Chapter 1

Improving transport cost-benefit analysis: Overview and findings

Daniel Veryard[1]

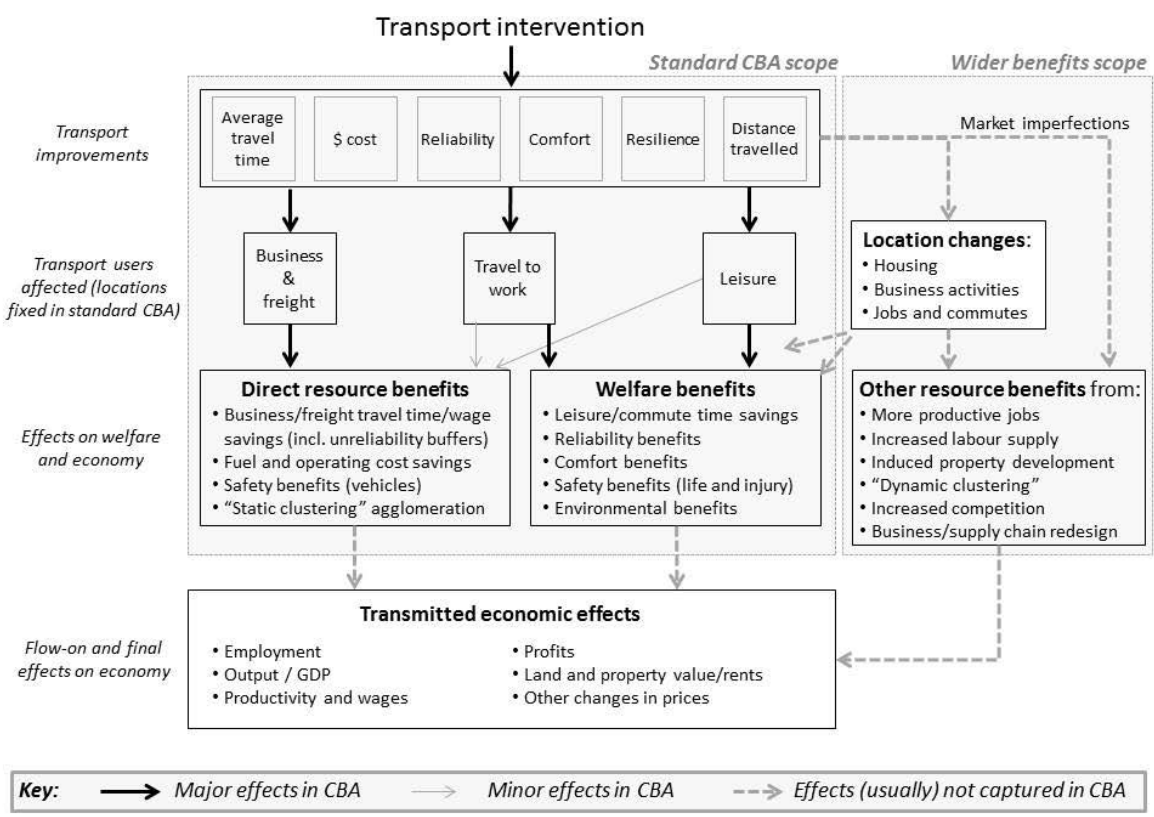
Socio-economic cost-benefit analysis (CBA) is a powerful framework that can be very useful to governments making investment decisions. However the standard application of transport CBA has room for improvement. This chapter describes efforts to improve the quality of transport CBA and its applicability 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

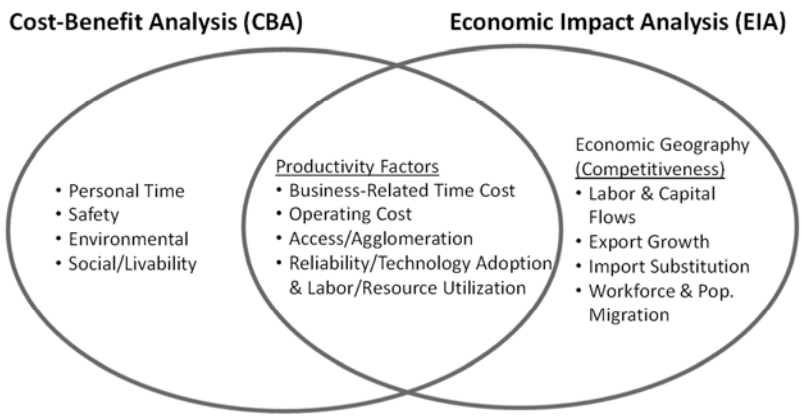
An obvious strategy to overcome the limitations of CBA is to complement it with an alternative appraisal framework. Two broad categories were discussed at the roundtable: economic impact analysis, which measures changes in business revenue, profits, wages and jobs for a specified area; and scoring methods,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

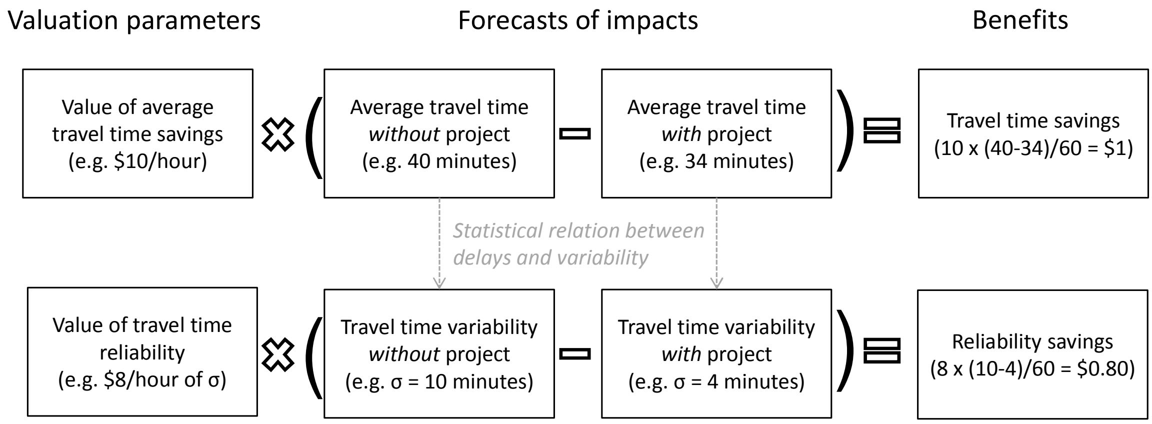
Incorporating reliability benefits in CBA

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).

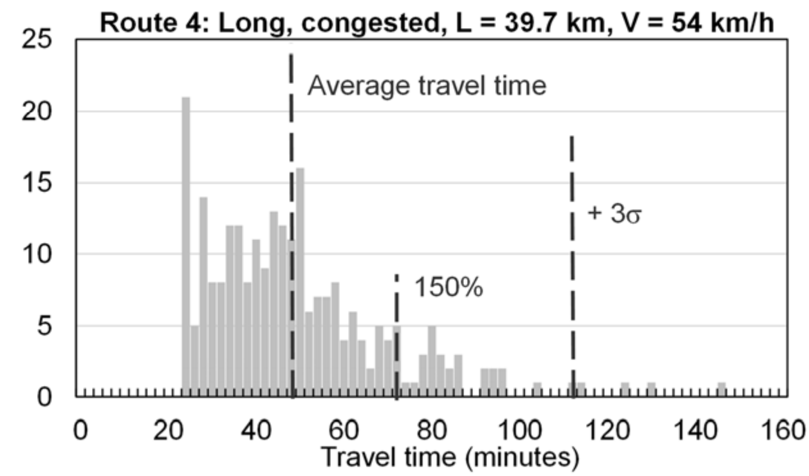
Extending the valuation approach from the case of passenger car travel is likely to be achievable with further research. In the case of public transport service reliability it may be that, as with *average* travel time, travellers may experience disproportionate disutility from *variation* in the components 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.

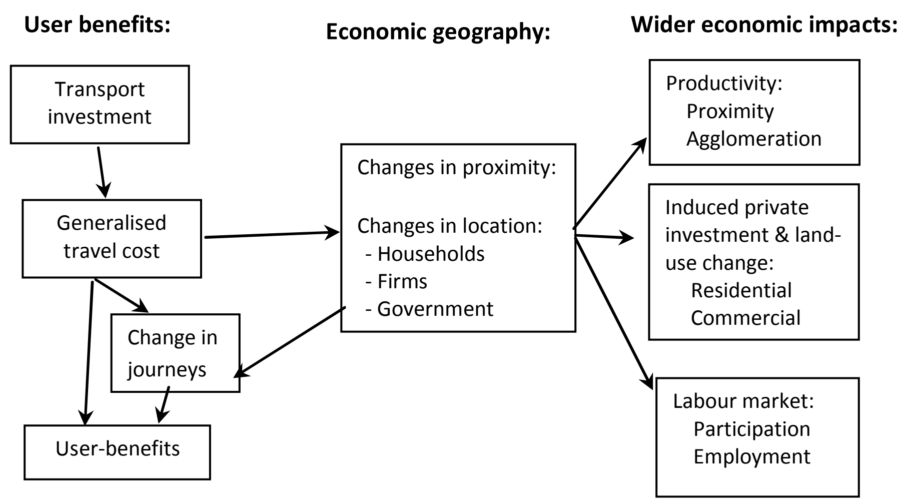
- Sharing experience and data – several roundtable participants voiced an interest in closer international collaboration of researchers to share techniques, data and results. The roundtable itself was seen as a useful step in this direction.

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.

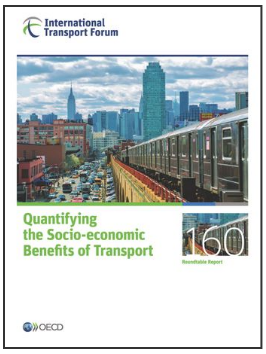
Volpe Transportation Systems Center (2012), Trends in statewide long-range transportation plans: core and emerging topics. https://www.planning.dot.gov/documents/State_plans_report_508_A.PDF (accessed 10 December 2015).

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



From:

Quantifying the Socio-economic Benefits of Transport

Access the complete publication at:
https://doi.org/10.1787/9789282108093-en

Please cite this chapter as:

Veryard, Daniel (2017), “Improving transport cost-benefit analysis: Overview and findings”, in International Transport Forum, *Quantifying the Socio-economic Benefits of Transport*, OECD Publishing, Paris.

DOI: https://doi.org/10.1787/9789282108093-2-en