (ed.), *Handbook of Regional and Urban Economics*, Volume 4, Elsevier.

Duranton, G. and M.A. Turner (2012), "Urban growth and transportation", *Review of Economic Studies*, Vol. 79/4, pp. 1407–1440.

Fujita, M, P. Krugman and A.J. Venables (1999), "The spatial economy: cities, regions and international trade", MIT press.

Geurs, K.T., B. Van Wee P. and Rietveld (2006), "Accessibility appraisal of integrated landuse-transport strategies: methodology and case study for the Netherlands Randstad area", *Environment and Planning B: Planning and Design*, Vol. 33/5, pp. 639-660.

Geurs, K., B. Zondag, G. De Jong and M. De Bok (2010), "Accessibility appraisal of landuse/transport policy strategies: More than just adding up travel-time savings", *Transportation Research Part D: Transport and Environment*, Vol. 15/7, pp. 382-393.

Glaeser, E.L., and J.D. Gottlieb (2009), "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States", *Journal of Economic Literature*, Vol. 47/4, pp. 983-1028.

Graham D., S. Gibbons and R. Martin (2010), "The spatial decay of agglomeration economies: Estimates for use in transport appraisal". Report for Department of Transport, October 2010.

Mackie, P., D. Graham and J. Laird (2011), "The direct and wider impacts of transport projects; a review", in De Palma, A, R. Lindsey, E. Quinet, and R. Vickerman (eds) (2011), *A Handbook of Transport Economics*. Edward Elgar, Cheltenham.

Mankiw N.G. and M. Whinston (1986), "Free Entry and Social Inefficiency", *Rand Journal of Economics*. Vol. 17, pp. 48-58.

Marshall, A (1890), *Principles of Economics*, London and New York: Macmillan

Martinez, F. and C. Araya (2000), "Transport and land-use benefits under location externalities", *Environment and Planning A*, Vol. 32/9, pp. 1611-1624.

Melo, P.C., D.J. Graham and R.B. Noland (2009), "A meta-analysis of estimates of urban agglomeration economies", *Regional Science and Urban Economics*, Vol. 39, pp. 332-342.

Melo P.C., D.J. Graham, R. Brage-Ardao (2013), "The productivity of transport infrastructure investment: A meta-analysis of empirical evidence", *Regional Science and Urban Economics*, Vol. 43, pp. 695-706.

Puga, D. and J. de la Roca (2012), "Learning by working in big cities" CEPR Discussion Paper 9243.

Redding, S.J. and M. Turner (2014), "Transportation costs and the spatial organisation of economic activity", No. w20235. National Bureau of Economic Research.

Rice, P.G., A.J. Venables and E. Patachini (2006), "Spatial determinants of productivity; analysis for UK regions", *Regional Science and Urban Economics*, Vol. 36, pp. 727-752.

Rosenthal, S.S. and W.C. Strange (2004), "Evidence on the Nature and Sources of Agglomeration Economies", in J. V. Henderson and J. F. Thisse (eds.), *Handbook of Regional and Urban Economics*, Volume 4, Elsevier.

Simmonds, D.C. (2012), "Developing land-use/transport economic efficiency appraisal" www.davidsimmonds.com/files/LUTEE-paper-for-ETC-v3-111012-1.pdf.

Spatial Economic Research Centre (SERC) (2009), "Strengthening the economic linkages between Leeds and Manchester", Report to Northern Way.

Venables, A.J., J.Laird and H.G. Overman (2015), "Transport Investment and Economic Performance", www.gov.uk/government/publications/transport-investment-and-economic-performance-tiep-report.

Venables, A.J. (2016), "Transport appraisal with land-use change", Working paper, Oxford.

Worsley, T. (2011), "The Evolution of London's Crossrail Scheme and the Development of the Department for Transport's Economic Appraisal Methods", *International Transport Forum Discussion Papers*, No. 2011/27, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/5kg0prk600jk-en

Appendices

Appendix I. Accessibility and productivity

Table A.1 reports elasticities of productivity with respect to economic mass. It is not intended as a definitive statement of parameter values but is indicative of the magnitudes and illustrative of the issues.

In the first block of the table the units of observation are places. Results are reported from survey article (Rosenthal and Strange, 2004), the US (Ciccone and Hall, 1996) and the UK (Rice et al., 2006). Controlling for skill and, in Rice et al. also factoring out differences in occupation structure, researchers find elasticities in the range 0.03-0.04. Rice et al. (2006) also estimate, rather than impose, the rate of spatial attenuation of effects; they tail off sharply beyond about 45 minutes driving time, i.e. are concentrated within travel-to-work distances.

The second block is representative of studies based on firm level data (for the UK, plants from the Annual Respondents' Database). The study by Graham et al. (2009) estimates productivity relationships by sector, using an *ATEM* computed for the location and sector of plants and offices. Elasticities of similar magnitude are derived from this work, and there is considerable heterogeneity, with effects largest in business services. The spatial decay factor was estimated separately for each sector and is largest in service activities, suggesting the incentive for tightly concentrated service clusters. This study provides the elasticities generally used in UK DfT appraisals.

The third block of Table A.1 reports results of estimating wage equations, i.e. looking at the determinants of the earnings of individual workers. The three studies indicated are for data from France, Spain and the UK. Working with individual data makes it possible to address the issue of 'people versus place' by using a fine level of worker level controls – generally skills, age and experience. Once again, elasticities of productivity are of similar size, with those for France (Combes et al., 2008) and Spain (Puga and Roca, 2012) at 0.046 and 0.05 respectively.

The studies in the third block contain two important extensions. One is that while some characteristics of individual workers are observable – their age, skill and experience – their innate ability is not. Bias is introduced if there is a selection effect such that people with high innate ability are more likely to move to large cities. Individual fixed effects control for this, with identification coming from tracking individuals who move. Estimates of this type are presented in the final row of each of these studies and in most cases markedly reduce the productivity elasticity. For Combes et al. (2008) and for Puga and Roca (2012) including these individual effects approximately halves the elasticity, although still leaving it within the range put forward in earlier studies.

The second extension is that the work by SERC (2009) has richer modelling of access to economic mass, constructing *ATEM* measures separately for two different modes of transport (car and rail) and estimating the joint effect of both measures on wages. Consistent with the results above, they find that controlling for the observable characteristics of individuals (and jobs) reduces the effect of access to economic mass (by somewhere between a quarter and a third). The effect of controlling for unobservable characteristics depends on whether one is considering the impact of accessibility by car or by train. For accessibility by car, allowing for sorting on the basis of unobserved characteristics increases the estimated effect (and turns it significant). In contrast, for accessibility by train allowing for sorting decreases the estimated effect by a factor of 3 (larger than the reduction found in studies that do not split by mode).

Table A.1. **Accessibility and productivity**

	Elasticity of productivity with respect to ATEM	**Controls**		**Distance measure / spatial decay**
Unit of observation: Places				
Rosenthal and Strange (2004)	0.05 - 0.11	--- Survey article ---		
Ciccone and Hall (1996)	0.03	Education level		Fixed
Rice, Venables and Pattachini (2006)	0.04	Occupation, skill		Travel time. Estimated.
Unit of observation: Firms				
Graham et al. (2009)	Econ average: 0.043 By sector: Manuf: 0.021 Construction: 0.034 Cons. servs: 0.024 Bus. servs: 0.083	Firm characteristics (e.g. firm age)		Geographical distance. Estimated.
Unit of observation: Workers		Observable (occupation, age, skill, experience) 2	Unobservable (individual fixed effects) 2	
Combes et al (2008)	0.035 0.024	√ X	X √	Fixed
Puga and Roca (2012)	0.046 0.023	√ √	X √	Fixed
SERC (2009)[1] Car	0.08 (not signif.) 0.05 (not signif.) 0.07	X √ √	X X √	GTC car. Fixed, reciprocal
SERC (2009)[1] Rail	0.258 0.17 0.05	X √ √	X X √	GTC rail Fixed, reciprocal

1: SERC 2009, columns 1, 5, 6 table 8 p49.
2: √, control included**:** X, control not included.

Appendix II. Investment and land-use change

Residential development: The change in welfare (to a first order approximation) is the rule of half, plus the quantity change times the average price cost margin,

$$\Delta W = \Delta T\{Q_0 + (Q_1 - Q_0)/2\} + (Q_1 - Q_0)(p_1 - c_1 + p_0 - c_0)/2 \quad (1)$$

The text refers to the two elements as UB and WB respectively. Rearranging,

$$\Delta W = -\Delta t.Q_0 + (Q_1 - Q_0)(p_1 - c_1 + p_0 - c_0 - \Delta t)/2 \quad (2)$$

Land value uplift is the change in price times initial quantity, plus the additional quantity times the new price minus its average opportunity cost (construction cost plus value of land in previous use):

$$\Delta V = (p_1 - p_0)Q_0 + (Q_1 - Q_0)(p_1 - (c_1 + c_0)/2) \tag{3}$$

Land value uplift measures the change in welfare, $\Delta V = \Delta W$ if and only if $p_1 - p_0 = \Delta t$, i.e. the change in price is equal to the user benefit, and not influenced by the change in quantity supplied.

Commercial land-use change: For an iso-elastic demand curve, $x = p^{-\sigma}$, expenditure is $px = p^{1-\sigma}$ and consumer surplus (CS) is the integral of the area below the demand curve and above price, $CS = p^{1-\sigma}/(1-\sigma)$, from which the ratio of consumer surplus to expenditure is 1/(1-σ). For fuller treatment, with many varieties and a spatial structure see Fujita et al. (1999).

Participants list

Pr. Jonas ELIASSON — Chair
Professor Transport Systems Analysis
Royal Institute of Technology
Teknikringen 10
100 44 STOCKHOLM
Sweden

Pr. Mogens FOSGERAU — Rapporteur
Dept. Transport Policy and Behaviour
Technical University of Denmark (DTU)
Wildersgade 60a 2-1
1408 COPENHAGEN K
Denmark

Dr. Marco KOUWENHOVEN — Rapporteur
Significance
Koninginnegracht 23
2514 AB THE HAGUE
Netherlands

Mr. Pim WARFFEMIUS — co-Rapporteur
Researcher - Project Manager, Transportation Economics
Netherlands Institute for Transport Policy Research (KiM)
Ministry of Infrastructure and the Environment
P.O. Box 20901
2500 EX DEN HAAG
Netherlands

Mr. Anthony VENABLES — Rapporteur
BP Professor of Economics
University of Oxford
University of Oxford
Old Indian Institute
34 Broad Street
OX1 3BD OXFORD
United Kingdom

Mr. Glen WEISBROD
President
Economic Development Research Group, Inc
155 Federal St., Suite 155
02110 BOSTON
United States

Rapporteur

Mr. Lars ROGNLIEN
Associate Director
EY
680 George Street
NSW 2000 SYDNEY
Australia

Pr. Michael A.P. TAYLOR
Transport Planning
Barbara Hardy Institute
University of South Australia, Building C
Manwson lakes Campus
GPO Box 2471
5001 ADELAIDE
Australia

Professor William ANDERSON
Director
Cross-Border Institute
University of Windsor
401 Sunset Avenue
N9B 3P4 WINDSOR
Canada

Pr. Juan-Carlos MUÑOZ
Department Transport Engineering and Logistics
Pontificia Universidad Católica de Chile
School of ENgineering
Escuela de Ingeniería.
Vicuña Mackenna 4860, Macul.
7820436 SANTIAGO
Chile

Mr. Per Skrumsager HANSEN
Senior Advisor
Ministry of Transport
Erhvervs- og Analysekontoret
Frederiksholms Kanal 27F
1220 COPENHAGEN K
Denmark

Ms. Ninette PILEGAARD
Senior Researcher
Department of Transport
Technical University of Denmark
Bygningstorvet, Building 116B
KGS. LYNGBY
Denmark

Pr. Hani MAHMASSANI
Civil and Environmental Engineering
Transportation Center
Northwestern University
300 Chambers Hall, 600 Foster
60208 EVANSTON
United States

Professor Pravin VARAIYA
Department of Electrical Engineering and Computer Science
Berkeley University of California
Office 271M Cory Hall
94720 BERKELEY
United States

Mr. Clifford WINSTON
Senior Fellow Economic Studies Program
The Brookings Institution
1775 Massachusetts Avenue, N.W.
WASHINGTON
United States

Dr. André DE PALMA
Ecole normale supérieure de Cachan
61 avenue du président Wilson
94230 CACHAN
France

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 VALLEE CEDEX 2
France

Mr. Jean-Claude PRAGER
Directeur
Société du Grand Paris
62 rue de Mirosmesnil
75008 PARIS
France

Professor Emile QUINET
ENPC - Ecole Nationale des Ponts et Chaussées (Université Paris Val-de-Marne)
28, rue des Saint Pères
75007 PARIS
France

Professor Hironori KATO
University of Tokyo
2-17-58 Higashi Kunitachi
186-0002 TOKYO
Japan

Professor Oded CATS
Department of Transport and Planning
Delft University of Technology Faculty of Civil Engineering and Geosciences Transport and Planning
Stevinweg 1
Bldg. 23, Stevinweg 1
DELFT
Netherlands

Mr. Christopher CAMPBELL
Economic Advisor
Department for Transport (DfT)
Great Minster House
76 Marsham Street
SW1P 4DR LONDON
United Kingdom

Professor Daniel GRAHAM
Department of Civil Engineering
Imperial College London
Skempton Building
South Kensington Campus
SW7 2AZ LONDON
United Kingdom

Mr. Joseph LOWE
Head of Economics Branch
Public Spending Group (2/N1)
HM Treasury
1, Horse Guards Road
SW1A 2HQ LONDON
United Kingdom

Professor Peter MACKIE
Institute for Transport Studies
University of Leeds
36 University Road
LS2 9JT LEEDS
United Kingdom

Professor Roger VICKERMAN
Director
Centre for European, Regional and Transport Economics
University of Kent
Keynes College
CT2 7NP CANTERBURY
United Kingdom

Mr. Tom WORSLEY
Head of Division
Integrated Transport Economics and Appraisal
Department for Transport
Great Minster House, Room 3/06A
76 Marsham Street
SW1P 4DR LONDON
United Kingdom

Mr. Gunnar ISACSSON
Swedish National Road and Transport Research Institute (VTI)
Olaus Magnus väg 35
SE–581 95 LINKÖPING
Sweden

Observer

Ms. Cécile CHÈZE
Doctorante en Sciences Economiques
Laboratoire d'Economie des Transports (LET)
France

OECD

Mr. Kurt VAN DENDER
Head of Tax and the Environment Unit
CTP/TPS/OECD
Annexe Delta 7160
2 rue André-Pascal

ITF Research and Policy

Mr. Stephen PERKINS
Head Research and Policy
2 rue André-Pascal
75016 Paris

Mr. Daniel VERYARD
Transport Analyst

Ms. Jagoda EGELAND
Transport Analyst

Mr. Lorenzo CASULLO
Transport Analyst

Ms. Lorna WILSON
Assistant

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 2016 04 1 P) ISBN 978-92-82-10808-6 – 2017

International Transport Forum

Quantifying the Socio-economic Benefits of Transport

This report sets out several of the recent advances to the quantification and valuation of economic benefits of transport and points to the most promising approaches. The sophistication of modern supply chains and the growing prominence of the services sector have increased the interest of decision-makers in economic benefits beyond those traditionally captured in transport appraisal. Recent advances in assessment methodologies, and their application in decision-making procedures, make it useful to review the state of the art in two areas in particular: reliability benefits and wider economic benefits that flow from transport-related development.

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.itf-oecd.org

9 789282 108086
(742016041P1)
ISBN 978-92-821-0808-6