Department for Transport

Strengthening the Links between Appraisal and Evaluation

July 2016

Contents

Executive summary

Appraisal (carried out ex ante) and evaluation (carried out ex post) are both valuable analytical tools to inform decisions about the effectiveness and value for money of transport policy or investments. Appraisal helps decision makers choose between options before implementation; evaluation can provide evidence of how an intervention has performed once implemented.

Despite the apparent similarities in the ultimate purpose of appraisal and evaluation, both forms of analysis tend to be carried completely independently of each other. This suggests missed opportunities in terms of learning from evaluation to improve appraisal methods; and using information and models from appraisals to maximise what can be learned from the evaluation.

This study was therefore commissioned by the Department for Transport (DfT) to understand, and deliver advice on, how the links between appraisal and evaluation can be strengthened.

Although this report makes recommendations to the DfT, other stakeholders may also find the analysis interesting.

The study's objectives

Links between appraisal and evaluation are currently generally considered relatively weak. This study therefore focuses on understanding the reasons why links are currently weak and exploring how they can be strengthened.

Links between appraisal and evaluation are considered from two perspectives:

- Link 1: exploring how to embed forward planning of evaluation at the appraisal or business case stage; and,

- Link 2: exploring how evaluations can inform appraisals

To inform our work, we have collated evidence from 9 interviews with transport evaluation experts alongside case study analysis of four transport interventions (across local transport, strategic roads and rail). We also provide new analysis to illustrate how the links could be improved.

Findings

We found that there are currently a number of barriers which weaken the links between appraisal and evaluation. These include:

1. Lack of evaluation capability: practitioners of evaluations may not receive sufficient training or be adequately aware of the full range of methods, tools and approaches that are available for carrying out a robust and proportionate evaluation.

2. Appraisal iteration: appraisals evolve over a number of months (or years) and various updates are carried out over time - identifying the appropriate appraisal to compare against outturns in evaluation can be difficult.

3. Complexity of transport models: transport models are very complex which can mean that it is extremely difficult to understand the reasons why observed outcomes differ markedly from anticipated outcomes in the appraisal.

4. Difficulty in accessing and re-using appraisal and modelling work: for an evaluation to explore whether transport outcomes anticipated in the appraisal have been realised, the detailed outputs (and inputs) of the appraisal model need to be stored. This is rarely done in practice. The outsourcing of modelling further raises the risk that such information is not stored and made available for evaluators.

5. Governance: evaluation can be a resource-intensive activity so senior buy-in is needed to release the necessary resources.

Practical challenges we found from our case study analysis corroborate these findings from the experts. In addition, we find that:

- There are challenges identifying appropriate 'without scheme' assumptions to be able to carry out a robust evaluation. For example, external changes are sometimes not accounted for and the 'without scheme' case assumed in the appraisal was in some cases very different to the 'without scheme' case in the evaluation. These issues make it difficult to compare actual outcomes with what was anticipated;

- Evaluation questions sometimes do not map back to the original objectives of the intervention. Hence the evaluation would not be able to uncover whether the objectives have been achieved;

- Wider economic outcomes can be extremely difficult and resource-intensive to evaluate. Therefore they may not be feasible or proportionate to evaluate even where these are a core objective of the intervention;

- Outcomes that were not anticipated at the point of implementation (often referred to as "emergent outcomes") are often not considered; and

- Learnability from evaluation could be limited where the evaluation questions do not link to the strategic objectives of the intervention; and if key outcomes are not robustly evaluated.

Strengthening the links between appraisal and evaluation

Where the barriers identified above can be overcome, there are likely to be benefits for the quality of both appraisals and evaluations.

For link 1: embedding forward planning of evaluation at the appraisal stage, we note that this can have a number of beneficial effects. These include:

- Using a logic model framework for the evaluation plan in line with best practice clarifies the theory of change. This can help to identify the outcomes that should be included in the appraisal;

- It can help ensure consistency in appraisal and evaluation where appropriate, and therefore facilitate the comparability of outturns and forecasts that is needed as part of an evaluation; and,

- It can prompt practitioners to store and collect the data needed for the evaluation.

For link 2: using evaluation to inform appraisal we have identified two particular aspects where evaluation can add value:

- Using evaluations to identify areas where appraisal tools are systematically and materially mis-forecasting outcomes of a scheme and hence warrant further investigation; and

- Using evaluations to inform some of the key assumptions used to forecast outcomes at the appraisal stage.

Given the availability of local and national guidance for practitioners that will support link 1, we focus in particular on actions to strengthen link 2.

Using evaluations to identify areas where appraisal tools are systematically and materially mis-forecasting outcomes of a scheme and hence warrant further investigation

Given the challenges around sharing information between practitioners of appraisals and evaluations, a 'handover pack' could be developed at the time of the appraisal to hand to the evaluator (or stored on a common platform for use by a future evaluator). This would include key information which should be reported routinely. This is important so that forecasts, and the basis on which they were developed, are understood when comparing them against actuals.

This handover information can be used to inform comparative analysis of forecasts with actual observations. In carrying out this analysis, where systematic forecasting errors are identified (beyond tolerable thresholds), the scheme and its modelling suite could be 'red flagged'. Having done this, the complexity of modelling means that a separate piece of more detailed analysis is needed to identify why there are discrepancies between forecasts and outturns.

Although it is possible to learn from individual scheme-level evaluations, there is likely to be greater scope for learning to inform appraisal if a portfolio of interventions is considered. This is because it provides a larger sample size from which to learn, and also allows various comparisons across scheme evaluations to be made to identify common issues.

Using evaluations to inform some of the key assumptions used to forecast outcomes at the appraisal stage

In order to use scheme-level evaluations effectively to inform appraisal assumptions, it is useful to follow a five step process:

- Step 1: identify where there are evidence gaps or a case for updating evidence used within appraisal processes;

- Step 2: consider the extent to which scheme-level evaluation is likely to be the appropriate method for developing that evidence;

- Step 3: develop specific questions on which the evaluations will focus;

- Step 4: design robust evaluation approaches to develop the required evidence; and,

- Step 5: collate and interpret the evidence and translate into appraisal guidance when appropriate.

For the purposes of this report, we tested this approach for real as an exercise with the DfT. This exercise is presented for illustrative purposes to demonstrate how it could work.

Recommendations

We make the following recommendations for DfT and its partners:

Implement appraisal handover packs

1   Design and implement the use of a handover pack as a method of passing forward appraisal information to the evaluator. Linked with this, ensure that externally commissioned modelling analysis is accompanied by logs that store relevant information to feed into the handover pack.

2   Include within the handover pack a requirement to store all sensitivity test runs and scenarios, not just the central case. (This could save the need for additional model runs as part of the evaluation which can be very costly and infeasible).

Develop portfolio approaches for evaluation

3   Consider developing consistent approaches to evaluating portfolios of schemes to identify areas where appraisal tools are systematically and materially mis-forecasting outcomes and hence warrant further investigation. This approach is already followed for major road schemes through POPE and, to some extent for Local Major Schemes.

4   Work alongside Highways England and local partners to explore ways in which the POPE programme and other evaluation series which use meta analyses,

such as the evaluation of Local major schemes, could be adapted to generate further evidence to inform validity tests of appraisal modelling in a systematic way.

Align evaluation work with appraisal evidence needs

5   Explore how best to capture evidence and learn from evaluations to be able to inform future appraisal methods

6   Repeat the process of engagement and consultation with DfT appraisal leads in order to identify, at a strategic level, other areas where evaluations can be used to fill knowledge gaps going forward.

7   Consider aligning DfT-commissioned evaluations (which will be carried out for other main purposes) to ensure that they add value by delivering evidence that will inform WebTAG guidance and models where this is feasible.

8   Commission relevant evaluations and store the findings systematically such that the evidence base can, over time, be translated into appraisal guidance.

1. This study

Introduction

Appraisal (carried out ex ante) and evaluation (carried out ex post) are both valuable analytical tools to inform decisions about the effectiveness and value for money of transport policy or investments. Appraisal helps decision makers choose between options before implementation; evaluation can provide evidence of how an intervention has performed once implemented, therefore allowing valuable learning for future decisions.

Despite the apparent similarities in the ultimate purpose of appraisal and evaluation, both forms of analysis tend to be carried completely independently of each other. This suggests missed opportunities in terms of learning from evaluation to improve appraisal methods; and using information and models from appraisals to maximise what can be learned from the evaluation.

This project was therefore commissioned by the Department for Transport (DfT) to understand, and deliver advice on, how the links between appraisal and evaluation can be strengthened.

Although this report makes recommendations to the DfT, other stakeholders may also find the analysis interesting.

Context

Evidence-based policy making is critical to ensure that investment decisions deliver the best returns on increasingly constrained public funds. Evidence on the value for money and impacts of an intervention are typically generated through two key forms of analysis: ex ante appraisal (as advocated in the HM Treasury Green Book) and ex post evaluation or benefits management (as advocated in the HM Treasury Magenta Book).

These forms of analysis can be summarised as follows:

- Ex ante appraisal: this provides an auditable means to compare the costs and benefits of different policy/investment options before a decision is made. Appraisal delivers evidence on the full range of costs and benefits of each shortlisted policy/investment option over its lifetime, relative to what would be expected to happen without the policy/investment. Appraisals may assess individual projects (such as infrastructure investments or new pricing regimes), policy changes (such as regulatory changes) or packages of measures.

- Ex post evaluation: evaluation is carried out after a policy/investment has been implemented. The evaluation can be used for a range of purposes, including to provide evidence about what outcomes have been observed, for whom, over what timescale and under what conditions. In contrast to the full range of costs and benefits considered within appraisals, evaluations are often much more targeted because best practice advocates that the analysis focuses on specific questions. These may relate to the way the policy/investment was delivered (a 'process' evaluation), the outcomes observed and how they are distributed across population groups or geographies (an 'impact' evaluation) or the actual costs and benefits (an 'economic' evaluation).

Clearly both forms of analysis have an important role to play in generating evidence to inform decisions. Yet, to date, greater emphasis has been placed on developing appraisal methods and tools to inform investment decisions, than on understanding what has been observed post-implementation.

This has been changing over recent years as practitioners increasingly recognise the value of being able to learn about what works and under what conditions. For example, scrutiny bodies such as the Public Accounts Committee and the National Audit Office[1] and leading academics regularly remind government of the need for ex post evaluation evidence. The value of such evidence was also noted in a recent OECD International Transport Forum roundtable focusing on ex-post assessment of transport interventions[2].

The DfT is building its evaluation capability and is currently implementing its monitoring and evaluation strategy[3], led by its new Evaluation Centre of Excellence. As part of its work, DfT has identified a need to explore how appraisal and evaluation could better inform each other. This project therefore focuses on investigating how the links between ex ante transport appraisal and ex post transport evaluation could be strengthened.

Objectives of this study

Links between appraisal and evaluation are currently generally considered relatively weak. This study therefore focuses on understanding the reasons why links are currently weak and exploring how they can be strengthened. In particular, we consider the links between appraisal and evaluation from two perspectives:

- Link 1: exploring how to embed forward planning of evaluation at the appraisal or business case stage; and,

- Link 2: exploring how evaluations can inform appraisals.

These links are shown in **Figure 1**.

---

[1] National Audit Office (2013) http://www.nao.org.uk/wp-content/uploads/2013/12/10331-001-Evaluation-in-government_NEW.pdf
[2] International Transport Forum (2014) http://www.internationaltransportforum.org/jtrc/DiscussionPapers/DP201419.pdf
[3] DfT (2013) Monitoring and Evaluation Strategy https://www.gov.uk/government/publications/monitoring-and-evaluation-strategy

Figure 1. Links between appraisal and evaluation
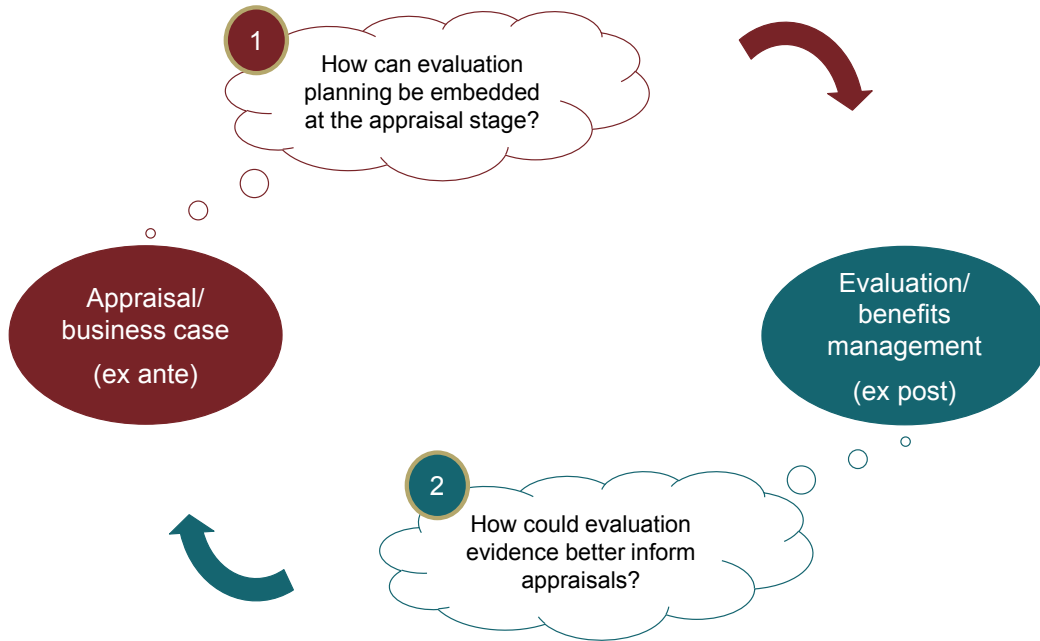
The first link relates to the best practice approach to policy evaluation: planning for evaluation must begin at the appraisal stage i.e. before implementation. This helps in a number of ways. For example, it allows the practitioner to identify what it is that will be the focus of the evaluation and therefore identify the data to be collected to inform a 'counterfactual' and to measure outcomes. Also, it helps to ensure that information generated at the appraisal stage can be stored appropriately to hand to the future evaluator to be validated as part of the evaluation.

The second link relates to the potential for the evaluation to inform appraisals. There are many questions that evaluations could be designed to address, so informing appraisals is just one and is the focus of this study. We consider this potential from two perspectives:

- Using evaluations to identify areas where appraisal tools are systematically and materially mis-forecasting outcomes of a scheme and hence warrant further investigation; and

- Using evaluations to inform some of the key assumptions used to forecast outcomes at the appraisal stage.

Structure of this report

This study is structured as follows:

- Chapter 2 describes the 'logic model' framework that underpins evaluation and appraisals. This describes the theory of change of an intervention;

- Chapter 3 assesses the strength of the current links between appraisal and evaluation and identifies barriers to those links being stronger; and,

- Chapter 4 focuses on using evaluation evidence to inform appraisals. This chapter concludes with some recommendations for DfT to consider.

2. The logic model framework

Introduction

To aid the practitioner in ensuring evaluations are carried out in a structured and evidence-based way, a logic model framework is advocated as best practice in the HM Treasury Magenta Book. This chapter describes the logic model framework and how it can be used to underpin evaluation analysis. This is described in further detail in the HM Treasury Magenta Book[4] and in "Logic Mapping: Hints and Tips" prepared for the DfT in 2010[5].

The logic model

Logic modelling is recommended as best practice in the HM Treasury Magenta Book (the government's guidance on monitoring and ex-post evaluation). It visually describes the 'theory of change' of an intervention by walking the reader through the process of how inputs (such as investment of capital, person time or other resources) are deployed and transformed to deliver outputs (physical observable outputs such as transport infrastructure or new technology). These are then expected to deliver outcomes for transport users (such as improvements in overcrowding or congestion) and for the wider economy, society or environment (for example, by unlocking business investment; creating opportunities for labour mobility; or, lowering emissions). These can then facilitate longer term impacts such as economic growth or housing developments.

The theory of change should of course also note the potential for adverse outcomes to arise and for unforeseen outcomes to be observed.

Figure 2 presents an illustrative logic model for a hypothetical transport infrastructure investment. The framework can be applied to any transport intervention. Key to the framework is the way it encourages the practitioner and policy-maker to articulate the vision for the outcomes and impacts of the intervention and the intervening outputs and activities that are needed to deliver them, including the level and nature of required inputs.

The stylised logic model in **Figure 2** shows that the outcomes and impacts can be considered in two categories. The first category refers to the direct outcomes and impacts for transport users. These refer to outcomes that accrue to those who travel on the transport system - some may be positive (for

---

[4] HM Treasury (2011) The Magenta Book
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/220542/magenta_book_combined.pdf
[5] Tavistock Institute (2010) Logic Mapping: Hints and Tips
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/3817/logicmapping.pdf

example if an intervention lowers overcrowding) but some may also be negative (such as if an intervention shifts activity from one part of the transport system so that some transport users gain at the expense of others).

The second category refers to the outcomes and impacts that accrue to the wider economy, society or environment. These would generally result from the outcomes for transport users, though importantly, in some cases they can accrue in anticipation of those transport outcomes and therefore be observed before the intervention has actually been implemented. This might be the case for example if developers speculate on the purchase of land in the expectation that land values may increase after the intervention is implemented. Importantly, these wider benefits are not necessarily additional to the direct transport user benefits, though in some cases - such as with the agglomeration benefits - they are genuinely additional.

Figure 2. Illustrative logic model
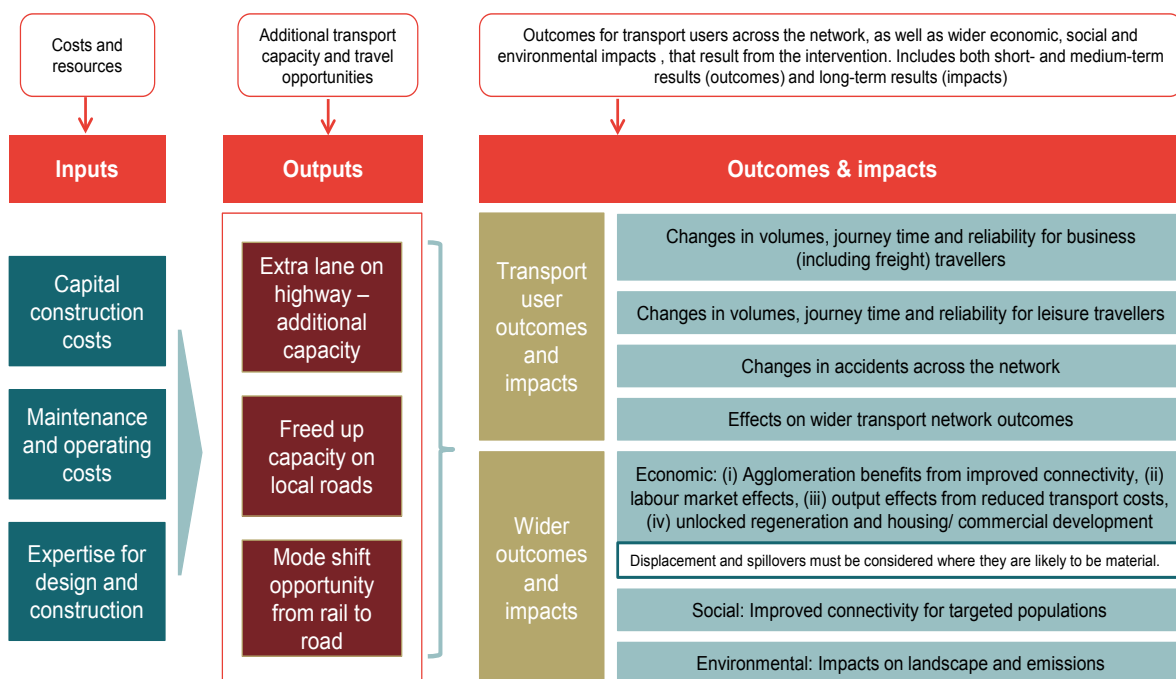
The logic model provides a useful framework to underpin the evaluation. It can assist the practitioner in identifying what questions they wish to ask of the evaluation. For example, there may be particular objectives for which evidence on whether they have been realised would be useful. The logic model also helps to identify outcomes that should be monitored and hence what data may help them do that. The practitioner should also consider the time period over which they would like to evaluate the intervention as some outcomes will be observed immediately, others may take several years, and some outcomes may arise before implementation (such as land value changes in anticipation of a future investment).

Importantly, the level of detail captured in the logic model should vary in line with proportionality considerations. Greater detail is likely to be warranted for interventions which are of a larger scale (in terms of the level of investment and

geographical reach) and for which outcomes are likely to be more diverse or complex.

This logic model framework underpins both appraisals and evaluations and therefore underpins the rest of this report. It should be noted that the principles of the logic model are consistent with those underpinning current DfT appraisal guidance for Technical Project Managers[6] and Senior Responsible Officers[7].

The next chapter explores the scoping evidence on the current links between appraisal and evaluation.

---

[6] See https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/427078/webtag-tag-guidance-for-the-technical-project-manager.pdf

[7] See https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/427074/tag-guidance-for-senior-responsible-officer.pdf

3. Understanding the current links between appraisal and evaluation

Introduction

To assess how the links between appraisal and evaluation could be strengthened, we need to first understand the strength of the current links; the barriers that hinder them from being stronger; and, the facilitators that would help overcome those barriers.

To explore these issues, this chapter describes the evidence we collated from two forms of investigation:

- First, we carried out nine interviews with evaluation experts. These were used to draw out experiences and observations having worked directly (as a practitioner) on transport evaluations.

- Second, we carried out four case studies of interventions for which documentation was available covering both the ex ante appraisal and 1-year after evaluation (one case study had only an appraisal available). The purpose of the case studies was to explore the extent to which the evaluation was appropriately planned as part of the business case (link 1); and the extent to which the evaluation could be used to inform future appraisals (link 2).

The findings of our scoping work are described below.

1. Learning from the experts

This part of our scoping work involved carrying out nine semi-structured interviews with academic evaluation experts; evaluation practitioners; and appraisal and evaluation policy leads. The purpose of these interviews was to understand their thoughts and experiences in relation to the links between appraisal and evaluation; their perception of the likely barriers hindering stronger links; and the facilitators that would help improve the links.

We have carried out a thematic analysis of the findings from the interviews. Five key themes emerged from this work:

- 1. Evaluation capability
- 2. Iterative development of transport appraisals

- 3. Complexity of transport models
- 4. Difficulty of accessing and re-using appraisal and modelling work
- 5. Governance

We cover each below.

1. Evaluation capability

Evaluation capability refers to the skills, knowledge and experience needed to plan, carry out and report a robust evaluation. Unless an evaluation is robust, it cannot credibly inform appraisals. We consider various aspects of evaluation capability below.

Knowledge, training and skills

Experts noted that although there are many courses and qualifications available for transport modellers and transport appraisal, evaluation is typically not covered in the same way. Practitioners therefore tend to lack evaluation-specific training.

A related observation, and in part a consequence of the above, is that appraisals and evaluations are carried out by separate groups of practitioners. This separation creates an inherent barrier to knowledge being passed on from the appraiser to the evaluator, and also means that the appraisal practitioner has little knowledge or understanding about what might be useful for the evaluator in the future. Therefore, records and information are regularly not logged in a form or location that is easily accessible at the time of the evaluation.

Delivering high quality evaluations requires very specific knowledge and experience, such as in relation to determining an appropriate counterfactual; interpreting the outcomes; monitoring unexpected outcomes etc. However, a lack of training in evaluation methods and how to set up the evaluation robustly could increase the risk that some evaluations are actually misleading. For example, an intervention which is intended to improve bus reliability on a particular route could be evaluated by looking at outcomes for that route only. However, improvements to that route could have been achieved by shifting congestion onto other routes. Therefore unless the evaluation scope has been defined sufficiently widely, findings could be flawed. This would risk poor policy decisions being made.

Guidance material

Most of the experts we consulted for this study referred to the potential for guidance materials to pose a challenge for practitioners' ability to deliver robust evaluations if the methods are not practical.

An example referred to is the evaluation guidance documents published by DfT, such as the "Local Sustainable Transport Fund Monitoring and Evaluation

Framework"[8]. These are intended to set out best practice requirements for successful bidders for central government funding (or other grants) to monitor and evaluate the outcomes of their funded interventions. However, although useful in setting out best practice, practitioners often find it difficult to implement this in practice because of the practical constraints on the ground. These can include capability, resources and others - for example, some data may not be available (such as if it is owned or collected by private entities) or may be disproportionately costly to collect[9].

Local practitioners noted that they have in some cases shifted towards benefits management as a way of monitoring the outcomes of interventions. One of the main drivers has been the steer on major projects from the IPA[10] and NAO. For example, DfT, Transport for London (TfL) and HS2 Ltd, among other bodies, have established internal benefits management functions. Benefits management is a structured approach to the identification, quantification, analysis, planning, tracking realisation and optimisation of benefits[11], which are the measurable improvements that a project seeks to deliver. This approach differs from evaluation in that it is a project management discipline, rather than an analytical one. Its main focus is on managing and tracking the delivery of anticipated benefits of a project. Benefits management and evaluation have a shared concern to understand the outcomes of a project (whether benefits or dis-benefits) while evaluation often also seeks to learn what worked well and what didn't, and under what conditions. Benefits management also centres on monitoring delivery of the project itself while evaluation tends to also look at wider evidence which may not be generated through project delivery.

2. Iterative development of transport appraisals

Experts also referred to the iterative process used to develop transport appraisals as being a barrier to linking appraisal and evaluation. By their nature, it can take several years from the time of the first options appraisal to the point of construction. Over that time, a scheme is appraised many times as it evolves and is finessed. It is therefore sometimes difficult to single out the final appraisal carried out prior to the start of works, against which the evaluation should be carried out.

3. Complexity of transport models

Transport user benefits sit at the heart of a transport appraisal and are assessed using (often complex) transport models. The models use a series of inputs and assumptions to generate their forecasts and, in turn, estimate the benefits and costs, relative to a 'do minimum' (without intervention) case.

The range of assumptions involved in transport modelling is wide so it is important that assumptions are evidence-based and realistic. Evaluation can

---

[8] Published by the Department for Transport in 2012: available at:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/35975/lstf-monitoring-evaluation-framework.pdf
[9] We understand DfT is currently reviewing its guidance for local schemes to consider how this can best be revised.
[10] The Major Projects Authority and Infrastructure UK merged in January 2016 to become the Infrastructure and Projects Authority
[11] This definition is taken from Jenner / APMG (2014) Managing Benefits, TSO, London.

therefore play a valuable role in corroborating the validity of key assumptions, although the level of work undertaken will be influenced by proportionality considerations, given the variance in complexity of models available.

The box below outlines the core approach to transport modelling.

Transport modelling

Transport models are developed to forecast the costs and benefits of an intervention relative to the case in which the intervention is not delivered. They are founded on the economic concepts of demand (covering the number of transport users and their behaviour in response to infrastructure provision) and supply (indicated by the cost of using the infrastructure). Underlying the forecasts are assumptions about:

- the number of potential users (population and economic growth are among the key drivers of travel demand);

- the behaviour of the users (such as the responsiveness - or elasticity - of their demand for travel with respect to the cost of making trips);

- the cost of using the infrastructure, which is related to the infrastructure provision;

Transport models forecast overall travel demand, which is driven by factors such as population growth, economic growth and travel costs; and how that demand is distributed across the transport network of interest[12]. This is done by drawing on our understanding of the value to different travellers of changes in travel time (and of other elements of the disutility of travel) and through our observations of behaviour and responsiveness to those changes. These behavioural responses allow us to then estimate changes in demand by mode, destination and route, and hence traffic volumes, congestion/crowding etc. Traffic models typically start from data on existing traffic flows and then predict how flows will change over time in the do-minimum (without a scheme) and do-something (with a scheme) case.

4. Difficulty of accessing and re-using appraisal and modelling work

Our work has identified a number of barriers which make it difficult for evaluators to access and re-use material used at the appraisal stage. These include:

- Staff continuity and outsourcing: where evaluations are being used to explore the extent to which the appraisal forecasts were relatively accurate, it is often helpful to be able to speak with the appraisal practitioner. However, given appraisals and associated modelling are often carried out by external parties who are sub-contracted to the project sponsors, it may not be possible to communicate with the appraisal lead if they have moved

---

[12] It is also worth noting that some models do not have demand responses, for example, if they use fixed trip matrices. Using evaluation to validate these models would therefore need to focus on those aspects of travel behaviour for which assumptions are made.

on to a different post. In addition, unless written into their contract, the appraisal consultants may have very little incentive to log or store all material and documentation that would be useful for the evaluation leads in the future. This loss of knowledge and documentation therefore represents a lost opportunity for learning.

- Data is costly to collect: a core principle guiding evaluation activity is that it must be proportionate to the scale of the intervention. Many of the experts we consulted for this study noted the high costs of collecting data for bespoke evaluation purposes and therefore their need to rely on operational data as far as possible. However, this increases the risk that the data is not adequately comprehensive or is not complete over time.

- Cost of model runs: a further issue widely noted by the experts we spoke to relates to the costs of transport model runs. Many of these models are bespoke to the individual intervention in question and are often operated by external consultants. Such models are typically very data intensive and absorb a significant resource in their development. Commissioning model runs to help validate their assumptions and outputs is rarely done, in part because of the significant costs associated with such activity.

- Guidance changes: central appraisal guidance changes over time. Schemes appraised using previous versions of guidance may therefore rely on very different underlying input data and assumptions to those carried out more recently.

- Developments in modelling and appraisal: development of software, modelling capability and new research[13] mean that models and their assumptions evolve over time. This may make direct comparisons (between appraisal and evaluation) difficult.

- Information volume: the sheer volume of information that would need to be stored (e.g. models and accompanying documents) to enable comparisons between the appraisal and evaluation stages is a barrier in itself.

5. Governance

One of the key facilitators to undertaking meaningful and robust evaluations that was noted by the experts we spoke to is having appropriate governance and senior level buy-in and support for evaluation or benefits management activity. Various different issues were raised in relation to governance and are described below.

Resources

A key barrier to carrying out evaluations and benefits management in practice is the level of resource required. Particularly in a world of increasingly constrained resources, unless senior management support the need to carry out evaluation and benefits management activity and actively use its findings to

---

[13] For example recent research into the value of time savings will (if implemented) require the updating of modelling and appraisal values.

inform future decision-making, it is not likely that sufficient resources will be made available relative to other priorities. Project sponsors and developers are not typically resourced to follow through their activities to post-opening evaluation[14] therefore separate resources are often needed.

The experts we spoke to suggested that the extent to which senior management teams are prepared to invest resources in evaluation activity can in some cases be influenced by the pressure they are under to demonstrate effective use of their investment capital. For example, Highways England is required to demonstrate that public funds are invested in a way which delivers value for money for the taxpayer. Its Post Opening Project Evaluation (POPE) programme run by a dedicated team is a key way to achieve this.

A benefit of having the POPE programme is that Highways England is in a position to ensure a consistent approach is used across its evaluation of major programmes. This includes both individual project-level evaluations, along with meta-analyses that consider tens of projects at the same time to learn lessons.

Devolved decision-making

Devolution of transport budgets in some areas of England means that some local areas have the responsibility for making decisions about how to allocate their available funds. To inform such decisions, some of the experts we spoke to for this study indicated that senior managers could therefore have a greater interest in evaluation and benefits management so that they can demonstrate the extent to which they are delivering value for money in their local areas. Such evidence is able to support the communication of funding decisions.

2. Case study analysis

The second form of scoping work we carried out to understand the current links between appraisals and evaluation was case study analysis.

For this we worked with the DfT and Highways England to identify four case study interventions across three major transport project spending areas: local transport, rail and strategic roads. Although we recognise this is a small number of case studies, there are features about these case studies that makes them relevant more widely, as explained below.

The purpose of the case studies was to allow us to learn about the extent to which links between appraisal and evaluation can be observed in practice.

The four case studies were:

1. Rochdale Interchange

2. Doncaster White Rose Way

---

[14] Obviously, for all transport investment, the operating performance of the transport infrastructure would be monitored, but this is likely to focus on operational efficiency and reliability, rather than taking a strategic assessment of overall benefits of the intervention.

3. Crossrail 2

4. A46 Newark to Widmerpool

Further detail on the process of selecting these case studies and our detailed analysis is in the annex.

Key findings from our analysis of the case studies

Our analysis of the case studies involved reviewing available documentation such as the business cases, evaluation plans, one year-evaluations and also interviews with the evaluation or business case leads.

Our case study analysis highlighted some cross-cutting themes, as summarised in Figure 3.

Figure 3. Case study analysis: common themes

Do-minimum / Counterfactuals	▪ Methods vary and they often rely on 'before versus after', sometimes without controls ▪ Do-min in the appraisal appears not to be directly drawn upon for the evaluations
Questions asked of the evaluation	▪ Questions do not always relate back to the (strategic) objectives of the intervention – the impact of this depends what we want to learn from the evaluation
Wider economy impacts	▪ Even if the objectives of the intervention should contribute to wider economic outcomes (e.g. local transport case studies), these effects are often difficult to appraise or evaluate, particularly for smaller schemes.
Ease of re-running the models	▪ The complexity of transport modelling is such that re-running the modelling is costly and not typically done – it would require highly specialised capability to do so. Plus models are often run by consultants
Modelling complexity	▪ The modelling is often comprised of a suite of interacting models therefore isolating the drivers of difference between forecasts and actuals is likely to be highly challenging. Intermediate outcome comparisons are more useful
Emergent benefits	▪ It is not clear how the evaluations pick up on emergent benefits – they are often not referred to or logged
Learnability	▪ The learnability from scheme evaluations is limited if evaluation questions do not link back to the strategic objectives of the intervention; if validation of outcomes anticipated in the appraisal is not possible; and if key outcomes are not evaluated. Examining portfolios of interventions through meta analyses (such as POPE or meta-analysis of local major scheme evaluations could provide greater insight on specific evaluation questions.

As shown above, several of our findings from the case study analysis corroborate what we found in the expert interviews. Firstly, the case studies highlighted the challenges for practitioners in developing the counterfactual. Often 'before versus after' methods are used but in some cases do not control

for confounding factors. This risks over or under estimating the impact of the intervention. We also find that the 'without intervention' case in the appraisal differs from the 'without intervention' case in the evaluation - therefore limiting the value of comparing the results.

Secondly, the complexity of the transport modelling means, as noted above, that it is difficult to understand reasons for the differences between forecasts and actuals. Thirdly, related to this, modelling complexity also makes it difficult and costly to re-run transport models.

The case studies also identified some other important issues as described below:

- Evaluation questions: We noted that although the one-year evaluations specified evaluation questions, they often did not relate directly back to the objectives of the intervention. This means that it is possible the evaluation would not uncover whether the objectives have been achieved.

- Wider economic impacts: These often appear as objectives of the case studies but often are not evaluated. This is because the evaluation of wider economic outcomes, such as reducing levels of deprivation or facilitating housing growth, is relatively data and resource intensive. It can also present many methodological challenges (such as defining the counterfactual etc.). We note however that qualitative interviews with stakeholders about wider economic outcomes were planned in some cases which could provide some insight (for example, whether businesses relocated because of an intervention) but are not likely to indicate the scale of wider economic impacts realised.

- 'Emergent outcomes': We found that in the case studies we investigated, no provision was made to account for such outcomes. This is likely to be in large part due to the difficulties in identifying such outcomes. For example, to identify emergent benefits, some form of evidence stream is needed that can pick up issues that may not be subject to active monitoring. Issues logs in project/benefits management and stakeholder research are among ways that this could be done.

- 'Learnability': We found that there can be constraints at the individual scheme level that can hinder learnability, not least the need for evaluation activity to be proportionate given the costs of data collection. However, meta analyses of a larger portfolio of programmes offers the potential to learn about specific evaluation questions by looking across a number of similar scheme evaluations. This approach is used for the POPE programme[15] and a recent meta-analysis of local transport major schemes[16]. By focusing the analysis on specified evaluation questions, a more in-depth set of learning can emerge[17].

[15] http://assets.highways.gov.uk/our-road-network/pope/major-schemes/pope-meta-2013-final-report.pdf
[16] https://www.gov.uk/government/publications/local-major-schemes-meta-evaluation
[17] This approach was in fact taken by the Highways Agency several years ago to improve its appraisal modelling of accident and collision outcomes.

3. What this means for the links between appraisal and evaluation

Our analysis in this chapter has confirmed that the links between appraisal and evaluation are currently hindered by a number of barriers.

Where these can be overcome, there are likely to be benefits for both the appraisal and the evaluation.

Likely benefits from overcoming barriers in link 1: embedding evaluation at the appraisal stage

Based on the evidence in this section, planning an evaluation at the appraisal stage would be beneficial in three key ways.

First, it can help clarify thinking around the anticipated outcomes that should be included in the appraisal. Constructing a logic model (as described in chapter 2) which sets out the causal chain of events from scheme inputs to scheme outcomes, is an effective way to visualise the theory of change underpinning the appraisal. This enables practitioners to think through and critically evaluate each link in the chain in a systematic manner, ensuring that the benefits modelled in the appraisal are well-founded and evidence-based. The logic model can then also be used to underpin the evaluation.

Second, it can help ensure consistency in appraisal and evaluation where appropriate, and therefore facilitate the comparability of outturns and forecasts that is needed as part of an evaluation. For example, early planning could help to align the 'without intervention' case in the appraisal (often referred to as the do-minimum) and the 'without intervention' case in the evaluation (often referred to as the counterfactual).

Third, it can prompt practitioners to store and collect the data needed for the evaluation, where a decision is made that such an evaluation should be taken forward. It is useful to store the key assumptions underlying the appraisal modelling (including different scenarios that have been modelled, such as those using low, medium and high economic growth rate assumptions), so that future evaluators can explore the causes of any discrepancies between outturns and forecasts. Appraisal assumptions around the timing of benefits can also be used to inform when evaluations should be carried out and hence the period of data collection. Further, by selecting indicators for the evaluation and assessing the existing data on these indicators, practitioners can identify gaps in the data that require bespoke data collection, including baseline data. It is crucial that this is done at the appraisal stage because any additional data must be collected before the scheme opens, and in some cases before the scheme is announced due to potential 'anticipation' effects - for instance, land values near a major transport scheme will often respond immediately after its announcement.

Likely benefits from overcoming barriers in link 2: using evaluation to inform appraisal

Enabling evaluation to be used to better inform appraisal is likely to lead to a range of potential benefits. Two particular benefits that have been alluded to in this chapter are the following:

- Using evaluations to identify areas where appraisal tools are systematically and materially mis-forecasting outcomes of a scheme and hence warrant further investigation; and

- Using evaluations to inform some of the key assumptions used to forecast outcomes at the appraisal stage.

We are aware that there are various guidance documents available that provide support for practitioners to embed evaluation planning at the appraisal stage. Some of these are currently being updated and developed further with practitioners. We do not therefore seek to provide any further guidance here, but would note the importance of ensuring guidance developed in the future remains pragmatic and supports proportionate evaluation activity.

There remain notable gaps, however, in guidance relating to how evaluation can inform appraisal. For this reason, we focus the remainder of this report on the latter, though fully recognising the importance of the former.

4. Using evaluation evidence to inform appraisal

Introduction

This chapter considers how evaluations can be used to inform appraisal (link 2). In particular, we consider two aspects:

- Using evaluations to identify areas where appraisal tools are systematically and materially mis-forecasting outcomes of a scheme and hence warrant further investigation; and

- Using evaluations to inform some of the key assumptions used to forecast outcomes at the appraisal stage.

Using evaluations to identify areas where appraisal tools warrant further investigation

Introduction

Given the significance of transport modelling at the appraisal stage, it is important that modelling outputs can be checked to ensure that they forecast with a reasonable degree of accuracy. Clearly models will never predict the future with 100% accuracy and some degree of uncertainty should be allowed for. However, systematic and material differences between forecasts and outturns (subject to certain tolerance levels) could potentially indicate weaknesses in the structure of the model or inadequate modelling assumptions. This would mean that decisions about which schemes are prioritised could be based on flawed information and errors could be repeated over time.

Using evaluations to compare actual outcomes with forecast outcomes can enable learning for future appraisals and could help minimise the risk of repeating forecasting errors over time.

Such an approach is already in place for Highways England major projects[18] (and to some extent Local Major Schemes[19]) which are evaluated at regular

---

[18] The current POPE methodology is specified in "Post Opening Project Evaluation (POPE) of Major Schemes Methodology Note" September 2014
[19] A Meta study[19] was also produced by DfT for Local Major Schemes (LMS) although this is less able to shed light on the reasons for discrepancies between forecasts and outturns due to the fact that there is considerable variation in evaluation practice and quality across these schemes.

intervals. In the case of highway interventions, a clearly specified methodology is applied to evaluate each major project (over £10m), and a meta-study is produced every two years. This approach applied by Highways England is known as Post Opening Project Evaluation (POPE) and is summarised in the box below.

A targeted meta-analysis approach could potentially be applied to other areas of transport investment to enhance learning from evaluation to inform appraisals. Such an approach would seek to determine whether discrepancies arise because of exogenous inputs, modelled capacity, behavioural assumptions or modelling methods.

Given the complexity of transport models and the difficulties in accessing and re-running them (as identified in the previous chapter), evaluations can serve a valuable role in being able to 'red flag' where there is a case for separate and more detailed investigation into the appraisal and modelling. That is, they can be used to identify where there are material differences between forecasts and actuals and that further attention is justified.

The remainder of this chapter explains how to use evaluations to 'red flag' issues in the modelling.

Post Opening Project Evaluation (POPE) Meta study

All Highways England major projects (with value greater than £10m) are evaluated one year after scheme opening and five years after scheme opening. Individual scheme evaluations consider the performance of the intervention against each of the objectives of the Appraisal Summary Table (environment, safety, economy, accessibility and integration) as well as scheme specific objectives. Evaluations also compare forecasts of transport user benefits from the appraisal with outturns (what has been observed in practice) for each of the key objectives and comment is provided on any material discrepancies that emerge.

The findings from individual schemes are then pooled together and analysed in a meta study produced every two years. This is used to identify common themes in the data that can then be used to examine the relationship between forecast outcomes and outturns.

The POPE Meta report devotes a chapter on forecasting accuracy which looks at (among other aspects):

- Traffic volumes;
- Journey time savings;
- Safety; and
- Scheme costs.

POPE's Meta study acknowledges that certain deviations from forecast are inevitable and only looks at differences over a certain tolerance level. For example, for traffic volumes a tolerance level of +/-15% is used so that schemes for which forecasts are within +/-15% of outturn are considered 'accurate' and only those where the discrepancy is greater than this are examined in further detail to understand the potential source of the discrepancy.

The latest published POPE Meta study was carried out in 2013 and looked at 75 schemes which opened between 2002 and 2010. The report found that around 65% of scheme appraisals accurately forecast traffic volumes (within the tolerance level) but that there is large variation between schemes. Predictions were less accurate for safety outcomes and scheme costs - for each of these aspects, around half of schemes were within the +/-15% tolerance level. The POPE Meta study concludes that the main reasons for material discrepancies are to do with strategic and local routing assumptions, land use assumptions, economic growth assumptions, other schemes and modelling errors.

Overcoming information barriers by developing a handover pack

We noted in chapter 3 that one of the barriers to better links between appraisal and evaluation is that separate teams are typically responsible for each type of analysis. This limits the flow of information from appraiser to evaluator. In addition, we identified that appraisal modelling is often outsourced which makes storing detailed inputs and assumptions from the appraisal modelling less likely.

To overcome these barriers, a 'handover pack' could be developed at the time of the appraisal to hand to the evaluator (or stored on a common platform for use by a future evaluator). This would include key information which should be reported routinely. This is important so that forecasts, and the basis on which they were developed, are understood when comparing them against actuals.

A key information checklist could be developed covering:

- Scheme information: key assumptions coded into the model on capacity, geographical coverage, interchanges, services, route choice options and changes in these over time;

- Exogenous assumptions (and the source data): land use planning assumptions; oil prices or fuel costs; fares, population and demographic change; GDP, employment and economic activity; values of time; and, other local policies or interventions assumed. Where land use transport interaction models are used, there are also likely to be exogenous inputs relating to commercial floor space availability and trade patterns, for example;

- Behavioural response assumptions: key elasticity parameters that feed into the model. In particular, this will include the elasticity of travel demand with respect to changes in generalised costs and the components of generalised costs as specified in the model (for example, in rail schemes the treatment of crowding in the model). Where land use transport interaction models are used then there are likely to be a much wider range of behavioural assumptions relating to the responsiveness of business investment or location decisions to changes in accessibility.

- Model information: model version, software used, updates and validation notes, storage location, dates run and modelled years, pricing base year, appraisal guidance version used and key assumptions and parameters.

- The forecasting report: this is a modelling report that is expected to be provided with a transport project business case; it includes forecasting assumptions on the supply and demand sides and their appropriateness, diagnostics of outturn results, the treatment of uncertainty and sensitivity testing[20].

The exact list in the checklist would vary according to the nature of the appraisal and modelling work that has been carried out. An assessment of what would be proportionate to include in the handover pack would also need to be made.

---

[20] See WebTAG Guidance for the Technical Project Manager for details.

The handover pack would be a good practice step which could be adopted in different ways for different types of scheme.

For rail franchising, a handover pack already exists as the process moves from a competition stage to the franchise contract management stage. So the contract manager for each franchise will receive a comprehensive "pack" setting out what to expect. Part of this will be the committed obligations (i.e. aspects of the service the winning bidder said that they would deliver in their bid) which will have been assessed in the bid process.

A suggested handover pack

As a starting point, a simple handover pack can be developed and used at the appraisal stage. This could be completed by the appraisal team and would then be handed over to evaluators. It would include actual data where feasible but also embedded documents, key sources used (e.g. WebTAG version), links to folders where input/assumption sheets are saved, folders where outputs are stored etc.

A suggested handover pack information checklist which sets out the type of information that could be gathered at the appraisal stage is shown in **Figure 4**. This is based on our assessment of the key information that would be required by evaluators to enable learning to inform appraisals. The checklist is a starting point which DfT and scheme promoters can build on and develop over time.

Figure 4. Information checklist for appraisers

Scheme information	Please record information below
Scheme name	
Date of completing the checklist	
Name of appraisal leads	
Stage of appraisal (i.e. outline business case, full business case etc)	
Appraisal period (e.g. 2015 to 2075)	
Assumed opening date	
Geographical area covered by the model	
Price base	
WebTAG version used	
Scenarios and sensitivities for which model outputs produced	

Key outputs and assumptions	Please record information below
Name of models used	
Types of models used (e.g. fixed demand, variable demand)	
Versions of models used	
Software used and format of outputs (excel, csv etc.)	
File of transport capacity coded into the model	
List of metrics that have been forecast in the appraisal (e.g. passenger numbers by mode; traffic volumes etc.)	
Output files for all forecasts: after year 1, 5 and whole appraisal period	
Input files for all assumptions and parameters (including values of time; population, land use planning, fuel costs etc.)	
Links to key external sources used (e.g. population data)	
Costs output files (split by capital and operating costs)	
Forecast report	

The completed checklist could potentially become part of the documentation reviewed for business case sign-off. The purpose of using a handover pack could be explained in the guidance for the technical project manager in WebTAG, for example, which has an annex listing the modelling documentation required at the business case submission stage.

Comparing forecasts from the appraisal with actual observations in the evaluation

The handover pack can be used by the evaluator to inform comparative analysis of forecasts with actual observations.

For each transport intervention, the starting point would be a high level comparison of forecasts and actual observations (also referred to as outturns) of key metrics and assumptions at the scheme level (these will vary by scheme but an illustrative example of what these might look like is provided in **Figure 5**). For example, evaluators can start by comparing traffic volumes and note whether outturn volumes roughly correspond to those forecast at the appraisal stage (or at least if the changes in those variables are as forecast). This would provide a high level starting point indicating where material discrepancies may lie. Then, this would go further into more disaggregated comparisons (e.g. by link, time of day) noting any systematic and material discrepancies between forecasts and outturns.

Figure 5. Illustrative example of checklist for evaluators

Key outputs and assumptions	Forecast	Outturn
Scheme opening date		
Expected scheme benefits (total)		
Expected scheme cost (total)		
Expected scheme benefits in 1st year after opening		
Expected scheme benefits in 5th year after opening		
Estimated scheme Benefit Cost ratio		
Estimated value of wider economic benefits (by type)		
Traffic volumes in 1st year after opening (potentially split by link)		
Traffic volumes in 5th year after opening (potentially split by link)		
Journey time savings (by time of day) in 1st year after opening (potentially split by link)		
Journey time savings (by time of day) in 5th year after opening (potentially split by link)		
Number of accidents in 1st year after opening (potentially split by link)		
Number of accidents in 5th year after opening (potentially split by link)		
Assumed level of economic growth (split by year) for the period considered in the appraisal		
Assumed trajectory of fuel prices (by year) for the period considered in the appraisal		
Population growth assumptions		
Other assumptions		
Are routing assumptions made at the appraisal stage broadly in line with outturn?	enter free text here	
Are land use assumptions made in the appraisal broadly in line with outturn?	enter free text here	
Are assumptions about other transport schemes likely to impact the modelled area made in the apprais	enter free text here	
Assessment of key factors driving significant (+/-15%) differences between forecasts and outturns		

It is important to note that most projects will have outturns which differ from those that were forecast to some extent. When considering whether variation is material and of concern one should check factors such as:

- Whether outturns[21] are within tolerance levels (for example, Highways England POPE analysis uses +/- 15%);

- Whether the key assumptions underpinning the appraisal were broadly correct (again within a certain tolerance range). Here, the evaluator should consider that:

  - if exogenous assumptions are not within the tolerance range (for example, if population growth is notably higher than expected) yet outturns are within tolerance levels, this raises questions about the validity of the appraisal (i.e. outturns were correct by chance); and,

  - if exogenous assumptions are not within the tolerance range and neither are outturns, this raises the question of how far the variation in outturn can be explained by the incorrect exogenous assumptions. This could in principle be checked by re-running the appraisal models (noting that this will not be practical or cost-effective in many cases, however) with corrected assumptions to assess whether the discrepancies between forecasts and outturns could be the result of other appraisal components such as model parameters, elasticities and values of time.

When comparisons between forecasts and outturns are made, it would be important to consider not just the central case modelled as part of the appraisal but also any sensitivity tests or scenarios that have been run. For example, if economic growth turns out to be lower than projected in the central case, it may be more appropriate to compare outturns with forecasts from a low growth scenario, provided one has been modelled. To enable these comparisons, it would be important for appraisers to store any sensitivity test runs or scenarios that have been examined as part of the appraisal. Doing so could save the need for further model runs which would have been needed to update exogenous assumptions.

In carrying out this analysis, where consistent errors are observed, this could serve as a 'red flag', identifying that there is a case for a more detailed investigation. The complexity of modelling means that a separate piece of more detailed analysis is needed to identify why there are discrepancies between forecasts and outturns. For example, if the database identifies that forecasts are consistently at least 30% higher than the actual traffic flows when using one form of modelling while forecasts are close to outturns when another form of modelling is used, then this will need a focused piece of work to examine the reasons why.

Similarly, if the central database identifies that traffic volumes are consistently over estimated for a certain type of scheme (e.g. junction) and accurate for another (e.g. bypass - new road) one could examine the differences between

---

[21] We note that comparing outturns and forecasts at a single point in time (say 1st year after opening) may not be sufficient to make judgements about the accuracy of forecasting. It may also be helpful to compare forecasts and outturns on a route over a longer horizon, say 3 or even 5 year period to ensure that any discrepancies that are picked up are genuine and not driven by outlier events.

the way in which these are modelled and identify possible drivers of discrepancies.

The overall approach for red-flagging where further investigation may be required is summarised in **Figure 6** below.

Figure 6. Summary of approach to red-flagging
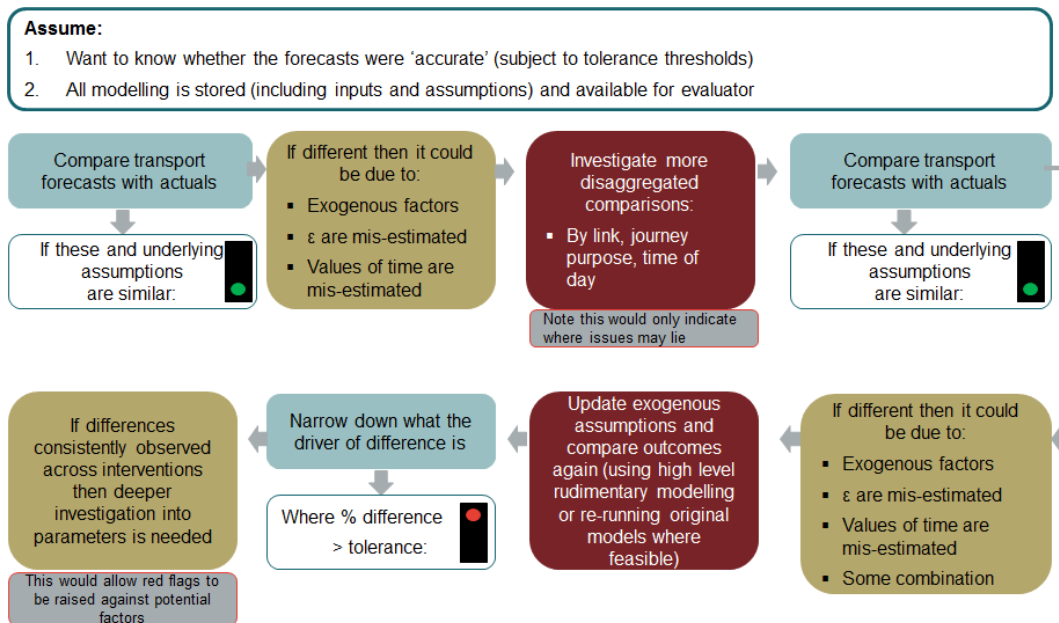
Using portfolios of project evaluations to learn about the validity of appraisal modelling

Although it is possible to learn from individual scheme-level evaluations, there is likely to be greater scope for learning to inform appraisal if a portfolio of interventions is considered. This is because it provides a larger sample size from which to learn, and also allows various comparisons across scheme evaluations to be made to identify common issues.

A portfolio of schemes could be identified at the appraisal stage for evaluation and the results of both the appraisal and evaluation could be logged centrally in a database. The database would effectively contain the information presented in **Figure 4** and **Figure 5**. At the high level this would include:

- the metrics under consideration (e.g. traffic volumes);

- appraisal baseline values and forecast values and associated difference between them over time for a particular scheme, link or time period;

- evaluation counterfactual value and outturn value and associated difference between them for the same scheme, link or time period;

- appraisal modelling information (model type, version, software, date of calibration, WebTAG version, exogenous assumptions, capacity assumptions, elasticity parameters, values of time etc.);

- description of known differences between appraisal and evaluation exogenous assumptions (for example, land planning assumptions, fuel costs or rail fares, GDP growth etc.); and,

- commentary on the extent to which forecasts differ from actuals and the disaggregated comparisons for which discrepancies are more significant (e.g. morning peak passenger volumes; traffic on a particular link etc.).

The schemes within each portfolio could have similar objectives and share common characteristics (e.g. major road schemes to enhance interurban capacity; local major transport schemes to relieve pinchpoint road congestion; and intra-urban rail infrastructure schemes). This would enable evaluators to take stock of any persistent or systematic material errors, identify the factors common to when those discrepancies occur and log them. The analysis would consider all scenarios estimated as part of the appraisal, rather than just focus on the central case.

Once a sufficient number of schemes have been examined in this way and a log/database created of the discrepancies and possible factors driving them, the information can be examined in more detail to identify patterns. This information can be interrogated to enable analysts to investigate the factors that are common when material discrepancies are observed across particular types of scheme, particular geographical areas, particular types of models, particular times of day outputs, and, particular types of link (e.g. dual carriageway etc.). When the factors common to systematic variations between forecasts and outturns have been identified or 'red flagged', those issues can be subject to separate detailed model validation research.

Recommendations - 1

As part of the process of verifying whether transport appraisal models are able to forecast transport outcomes with an appropriate level of accuracy, we make the following recommendations for DfT and its partners:

Implement appraisal handover packs

- Design and implement the use of a handover pack as a method of passing forward appraisal information to the evaluator. Linked with this, ensure that externally commissioned modelling logs and stores relevant information to feed into the handover pack.

- Include within the handover pack a requirement to store all sensitivity test runs and scenarios, not just the central case. (This could save the need for additional model runs as part of the evaluation which can be very costly and infeasible).

Develop portfolio approaches for evaluation

- Consider developing consistent approaches to evaluating portfolios of schemes to identify areas where appraisal tools are systematically and materially mis-forecasting outcomes and hence warrant further investigation. This approach is already followed for major road schemes through POPE and, to some extent for Local Major Schemes.

- Work alongside Highways England and local partners to explore ways in which the POPE programme and other evaluation series which use meta analyses, such as the evaluation of Local major schemes, could be adapted to generate further evidence to inform validity tests of appraisal modelling in a systematic way.

Align evaluation work with appraisal evidence needs

- Explore how best to capture evidence and learn from evaluations to be able to inform future appraisal methods.

Using evaluations to inform appraisal

Introduction

Appraisals rely on a number of evidence-based assumptions. Although evaluation can help us learn about a wide range of factors, this section focuses on the role that evaluation can play specifically in relation to informing the development of appraisal assumptions. In this context, evaluations can in principle be used in two key ways.

- First, they could help to corroborate some of the key appraisal assumptions relating to transport user benefits;

- Second, they could potentially go further by providing additional evidence to inform appraisal areas where there are currently gaps in knowledge.

It is however important to recognise that scheme-level evaluation cannot be used to address all possible evidence gaps and that independent standalone research studies may be better placed to address some. For example, scheme-level evaluations are not likely to be well placed to inform elasticity estimates (such as agglomeration elasticities for example) as studies using robust quantitative techniques such as econometric modelling are likely to be more fruitful.

Identifying where evaluation can inform appraisal

In order to use scheme-level evaluations effectively to inform appraisal assumptions, it is useful to follow a five step process:

- Step 1: identify where there are evidence gaps or a case for updating evidence used within appraisal processes;

- Step 2: consider the extent to which scheme-level evaluation is likely to be the appropriate method for developing that evidence;

- Step 3: develop specific questions on which the evaluations will focus;

- Step 4: design robust evaluation approaches to develop the required evidence; and,

- Step 5: collate and interpret the evidence and translate into appraisal guidance when appropriate.

For the purposes of this report, we tested this approach for real as an exercise with the DfT. This example focuses on transport impacts on the economy, which is an emerging area where there are sufficient deficiencies in the evidence base and where evaluation would be extremely useful in improving understanding of these impacts. This example has been chosen as an illustration and it is not intended to imply that this issue is of particular

importance. Indeed, evaluation could be used in a similar way to check more standard modelling assumptions, such as outturn elasticities of transport demand, for example.

This exercise is presented for illustrative purposes to demonstrate how it could work. The five steps are described next.

Step 1: Identify where there are evidence gaps or a case for updating evidence used within appraisal processes

In this first step, we were keen to seek a strategic view as to the areas of the current WebTAG appraisal guidance where DfT felt that there were evidence gaps, or where new evidence would be valuable. We therefore engaged with experts in DfT's Transport Appraisal and Strategic Modelling (TASM) team - the guardians of WebTAG appraisal guidance.

The TASM team identified transport and the economy as an area where improving appraisal methods and gathering stronger evidence will be highly beneficial. Four aspects of this area were suggested for consideration:

- Estimating displacement effects for wider economic impacts;

- Updating agglomeration elasticity estimates;

- Reviewing the current dependency test in 'dependent development' guidance; and,

- Obtaining benchmark estimates for total public sector spending required to enable dependent developments (not just transport spending).

It should be noted that this approach could be applied equally well to building evidence using evaluations for environmental or social impacts.

Step 2: Consider the extent to which scheme-level evaluation is likely to be the appropriate method for developing that evidence

The next step was to consider which of the evidence gaps could be effectively informed by evaluation evidence. In this example, having considered the practicalities of the four aspects of transport and the economy that were identified, greatest value-added could be achieved by focusing on the use of evaluation to inform assumptions around displacement effects.

In line with the HM Treasury Green Book, displacement concerns the degree to which an intervention facilitates outcomes in one area at the expense of reductions in activity elsewhere in the economy, where those reductions would not have occurred if the intervention had not been made.

The worked example we use in steps 3 to 5 therefore relates to the use of evaluation to inform assumptions of displacement of economic activity.

Step 3: Develop specific questions on which the evaluations will focus

This step involves ensuring sufficient understanding of the nature of the issue - in this case potential displacement effects - to be able to identify the specific questions that could be asked of the evaluation.

Transport improvements can lower the costs of travelling from one area to another and effectively increase the pool of workers that are accessible to an area. Displacement can occur when improvements in transport links lead to jobs relocating from one area to another rather than new jobs being created, although the value of displacement will depend also on the value of jobs[22].

Best practice evaluation guidance recommends that displacement effects need to be accounted for where they could be material. With increasingly devolved budgets and therefore appraisal at a local level, central decision-makers will be increasingly mindful of the potential for these effects to occur.

In general a number of factors are likely to determine the extent to which displacement effects may be a cause for concern. These include:

- The degree of spare capacity in the local economy: the extent to which affected labour markets have currently unoccupied workers;

- The skills composition of the local economy and neighbouring local economies: the extent to which the local labour market has the skills to meet the requirements of expanding or new businesses;

- The types of businesses serving the local and neighbouring economies: the extent to which businesses affected by the transport scheme serve local or global markets; and

- Travel to work patterns and the degree to which a transport scheme affects the generalised cost of travel[23]: i.e. the ability and willingness of local workers to access wider labour markets.

Exactly how these factors affect displacement is an empirical matter and there is currently a lack of evidence in this regard. Evaluations can help fill this gap by examining how the estimated magnitude of displacement effects varies from scheme to scheme. Displacement estimates from evaluations can then be used to assess the accuracy of appraisal estimates where possible. For example, where jobs are observed or expect to move as a result of a transport intervention[24], evaluation evidence could be used to assess whether these impacts have been accurately estimated by assessing:

---

[22] If workers move to jobs that are better suited to their skillset and this makes them more productive, this would limit the degree of displacement resulting from a transport scheme.
[23] A weighted sum of time and other costs of travel which can be measured in units of money or time.
[24] See WebTAG Unit A2.1 available here:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/427091/webtag-tag-unit-a2-1-wider-impacts.pdf

- Whether ex ante estimates for the number of people relocating as a response to the transport scheme are accurate (and whether the before and after locations are correct);

- Whether it's appropriate to assume the GDP per worker of these individuals is the (mean) average for their local authorities both before and after relocating (as currently recommended in WebTAG); and

- Whether the increase in tax from these individuals has been accurately appraised (currently recommended to be 30% of the increase in GDP).

Questions to ask of the evaluation therefore could include:

i. What is the estimated magnitude of displacement effects of the intervention?

ii. What are the factors driving the observed displacement?

iii. Which surrounding areas are likely to be affected by displacement in the presence of the transport intervention?

iv. What are the conditions under which displacement is likely to be more significant?

In step 4, we provide practical guidance on how a detailed investigation of a topic - in this example displacement - can be used in order to inform current knowledge gaps.

Step 4: design robust evaluation approaches to develop the required evidence

We have developed an approach to evaluating displacement effects such that the evidence can then be used to inform appraisals. This is described below.

Whether displacement occurs as a result of a transport intervention is an empirical matter which can be examined through a combination of quantitative and qualitative analysis. This could involve a series of steps:

- As a starting point, evaluators would need to define the geographic area of impact - the 'treatment' area - where the transport scheme is expected to have a direct economic effect.

- The next step would be to develop a hypothesis about which areas displacement might come from. The 'displaced' areas would typically be immediately adjacent to the treatment area but need not be.

- Then we would gather empirical evidence (administrative and survey data) to characterise the treatment and displaced areas over a period of time covering a sufficiently wide time window to study trends before and after the transport intervention. Characterising the areas includes looking at:
  - Local labour market composition in terms of employment status (employment, unemployment, inactivity), skills (education levels, qualifications) and demographics (age, ethnicity, gender etc.).

- Local employer composition in terms of sectors of employment, size of businesses and types of businesses (serving local markets or national/global markets).

- The next step would be to examine employment inflows/outflows in 'treatment' and 'displaced' areas after the transport intervention (or after the transport intervention announcement[25]) through primary data collection. The advantage of primary data is that it allows researchers to better understand causality and attribution.

We describe this process in detail in the annex.

Step 5: collate and interpret the evidence and translate into appraisal guidance when appropriate

This step considers the process of collating the evidence from across scheme-level evaluations and how it can be interpreted for use in appraisals.

Studying displacement effects can be resource intensive and requires considerable efforts by evaluators in order to obtain meaningful estimates. These efforts need to be proportionate to the cost and nature of the scheme in question. Displacement effects should be considered in detail in evaluations where this is proportionate and where displacement considerations are likely to be material (as per the best practice appraisal and evaluation guidance by HMT). Schemes that are expected to have little impact on employment need not look at displacement in any detail while those for which job creation is a major strategic objective should.

Displacement effects will inevitably be subject to some degree of uncertainty and it is likely that the most robust studies will produce a range of estimates. This is useful as it gives a sense of the likely influence of displacement on overall scheme performance in the 'best case' and 'worst case' scenarios, plus others in between. This helps to interpret the analysis appropriately by not implying spurious accuracy in the estimates, but by being clear on the underlying assumptions under each estimate produced.

Furthermore, as previously discussed the degree of displacement will also vary depending on a number of factors such as the type of transport scheme and local economic conditions. It therefore does not seem appropriate that a single displacement estimate should be recommended in appraisal guidance for application to all schemes. Exploring how such displacement effects vary across schemes, or whether systematic biases exist across schemes (in terms of assumed displacement effects at appraisal stage), would be interesting and informative.

Ideally, to translate evaluation evidence into appraisal guidance, one would obtain estimates covering how displacement varies by:

---

[25] Noting that sometimes effects of transport interventions can begin in anticipation of the intervention so before it is actually complete.

- Type of scheme: some schemes will have greater potential to generate significant employment effects (relative to the value of the scheme);

- Location: the level of displacement will vary according to local labour market conditions;

- Size of scheme: the level of displacement will vary according to the scale or the reach of the impact of the scheme; and

- Scheme mode: whether displacement is more likely for, say, road rather than rail interventions.

This example is illustrative but highlights the considerations that are important for the practitioner when generating and interpreting evaluation evidence to inform appraisals.

Conclusions

Policy makers are interested in the effects of transport schemes on the economy, environment and society. This chapter has shown how evaluation can help to inform:

- The degree to which transport appraisal models accurately forecast future transport outcomes and where 'red flags' need to be raised against issues that merit further investigation;

- How evaluation can be used generate evidence to inform appraisal assumptions (using the issue of displacement effects as an example).

Recommendations - 2

We recommend that DfT and partners consider the following:

Align evaluation work with appraisal evidence needs

- Repeat this process of engagement and consultation with DfT appraisal leads in order to identify, at a strategic level, other areas where evaluations can be used to fill knowledge gaps going forward.

- Consider aligning DfT evaluations (which will be carried out for other main purposes) to ensure that they add value by delivering evidence that will inform WebTAG guidance and models where this is feasible.

- Commission relevant evaluations and store the findings systematically such that the evidence base can, over time, be translated into appraisal guidance.

Annex

Chapter 3: Case study selection

This section provides further detail on the case studies used in our analysis.

For this stage of work we used four case study interventions across three main transport spending areas (local transport, rail and strategic roads). The case studies were intended to differ in their scale and focus and stage of appraisal so that we have the opportunity to make observations from various perspectives. Obviously, we recognise that three case studies is a small sample size, but it is sufficient to identify some cross-cutting issues and also to highlight particular issues that DfT or others may wish to examine in more depth through separate work.

To select the case studies, we developed selection criteria which were:

- They have a robust business case (either outline business case or full business case) or detailed economic analysis which is typically used in a business case available;

- They have a modelling framework available for us to explore;

- The business case leads are available for us to talk to; and,

- They cover different areas of the DfT portfolio as far as possible including local transport, rail and strategic roads.

Using these criteria we worked with policy and analysis teams across DfT and Highways England to identify appropriate case studies.

Figure 7. Case studies selected for this study

	Case study	Rationale for case study choice
Local transport	Rochdale Interchange (opened 2013) Doncaster White Rose Way (opened 2013)	• Appraisal and 1-year evaluation documentation available for both (few schemes had both) • Rochdale Interchange is part of a wider programme so scope to inform other appraisals/ evaluations. Doncaster White Rose Way offers a road alternative • Allows us to explore the flow from appraisal to evaluation in practice
Rail	Crossrail 2 (business case stage)	• Economic analysis information is available • Modelling is innovative and well documented so offers scope to learn for other interventions • Key project objective is WEBs so opportunities to learn how they have been forecast and how they could be evaluated (and embed evaluation thinking at this early stage)
Highways England	A46 Newark to Widmerpool (opened 2012)	• Appraisal and 1-year evaluation documentation available for both (few schemes had both) • POPE is well-regarded and well-established so offers good opportunities to learn from this case study • Scope to help POPE leads in considering how WEBs could be included in POPE as WebTAG evolves

As shown in Figure 7, the case studies allow us to explore interesting issues in each case. Of particular note are the following advantages of this selection:

- The local transport interventions have the flow of three forms of documentation relating to the links between appraisal and evaluation. That is we had access to the appraisal (business case); the evaluation plan; and, the one-year evaluation report.

- The rail intervention is a large transformational project which is still in its relatively early stages of appraisal because the strategic outline business case is currently being prepared for completion in March 2017. Given one of its main focuses is on facilitating wider economic benefits, this offered a good opportunity to explore how the evaluation could be planned at this early stage to consider the realisation of wider economic benefits.

- The Highways England case study uses the Post Opening Project Evaluation (POPE) methodology which is widely used for all major schemes in which Highways England invests. Findings and observations are therefore likely to be relevant for other schemes that go through the POPE process (there have been around 80 schemes in the last two years).

Overview of the case study schemes

A brief overview of each of the case studies is below.

1. Rochdale interchange

This intervention involves the construction of a new interchange in Rochdale to serve bus passengers and passengers from the adjacent Town Centre Metrolink terminus. It involves the relocation of the bus station from one side of Smith Street to the opposite side, at a riverside site. The relocation of the existing bus station frees up land that is being used in the wider redevelopment of the town centre.

The scheme was delivered by Transport for Greater Manchester, with £11.5 million funding from DfT, and it was operational from November 2013. This intervention formed part of a wider programme of interchange upgrades across Greater Manchester which was estimated to cost around £100 million in total.

2. Doncaster White Rose Way

This local major scheme involved widening 1.9km of existing single carriageway to dual carriageway and the replacement of 2 existing roundabouts with high capacity signalised junctions. It also involved the construction of cycleway and footway routes, linked to the proposed Greenway Cycle Route. A further aspect of the intervention was a new bridge over the East Coast Main Line.

The scheme was delivered by Doncaster Metropolitan Borough Council (DMBC) with £32m funding from the European Regional Development Fund, DfT, DMBC and the local transport plan. It was operational from May 2013 and formed part of the wider Doncaster Network Management Strategy.

3. Crossrail 2

Crossrail 2 (CR2) is a proposed new railway that represents a transformational investment in the UK's railway network serving London and the wider South East. It would connect the National Rail networks in Surrey and Hertfordshire via new tunnels and stations between Wimbledon in the south and Tottenham Hale in the north, linking in with London underground, London Overground, Crossrail 1, national and international rail services.

CR2 will help relieve congestion on busy routes into central London and support economic development in and around the capital, specifically targeting some key opportunity areas, such as the Upper Lea Valley, Victoria and King's Cross. It would create better connections across the South East and the whole country, with a new Euston St Pancras station providing direct access to the new High Speed 2 line, Thameslink and Eurostar services.

If approved, it would be expected to open in 2030 with a capital cost of between £27 and £32 billion in 2014 prices.

4. A46 Newark to Widmerpool

The purpose of the scheme was to provide a 17.5 mile (28km) section of dual carriageway to replace a substandard section of single carriageway trunk road which was prone to regular congestion and delays and had a poor safety record. Numerous low quality junctions were replaced by grade separated junctions, allowing the A46 to have no junctions for through traffic to stop at. The scheme involved the construction of 8 full or compact grade separated junctions and 13 bridges and underpasses.

The scheme was originally planned to open in 2016 but in fact opened in April 2012 (investment was brought forward). The total cost was £369.7 million (2013 prices).

Chapter 4: Evaluating the nature and scale of displacement effects

In chapter 4, we set out an approach to estimate the nature and scale of displacement effects. This annex provides more detail for the interested reader.

Data sources needed for the analysis

The table below presents the data that would be needed for the analysis.

Figure 8. Summary of data sources

|	Data source	Time period covered	Geographic level of data	Definitions	Comments
Demographic data					
age	Quarterly Labour Force Survey	1992-2015	GOR, UA, LA, NUTS, Electoral wards, Census output areas (and others)	Government Office Region (GOR), Unitary Authority (UA), Local Authority (LA),	QLFS variables at low levels of geographical disaggregation only available for secure access QLFS and are not available in standard QLFS data sets.
gender	Quarterly Labour Force Survey	1992-2015			
ethnicity	Quarterly Labour Force Survey	1992-2015			
Skills data					
qualifications	Quarterly Labour Force Survey	1992-2015			
education level	Quarterly Labour Force Survey	1992-2015			
experience	Quarterly Labour Force Survey	1992-2015			
Labour market data					
employment	Quarterly Labour Force Survey	1992-2015			
unemployment	Quarterly Labour Force Survey	1992-2015			
inactivity	Quarterly Labour Force Survey	1992-2015			
employmen by sector	Quarterly Labour Force Survey	1992-2015			
claimant count	Nomis	2013-2016			
Employer data					
size of businesses	BSD/IDBR	1997-2014	Postcode		Data only available at secure location
types of businesses	BSD/IDBR, Quarterly Labour Force Survey	1997-2014	Postcode		ONS Virtual Micro-data Laboratory
employment flows	Primary data collection				Survey of local businesses

Key outcome measures

Labour market displacement analysis can look at the number of jobs as the key metric of interest. However, the value of jobs would also need to be considered - if a job is lost in locality A but a more valuable job (as indicated by average earnings data) is created in locality B, the displacement effect would be less than 100%.

Level/type of analysis

The analyses need to be carried out at the very local level. The exact geographical boundaries to be applied will depend on the schemes examined. The appropriate level of disaggregation will be determined by the size of the scheme and the geographic area it is likely to have an impact on, i.e. the size of the 'treatment' area. The 'displaced' area again may vary from scheme to scheme but in most cases, it would be expected to be an area adjacent to the treatment area. A possible approach for defining the 'displaced' area is to draw a buffer around the treatment area and define all areas that fall within that buffer as the 'displaced' area. The exact width of this buffer will vary from scheme to scheme. In densely populated areas a relatively narrow buffer (1/2 mile to 1 mile) may be appropriate while in sparsely populated areas a wider buffer may be chosen (2 miles to 3 miles). A stylised illustration of this is shown in **Figure 9** although we note that in practice the shape of the buffer need not be circular and may well be determined by the standard statistical geographies for which data is available.

Figure 9. Defining the treatment and displaced area - illustration
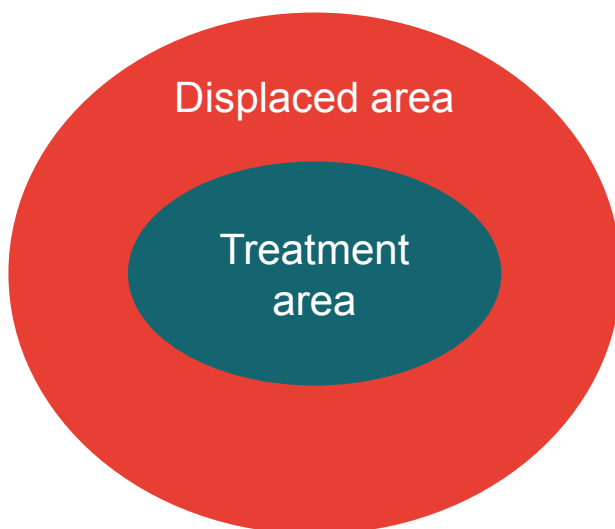
Once the 'treatment' and 'displaced' areas have been defined, researchers can match local economy data[26] to them in order to describe their main characteristics and perform analysis of trends before and after the transport intervention.

Below we illustrate how the type of analysis described in chapter 4 can be carried out in practice and what it may mean for displacement. In this example we compare two areas which are reasonably similar in terms of size and labour market composition, both in terms of employment status and skills. Business composition in the treatment area is quite mixed with roughly half of businesses providing goods and services to the local market with the other half producing tradable products.

In this example (illustrated in **Figure 10**) we compare a treatment area which pre-scheme had a relatively flat employment growth profile (0.2% p.a.) with an adjacent area with a relatively high employment growth profile (1.9% p.a.) focusing on number of jobs as the key metric of interest. Following the opening of the transport scheme the employment profile of the treatment area changes such that employment increases by an average of 1.9% p.a. in the first five years. Conversely, employment growth slows considerably in the adjacent area to an average of 0.7% p.a. in the five years after scheme opening.

Using pre-scheme employment trends in both treatment and displaced area, researchers can project what the expected rates of employment growth would have been in both the treatment and displaced areas had the scheme not taken place. The forecast levels of employment are depicted with the dashed lines in **Figure 10**. Based on pre-scheme trends employment would have been expected to increase by 1 job in the treatment area, from 150 to 151. The expected employment increase in the displaced area based on past trends would be expected to have been 11 jobs, from 116 to 127.

Comparing the actual and forecast job growth in the treatment and displaced areas gives us an idea of the likely level of displacement that might have occurred as a result of this notional transport scheme. In the treatment area, employment increased by 15 jobs relative to a forecast of 1, a net gain of 14 jobs while in the displaced (adjacent) area employment increased by 4 jobs against an expected increase of 11, a net loss of 7 jobs. In total, this suggests that half of the employment increase in the treatment area may be as a result of displacement. Of course, this analysis can be extended to account for the average productivity of workers and how that has changed as a result of the transport scheme. For example, if the 7 workers who relocated jobs ended up in occupations which are better suited to their skillset and hence increased their average productivity, the level of estimated displacement would fall.

This example provides a simplistic illustration of how evaluators could approach the estimation of displacement effects for individual schemes based on a

---

[26] The level of data required will vary by scheme. For schemes affecting very small areas it may be necessary to analyse data at the Census Output area level while for larger schemes it may be sufficient to look at Local Authority level data.

comparison between actual and forecast changes in employment where forecasts are derived based on past trends[27]. The advantage of this macro level approach is that it can be implemented using existing data. The disadvantage is that since localities are often exposed to numerous interventions at the same time, it may be difficult to attribute observed changes to a specific transport intervention. In practice the past does not always accurately predict the future so forecasting employment may require taking into account additional factors. In particular, where local economies benefit from specific other investments (or indeed divestments) which are likely to impact employment it is important to take these into account so that the effects identified in this analysis can be attributed to the transport scheme as far as possible (recognising that in practice attribution is difficult using this macro level approach).

Figure 10. Illustration of displacement analysis
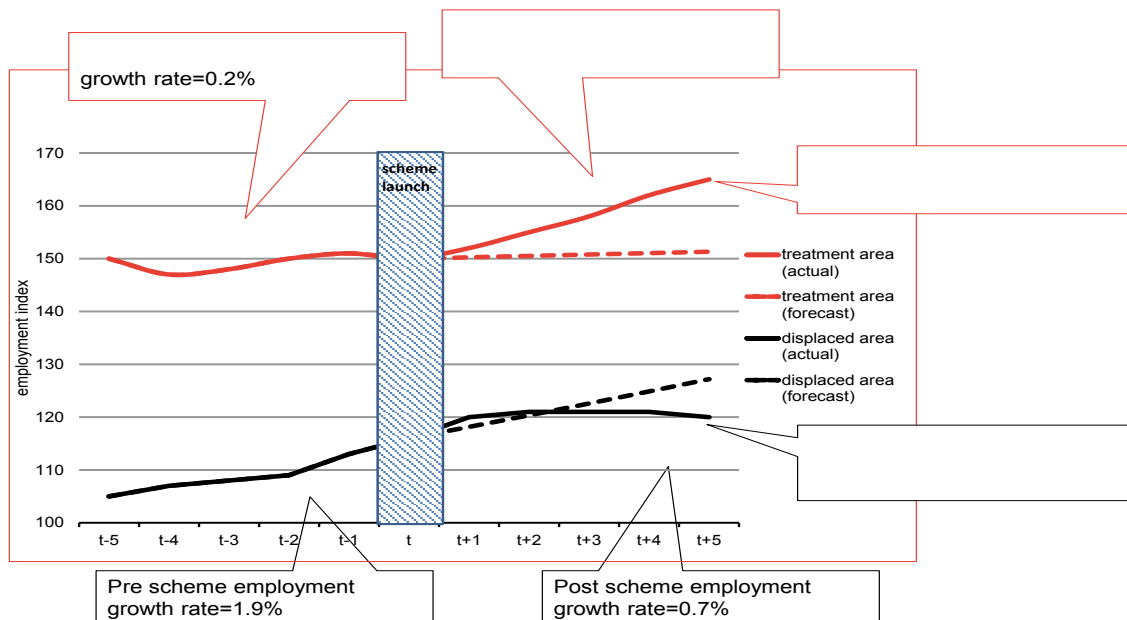
An alternative approach, which can be used to complement the above analysis (or indeed be implemented as a standalone assessment of displacement) is a micro level analysis where individual beneficiaries (businesses) of a transport intervention are identified and tracked over time using firm-level data to identify business creation and expansion as well as company deaths in the affected areas. In order to carry out this type of analysis evaluators would need to gather information on the population of businesses in the treatment and displaced area. This can be done using secondary data such as the Business

---

[27] An alternative approach would involve finding a 'counterfactual' area which is similar to the displaced area (but did not benefit from a transport scheme) and use employment data from that area to estimate what would have happened in the displaced area had it not been for the transport scheme. This method can be more robust than a trend analysis (described above) provided that a suitable counterfactual area can be found which is in practice very challenging.

Structure Database (BSD) which contains information on the population of businesses in the UK (approximately 98% of businesses based on turnover). The data contains information at the plant level (i.e. local units of national chains) but also includes data on the wider firm that a plant belongs to. The dataset provides information on the location of plants (at the postcode level), the sector of the economy in which they are classified (using the Office for National Statistics (ONS) Standard Industry Classification (SIC) codes) and their levels of employment. The BSD makes it feasible for researchers to identify the population of businesses in the treatment and displaced areas and then examine how their employment changes over time. Additional data on the firms in question and the flows of employment to and from them can be studied in detail by collecting primary data by means of a survey. The survey could be tailored to the specific question that the researcher is looking to explore but would include questions which try to unpick:

- The reasons for any observed changes in employment (or lack thereof) and the extent to which these can be linked to a transport intervention; and

- The extent to which any changes in employment are due to firms hiring currently inactive/unemployed local workers or moving workers who are already employed elsewhere.

The advantage of the micro approach is that it is more focused than the macro approach (as individual firms can be identified) and is better for showing attribution - the researcher is able to link changes in business activity to the transport intervention and displacement effects can be demonstrated explicitly through employer surveys. The disadvantage of the approach is that primary data collection and tracking individual firms over time requires more resource and technical capability - using data sets such as the BSD can only be accessed from secure ONS locations and using them requires considerable technical expertise as these are large data sets that require careful manipulation to obtain meaningful results.

Whichever approach (macro or micro) researchers choose to follow, it is important to first carry out descriptive analyses of affected areas which can help identify the conditions under which displacement is more likely. For example, we previously noted that displacement is in general less likely if local labour markets have spare capacity and the skill composition of the labour force is matched to the skill requirements of local businesses. It is also true that displacement is more likely to occur in areas served predominantly by businesses serving local markets. Carrying out a detailed description of the treatment and (likely) displaced areas can go a long way towards identifying the magnitude of any possible displacement effects. It is also a useful cross check against which the results from more detailed empirical analyses (outlined above) can be compared to ensure they fall within reasonable limits and are therefore robust.

For example, if a treatment area is characterised by businesses which are heavy exporters, an expansion in their activity is likely to be associated with

relatively small displacement effects because those businesses are likely to be able to expand their market reach rather than simply shift the markets they serve. So in this example, if 80% of production is exported we would expect to find displacement effects of around 20%. If on the other hand the treatment area is characterised by local businesses which serve only local markets then we would expect estimated displacement effects to be large. Based on this type of descriptive analysis, researchers could formulate plausible scenarios for the levels of possible displacement asking "what if 90% / 80%/ 60% of the employment change is displaced?" type questions and then comparing the results from scenario analysis to empirical estimates based on more rigorous data work. In our view, this descriptive approach is necessary but not sufficient to estimate displacement effects. It is there to complement more rigorous data work (described above) and to set the bounds within which we would expect displacement effects to fall, but does not, on its own, provide convincing estimates of displacement effects.

Finally, it may be possible to demonstrate the displacement effects (or lack thereof) of a scheme through more qualitative evidence such as case studies. These may not be sufficient on their own but could prove to be useful additions to quantitative analysis. Case studies are often qualitative and focussed on specific units (areas, businesses) and so by their very nature may not be representative. Nonetheless, they are a very useful tool for identifying attribution (the specific channels through which a transport scheme affects business decision making), for identifying employment flows, and for adding colour and intuition to quantitative work which can sometimes be more difficult for wider stakeholders to relate to.

Bibliography

Guidance on logic mapping:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/3817/logicmapping.pdf

Guidance for transport impact evaluations:

http://webarchive.nationalarchives.gov.uk/20110202223908/http://www.dft.gov.uk/pgr/evaluation/evaluationguidance/transportimpact/pdf/guidance.pdf

Monitoring and Evaluation Framework for Local Authority major Schemes
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/9154/la-major-schemes-monitoring-evaluation.pdf

Monitoring and Evaluation Framework for Local Sustainable Transport Fund:

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/35975/lstf-monitoring-evaluation-framework.pdf

WebTAG Unit A2.1 Wider Impacts guidance:

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/427091/webtag-tag-unit-a2-1-wider-impacts.pdf