TRANSPORT SCOTLAND:
Guidance for the evaluation of rail projects

CONTENTS

1. Introduction

1.1 This guidance has been developed by Transport Scotland analysts to inform evaluation of rail projects. The guidance sets out the key issues to consider and suggested steps to follow, drawing on case study examples where possible.

1.2 To inform the development of this guidance three pilot evaluation projects were commissioned by Transport Scotland in 2013 to 2014 (Evaluation of Laurencekirk Railway Station, Evaluation of Larkhall-Milngavie Railway Line, and Evaluation of Airdrie-Bathgate Line[1]). Two workshops were also held in March 2013 and November 2014 with rail practitioners, academics and policy officials to discuss areas to consider when carrying out rail evaluations, and also to reflect on learning from the three pilot projects.

1.3 The guidance is divided into different sections reflecting the different stages of evaluating a rail project, beginning with considerations to be made at the appraisal stage, followed by issues to consider when planning an evaluation, identifying data sources, through to the analysis and reporting of the data.

1.4 This guidance will be updated to reflect feedback received and learning from evaluations which Transport Scotland and other transport funders commission.

[1] The final reports of all three evaluations are available from the Transport Scotland website.

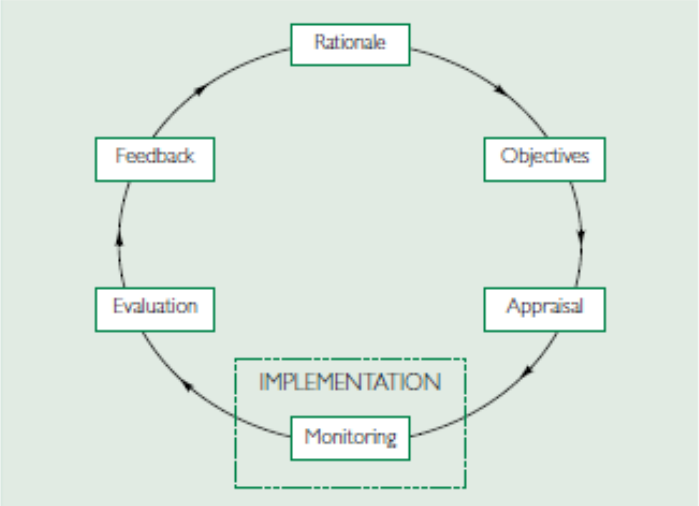
2. Background

2.1 The last 20 years has seen record investment in Scotland's railways which has led to the opening of new railway stations, new lines and improvements to existing infrastructure. Traditionally a great deal of time is invested in appraising such projects and setting objectives at the start but less work has been carried out on ex-post evaluations to assess whether they have met their objectives and have been a sound investment of public money. Guidance on evaluating Scotland's Motorway and Trunk Road Programme has existed for a number of years in the form of STRIPE[2].

2.2 This section outlines the importance of evaluation, highlighting that it is a key stage in the policy cycle. It also introduces existing high level evaluation guidance for the public sector.

Importance of evaluation

2.3 The policy cycle highlights the stages a policy, project, or programme should follow to ensure that it is successfully implemented and evaluated. Referring to the example policy cycle (see below), new rail projects have traditionally been well developed at the objective setting and appraisal stages, but less so for the later stages around monitoring and evaluation.



[2]http://www.transportscotland.gov.uk/system/files/uploaded_content/documents/tsc_basic_pages/Road/STRIPE_Trunk_Road_Project_Evaluation_Guidance.pdf

2.4 It is important to monitor and evaluate to test whether the project has been a sound investment of public money, to assess what the outcomes are compared to the objectives, and provide evidence for future interventions and investments..

2.5 The box below highlights findings from a literature review of existing rail evaluation projects.

Case Study highlighting literature review of existing rail evaluation projects:

To inform the first version of this guidance a literature review of existing rail evaluation projects was conducted. Very few government funded ex-post evaluations of rail projects had been carried out at a Scottish, UK or even at European level. Most evaluations of rail projects were carried out only one to two years after roll-out, resulting in patchy information on the impact of schemes . Studies which had looked at the impact over a much longer time period had tended to be academic studies, with none or little government involvement. Where they had been conducted they had tended to focus on large scale rail investments such as High Speed Rail rather than on smaller scale projects such as the electrification or extension of lines. The findings of the evaluations typically showed that ex-ante appraisals were often optimistic on levels of costs, and could significantly under- and over-forecast demand[3.] Such evaluations tended to ignore the impact rail projects had on cutting carbon emissions, and also in reducing transport congestion levels. It was also noted that there are international differences in the way ex-post evaluations are conducted with North American studies traditionally focusing on the impact on commercial property markets, and European and UK research on changes in demand patterns, and modal shift.

Green Book / Magenta Book / Scottish Transport Appraisal Guidance (STAG)

2.6 The importance of evaluation is also highlighted in two UK Government documents. The Green Book[4] provides clear, if high level, guidance on evaluation and its importance in feeding back lessons into the decision making process. More detailed, but still general, guidance on evaluation is available in the Magenta Book[5].

[3] http://www.transportscotland.gov.uk/strategy-and-research/publications-and-consultations/New-Stations-Study

[4] https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-governent

[5] https://www.gov.uk/government/publications/the-magenta-book

2.7 Transport Scotland's Scottish Transport Appraisal Guidance (STAG[6]) emphasises the importance of evaluation and recommends that an evaluation plan be drawn up as part of the development of the business case for a transport project. However the document is also high level and less prescriptive than the guidance for ex ante appraisal.

2.8 Against this background, this guidance has been developed to provide more detailed information on the key issues to consider. The development of the guidance has been informed by three pilot evaluation projects, as set out in the box below.

Background information on three pilot projects:

To test and refine our developing TS Rail Evaluation guidance we commissioned three pilot evaluations in 2013/14 of recent rail projects in Scotland:

1. Laurencekirk Railway Station;
2. Larkhall – Milngavie line;
3. Airdrie – Bathgate line;

As well as ascertaining whether the projects had met / were meeting project objectives, the contractors were asked to provide recommendations on guidance material to be included in this document on the evaluation of rail projects.

2.9 Before considering the stages and techniques to consider when carrying out an evaluation, the next section outlines how the appraisal stage can assist with an evaluation.

[6] http://www.transportscotland.gov.uk/stag

3. Appraisal

3.1 Consideration of the needs of evaluation should be made at the appraisal stage. Many of the steps for ensuring good practice in appraisal will assist with the project evaluation and are described below.

Good practice in appraisal

Project Objectives

3.2 As recommended in STAG it is imperative that the Transport Planning Objectives for the project are expressed with SMART principles in mind. These objectives should be established prior to the initial appraisal and be used in the subsequent evaluation. Objective-setting will be more effective if it involves all project stakeholders:

Specific – objective should specify exactly what the project will achieve in unambiguous terms;

Measurable – objective should be quantitative and measurable, against which the project can be evaluated in the future;

Attainable – objective should have a realistic chance of success;

Relevant – objective should have a clear purpose and benefit;

Timed – objective should specify the timescales for accomplishment.

3.3 Establishing achievable 'SMART' objectives is a 'balancing act'. On the one hand, objectives should be challenging and stimulating, but on the other they should also be realistic and take into account forecasting uncertainties.

3.4 At the same time as developing 'SMART' project objectives an evaluation plan identifying the future data requirement of the project should be developed. This should set out how the project objectives will be measured and frequency of the data collection. Consideration should be given in the plan to establishing a control group and a baseline. Further information on developing such a plan is given in Section 4.

Case Study from Laurencekirk Railway Station Evaluation on SMART objectives:
The Evaluation of Laurencekirk Railway Station highlighted that the objectives for the project did not meet all of the SMART (Specific,

Measurable, Achievable, Relevant and Time bound) criteria as set out in STAG. In particular, the objectives were not quantified and time bound making it difficult for some of the objectives to be assessed. The researchers acknowledged that it is not always possible to have objectives which are fully SMART.

Project Documentation

3.5 All documentation used to inform the appraisal should be archived and transferred to the stewardship of the sponsoring organisation or Transport Scotland, if the agency has been involved in delivery. An inventory should be prepared of all the material.

Demand Modelling

3.6 The appraisal stage will also typically involve the modelling of passenger demand for the project. To assist with the evaluation a record of the demand forecast methodology should be kept along with the assumptions used. Details of the information which should be collected is given below:

Case study from Larkhall-Milngavie Evaluation on demand modelling:
As part of the Larkhall-Milngavie Evaluation a review of the passenger forecasts and the modelling methodology was carried out. The forecasts for the line were contained in a Strathclyde Passenger Transport (SPT) Modelling report produced in 2000. It was found that there was significant differences between the demand forecasts and actual demand for the four new stations. However they were unable to single out any one key reasons for the difference. It was further complicated by the absence of documentation in the SPT Modelling Report verifying key modelling assumptions or to determine whether impacts such as demand abstraction and demand build-up factors were applied.

i. Choice of Demand Forecasting Methodology

3.7 It is good practice if, before any modelling begins, an assessment of the various modelling approaches available and the pros and cons of each one is undertaken, noting guidance set out in the Station Usage and Demand Forecasts for Newly Opened Railway Lines and Stations[7] report. As the document highlights clear justification for one approach over alternatives should be documented.

[7] http://www.transportscotland.gov.uk/report/new-stations-study-3677

ii. Documentation of Forecasts and Assumptions Applied

3.8 A modelling Record of Assumptions should be produced. This should include:

- a detailed description of the modelling methodology;
- all exogenous factors assumed including their source;
- the levels of demand abstraction and mode switching assumed, including justification and/or evidence for the assumptions made;
- all key modelling parameters including annualisation and ramp-up factors;
- a clear description of the rail service pattern (frequency and journey times) and fares assumed; and any assumptions made about the Reference Case and Do Something changes affecting the competing modes (e.g. journey time, frequency and/or fares).

3.9 Additionally, a full suite of model outputs (ideally in a simple comma-separated variable or equivalent 'flat' format) will facilitate a future comparison of the forecasts against actual demand by an evaluation. This should include comprehensive annual demand forecasts for all stations impacted by any change (i.e. all stations in the study area), and for an extended time period (e.g. for 15 years post-implementation). This will allow the evaluation to compare forecasts against actual demand.

iii. Sensitivity and Risk Analysis

3.10 To account for the impact of potentially volatile drivers such as economic growth on demand forecasts, it is best practice to prepare a range of forecasts to reflect the inherent risk in using a particular set of exogenous drivers. In addition to the 'Central' forecast (based on the most likely or mid-range scenario), typically two sensitivity tests should be performed: 'Low' and 'High'. As already outlined a log of these tests should be kept to inform a future evaluation.

3.11 The 'Low' sensitivity test could, for example, take the more pessimistic forecasts available for GDP, employment and population growth; the 'High' would take the more optimistic forecasts. These could be either drawn from different forecasters (e.g. Experian, Oxford Economics or CEBR), or by adjusting the 'Central' forecasts e.g. +0.5% and -0.5% per annum for the 'High' and 'Low' scenarios respectively.

3.12 In addition to sensitivity testing of the main exogenous drivers, tests should also be conducted around other factors that may influence the demand predicted for the scheme. For example:

- with and without park and ride facilities;
- the inclusion of a significant competitor response (from bus/coach operators);

3.13 If the forecast demand at a particular station is significantly influenced by additional local housing or commercial development, the impact of the likelihood of this development going ahead or going ahead on a smaller scale should also be explored in sensitivity testing.

3.14 As already outlined a log of these tests should be kept to inform a future evaluation.

4. Study Design

4.1 The previous section highlighted how the needs of the evaluation should be considered at the appraisal stage. This section outlines steps to consider early on in the policy cycle to assist with the design of an evaluation study, starting with drafting an evaluation plan. As well as identifying the timing of when to carry out an evaluation, the planning stage should identify how the study will attribute any change as a result of the intervention. A number of suggested methods for doing so are provided.

Evaluation plan

4.2 As has been highlighted in Section 3 consideration of the requirements of a project evaluation should be made early on, ideally at the appraisal stage (as recommended in STAG). As well as ensuring that project objectives are SMART, an evaluation plan for the project should be developed at project inception. The HM Treasury Magenta Book includes a number of useful tips for developing an evaluation plan this includes considering the audience for the evaluation, setting out the evaluation approach, identifying data requirements, obtaining necessary resources, and the dissemination of findings from the evaluation.

Intervention logic model

4.3 To assist the drafting of the evaluation plan, a logic model / intervention logic chain can help in mapping out the inputs, outlining the context in which the project was delivered, and the expected outputs, outcomes and impacts of the rail scheme. The identification of these can help in deciding which research approach should be used as well as help to understand the key outcomes to be measured. These outcomes should cover the transport planning objectives and the STAG criteria, including a wide range of economic, social and environmental outcomes.

Timing

4.4 Decisions on the timing of when the evaluation should take place should be made early on in the awareness that if it is carried out too soon some or all of the anticipated impacts may not have materialised. Alternatively, if left too late it may be difficult to directly attribute impacts to the project.

4.5 It is recommended that project evaluations be carried out at different time intervals such as at 1 year and 3 or 5 years after completion. To assist with the planning of, and deciding when to carry out an evaluation we have distinguished between two different types or stages of evaluation (Stage 1 and

2 evaluations - see below). Ideally both stages of evaluation should be carried out and consideration needs to be given in the evaluation plan to the requirements for each. In addition to these two stages, details of carrying out a land use assessment is also given which would not usually be conducted until 10-15 years after completion.

Recommended timescales for evaluation studies:

Stage 1 Evaluation

Usually carried out around a year following the completion of a project a Stage 1 evaluation is a high level assessment of the extent to which the project is on track to reach its objectives, through the examination of relevant monitoring data (usage/patronage, journey times, journey time reliability and abstraction/displacement from other modes and routes). An assessment of outturn versus predicted cost can also be made. This may also provide an opportunity to assess mitigation measures (to prevent, reduce and where possible remedy or offset any significant adverse impacts on the environment). An assessment of the delivery of the project usually through interviews with key members of staff should also be included (a process evaluation, lessons learned exercise).

Stage 2 Evaluation

It is recommended that a more comprehensive evaluation be carried out when a project has had sufficient time to bed in, typically 3 - 5 years after the completion of a project. It will draw upon the existing types of data outlined for Stage 1 but also collect primary data to assess if the project has achieved its objectives. Surveys of passengers and stakeholders would typically be conducted. Where possible this data should be compared with information collected as part of the baseline studies. It also recognises that there usually is a time lag in secondary data sources being published so by carrying out the evaluation 5-10 years after completion there will be usually be a better selection of data available. Additional elements such a recalculation of Benefit Cost Ratio may also be included, along with an in depth assessment of Wider Economic Benefits and Accessibility.

Land use assessment

Land use changes in response to major transport interventions tend to take a long time to more fully feed through. Land-use experts suggest an assessment after 10-15 years as being appropriate. It is envisaged that using survey techniques to conduct such an assessment would be ineffective after such a length of time and so techniques such as geographic information system (GIS) mapping are employed to compare land-use over time. Mapping land-use at the time the project is announced will enable the best baseline comparison in this respect.

Process Evaluation

4.6 It is important that the views of key stakeholders involved in the delivery of the projects are collected to ensure that lessons are learned for the future. STAG is clear that process evaluation should be undertaken. This information can be collected through face-to-face interviews, or by more structured methods such as questionnaires. This stage should ideally be scheduled in the Evaluation Plan for soon after project completion. This will ensure that most of the key individuals involved in delivering the project are still available to provide their input and that all issues encountered can be easily recalled and recorded. The process should be carried out in a constructive manner in order to encourage openness and honesty about what went well and what aspects could have been improved upon.

4.7 It is therefore recommended that the Process Evaluation is conducted within six months of project completion, with stakeholders primed to note developments during delivery where possible.

Proportionality

4.8 At the planning stage thought should be given to the scale of the evaluation and the extent to which the evaluation plan is proportionate. The cost of the rail project will have an impact on the potential scale of the evaluation. For instance a multi-method evaluation of a major new rail line may be justified but such an approach may not be appropriate for minor enhancements to a station. However, detailed evaluation of low-cost investments may be justified if results are generalizable, e.g. with respect to rural stations.

Determining Attribution

4.9 When drafting the evaluation plan particular consideration should be given to how to attribute any change as a result of the intervention. The following four areas should be considered when designing the evaluation:

(i) Counterfactual

4.10 As the Magenta Book highlights, a key characteristic of a good impact evaluation is that it recognises that most outcomes are affected by a range of factors, not just the policy or intervention itself. To test the extent to which a project was responsible for the change, it is necessary to estimate – usually on the basis of statistical analysis of quantitative data, what would have happened in the absence of the policy also known as the counterfactual. Establishing the counterfactual is not easy, since by definition it cannot be observed – it is what

would have happened if the project had not gone ahead. A strong evaluation is one which is successful in isolating the effect of the project from all other potential influences, thereby producing a good estimate of the counterfactual. The original business case for the project should include some estimates of this and forecast the difference the project might make. Having a baseline can also assist with establishing a counterfactual.

Case Study from report on ex-post evaluations of EU funded transport projects:
In the Frontier Economics et al (2011[8]) report on ex-post evaluations of 10 EU funded transport projects from 2000-06, one of their key observations is that such studies are fraught with the difficulty of identifying the counterfactual. They suggest that one of the problems is that, at the initial appraisal stage, inappropriate assumptions regarding travel demand and growth are often made in the 'without' scenarios. This in turn leads to difficulties in conducting an ex-post evaluation, and may also lead to difficulties in identifying the distribution of benefits between operators and users. By not properly specifying a counterfactual at the start of a project, when unexpected events may occur, such as economic downturn or changes to transport network, it is difficult to disentangle this later on in the evaluation.

(ii) Using Control / Comparison sites

4.11 The use of control and comparison sites map help to understand whether any changes associated with a rail project have occurred directly because of the intervention. For example higher property prices close to a station may be due to the station's proximity to other amenities rather than to the station itself. However establishing a control site may be difficult to achieve due to identifying areas which have not had exposure to similar interventions. For instance neighbouring areas may be affected due to 'spill over' effects from the project. It is recommended that the default control group for rail evaluation projects be data from a national / Scotland wide level.

Case Study from Laurencekirk and Larkhall-Milngavie evaluation on control groups:
For both the Laurencekirk and Larkhall-Milngavie pilot evaluations, Scotland wide data was used as a control group. For the former study, data from Laurencekirk and Aberdeenshire was compared against

[8] http://ec.europa.eu/regional_policy/sources/docgener/evaluation/pdf/expost2006/wpb_final_report.pdf

Scotland-wide data. With the Larkhall-Milngavie evaluation due to the large geographical coverage of the project (it extended over three local authorities: Glasgow, East Dunbartonshire and South Lanarkshire), the national average (i.e. Scotland as a whole) was selected as the most appropriate control group for socio-economic comparisons. However, to assess some demand impacts, another line in the Glasgow area unaffected by major rail improvements was selected as a control.

4.12 Comparison sites may provide an insight into whether a change can be directly attributed to a project or can be replicated elsewhere. Again it may prove difficult to find sites which directly match the intervention. Some general tips in identifying comparison sites include:

- choosing stations which are close as possible to the new stations (but also bearing in mind the above issue on 'spill over' effects). They should be as similar as possible in size, type of area served, rail service level, car park size etc. to the new stations being evaluated.
- areas with 'unusual' features (for example those serving out-of-town shopping centres) should also be avoided.
- recent demand trends at comparison stations should be checked – if there are large 'spikes' or 'troughs' in usage then there may be external factors at play which reduce comparability with other stations.
- locations where there has been a recent change in service or facility provision should be avoided, as should stations where there has been significant recent development in the catchment area (e.g. extensive housing construction, new sports stadium etc.) (Blainey 2013).

Case study from 'Gateways to Prosperity' article:
Control areas were used by Simon Blainey and John Preston (2010)[9] to establish whether changes in the economic and social circumstances of areas with new stations could be directly linked to this intervention. They collected information on 13 stations in South Wales which opened between 1982 and 1990, and then identified 10 sites which had been considered for a new station but where no construction had taken place. They used a variety of datasets from the 1981, 1991 and 2001 census, from Neighbourhood Statistics, the Land Registry, and the Ordnance

[9] Blainey SP & Preston JM (2010) *Gateways to Prosperity? The Long Term Impacts of New Local Railway Stations*, European Transport Conference, Glasgow, 11-13 October 2010.

Survey to establish whether the stations had an impact on employment, property prices and prices in the areas. Their analysis highlighted that in the two decades following station opening the population grew up to 8% more in the area around new stations than in control areas, with the difference only being significant in zones 2-5km for new stations. Findings with regards to household spaces were similar, with on average 600 additional household spaces created over the two decades in the zone 2-5 km from the new station than in the control zones. More significant was the finding that the opening of a new station leads to a 7-10% increase in house prices in the station's postcode sector.

(iii) Baseline

4.13 Establishing a baseline, and collecting data at a set point before an intervention will also assist in establishing the impact of a project ('before'). To ensure that a baseline can be established at the start, consideration needs to be given early on before a project commences of such data requirements. For instance, will existing surveys routinely collect the required information or will new primary data collection be required?

4.14 Ideally some baseline evidence should be collected soon after the project has been announced as individual behaviours, property prices or land use may be impacted immediately. This needs to be balanced against the fact that there will be a gap of some years between announcement and opening. Most evidence, such as survey evidence on travel and employment patterns, is best collected soon before opening. Elements to include in a baseline study may include drawing upon existing data sources as well as carrying out primary research with potential users of new infrastructure / services, and other groups such as businesses. Further information on data sources is in Section 5.

Case study from Laurencekirk and Larkhall-Milngavie evaluations on the difficulties of not having a baseline:
As baseline fieldwork had not been carried out for both the Laurencekirk and the Larkhall-Milngavie projects it was difficult attributing change, so the outcome evaluations depended on the memories of users and residents. Passengers who completed the online survey were asked about their behaviour before the station was reopened, and what they would have done had the station not reopened. Whilst this worked well for the Laurencekirk study (we assume that this is because the station had been operational for less than 5 years), there were difficulties in using this approach for the Larkhall-Milngavie project. There are also are inherent risks with the accuracy and reliability of the responses collected using this approach.

(iv) Determining outcomes

4.15 As was highlighted at the start of this section, producing a logic model / intervention logic chain of the rail project can assist with identifying expected outputs, outcomes and impacts of the scheme. The mapping exercise can identify what data needs to be collected on factors which may determine outcomes. Consideration at the planning stage should be given to whether this data will allow for statistical analysis to be carried out which will allow these factors to be controlled, allowing for effects to be disentangled.

Recalculating the Benefit Cost Ratio

4.16 Consideration should be made to whether the evaluation should include a recalculation of the Benefit Cost Ratio (BCR). This may be particularly relevant if the original BCR at appraisal stage had been unable to include the full range of costs and benefits or if the BCR had been low or high. Recalculation can also provide important insight into calculating BCRs for future rail projects. It would usually only be appropriate to carry out such an exercise for Stage 2 evaluations.

4.17 Projects which had been commissioned prior to 2010 are also likely to have been appraised using a 30 year appraisal period, rather than the current preferred period of 60 years. For these projects it may be worth considering recalculating the BCR using the 60 year period and taking into account other changes to appraisal methodology, if proportionate.

4.18 At the research planning stage consideration should be given to what data will be required to recalculate the BCR, and also whether such information is easy to access. For instance, outturn values to replace some of the forecast values will also be available for projects which have been operational for a few years.

4.19 Further information on recalculating the BCR is given in the Analysis and Reporting Section (Section 6).

5. Data Sources

5.1 At the study design phase the evaluation plan should identify the secondary and primary data requirements of the project. This section sets out a number of key pre-existing data sources (secondary data) which could be used, along with primary data collection methods. Section 6 sets out how the data collected should be analysed.

Secondary data

5.2 There are a number of secondary data sources both at a Scottish and UK level which could be drawn upon to inform evaluations of rail projects. Details of some of the key sources are given below but there are many others:

Scottish Transport Statistics

http://www.transportscotland.gov.uk/strategy-and-research/publications-and-consultations/j251205-000.htm

5.3 This annual publication draws together data on rail services in Scotland (along with information on other modes of transport).

National Passenger Survey

http://www.passengerfocus.org.uk/research/publications/national-passenger-survey-spring-2012-main-report

5.4 Passenger Focus's 'National Passenger Survey' is the main passenger survey to measure passenger satisfaction in the UK. This annual survey consults with around 50,000 passengers each year across the UK to produce the National Passenger Survey (NPS).

ORR's National Rail portal

http://dataportal.orr.gov.uk/

5.5 This portal collects together a selection of secondary data sources on rail, for example on station usage, and allows for personalised reports to be generated.

Rail data

5.6 Ticket sales data are an excellent source of initial data, and the rail ticket database LENNON holds information on all National Rail tickets purchased in the UK from 2001. If actual patronage has met or exceeded that forecast then ticket sales data are a good early indicator this. This should also involve

extraction of data to assess abstraction from other rail routes.
Data on revenues and ticket types should also be extracted. Given the extent to which they dominate the typical business case, actual journey time changes should also be assessed.

5.7 Identification of the type of rolling stock used can help assess CO2 impacts, as consumption and therefore emissions can vary for different types. International datasets on the rail projects could be used for comparison purposes. For instance UIC (the international union of the railway) has a database of the costs (mainly ex-ante estimates) associated with 166 high speed rail projects from around the world.

National Rail Travel Survey

5.8 The National Rail Travel Survey (NRTS) presented a comprehensive picture of weekday rail travel across the whole of Great Britain, aiming to represent all rail travel at all 2,500 stations in Great Britain on a typical weekday. It may be a useful source of background baseline information depending on timing as surveys were conducted in 2004-05.

5.9 Information was collected from passengers by self-completion questionnaire, covering the following topics:

- rail stations used
- time of travel
- access and egress modes
- origin and destination addresses
- trip purposes
- ticketing information
- demographic information

Other modes

5.10 It will likely be necessary to source data on users of other modes, to assess any modal shift from private motor vehicles (and test for any CO2 reductions) and any abstraction from other public transport modes to test for net changes to private sector (i.e. bus) revenues. TS hold monitoring data, including on numbers of road users though bus patronage data is held by commercial bus companies and may be more difficult to obtain for specific routes.

5.11 Data on other modes of travel collected by Transport Scotland's LATIS (Land-use And Transport Integration in Scotland) may also be useful. The service is offered to central and local government officials to assist with the appraisal of transport projects and programmes. It can be used to formulate and guide the

deployment of transport, planning and the environmental policies. LATIS has a robust database of transport, land-use and demographic data which is linked to a multi-modal transport and land-use modelling.

Non-transport data

5.12 For schemes with objectives relating to, for example, regeneration, accessibility or social inclusion, it may be necessary to source non transport data such as the following:

- Local area population and employment data
- Local area GVA data
- Business statistics

5.13 The following websites provide useful information on the following areas:

Registers of Scotland Residential Property Price Database

https://www.eservices.ros.gov.uk/shp/ros/shp/presentation/ui/pageflows/search.do;jsessionid=nvL9Kw2pL2jbPpPn3Q5nB0VtVrF9L1LPrbVd1tdVHf9JBTrx1Hkr!1635453261

5.14 The Register of Scotland property price database provides information on residential property prices. This is a useful resource for understanding how rail schemes may impact on property prices. House price data at postcode level can be purchased from the Registers of Scotland to allow small area analysis.

SG Property Price Time Series Data

http://www.scotland.gov.uk/Topics/Statistics/Browse/Housing-Regeneration/HSfS/KeyInfo

5.15 SG also publish a time series from the 1st Quarter 2003/04 to the latest quarter published for mean average price, median average price, volume of sales and market value by local authority.

NOMIS

https://www.nomisweb.co.uk/home/newuser.asp

5.16 The NOMIS online portal provides up-to-date labour market statistics from official sources for local areas throughout the UK. These surveys include the Labour Force Survey (LFS), claimant count, Business Register and Employment Survey (BRES), New Earnings Survey (NES), and the 2011, 2001, 1991 and 1981 Censuses of Population.

Small Area Income Estimates

http://www.improvementservice.org.uk/income-modelling-project.html

5.17 The Improvement Service published in 2013 small area income estimates for Scotland.

Scottish Neighbourhood Statistics

http://www.sns.gov.uk/default.aspx

5.18 This site gives access to a wide range of information for any area of Scotland from a Local Authority right down to Data Zones of more than 750 households. Information which is reported on includes population, economic activity, benefits and tax credits claimed, health, education, and housing and physical environment.

5.19 It will be important to bear in mind a number of caveats relating to these secondary data. First and foremost, observed impacts will not necessary have been caused by the transport intervention, so statements of causality should be made with great care. Displacement of activity from other areas should be taken into account (though such redistribution may be a valid objective). Second, the reliability of these data (e.g. GVA) decreases considerably at lower levels of disaggregation. Finally, some datasets are subject to significant time lags, sometimes of several years, which prevent their meaningful use within standard evaluation timeframes.

Primary data

5.20 In order to probe impacts more fully it may be necessary to conduct some primary research to collect new data. For large schemes involving significant investment this should be mandatory. Examples of primary data collection exercises are given below:

On-board or at station surveys

5.21 On-board surveys of passengers are a widely used technique to collect information not readily available through secondary data sources such as ticket sales. These structured surveys can collect demographic information on passengers, reasons for journeys, changes in travel behaviours, and satisfaction with services. Such surveys are typically administered by a researcher on the train who will talk through the questionnaire. Alternatively paper surveys can be left on seats on the train, or distributed during a journey

with the option of handing the survey in when leaving the train or posting back at a later date. Information to be collected by such surveys could include:

- Journey purpose (e.g. % of business, commuting and leisure travellers)
- Journey patterns in the absence of new service
- Potential agglomeration impacts:
- Moved job as a result of new service (impact on earnings)
- Moved house as a result
- Found new markets
- Found new business partners
- Increased labour supply (moved into employment, increased hours etc)
- Accessibility impacts (% with no car etc.)
- Equalities monitoring

Travel Diaries

5.22 Travel diaries can be a useful method for collecting in-depth information on travel behaviours. Respondents are encouraged to fill in information for all journeys and modes of transport over a particular period of time This for instance can show how respondents travel to and from a station, and can be used to highlight changes before and after an intervention. A travel diary over a longer period of time (7 days for instance) is more likely to provide richer information about travel patterns, however longer diaries will add to the cost of collecting the information and respondents may have difficulty in recalling all information about historic travel accurately Incentives are often used to encourage respondents to complete diaries.

5.23 This approach is used by the Scottish Household Survey (SHS) Travel Diary in collecting the travel behaviours for a random adult from the household sample. The SHS Travel Diary collects information on the purpose, origin, destination, duration, distance and mode of travel for all of the journeys made by the respondent the day prior to the interview. This information can be used alongside demographic and behavioural information about the random adult to provide estimates of travel patterns for particular subsets of the Scottish population. Transport Scotland publish annual estimates of travel from the SHS, including travel diary information in Transport and Travel in Scotland[10].

[10] See http://bit.ly/TSStats-TATIS.

Online surveys and data collection tools

5.24 On-line survey and data collection tools can be a useful method for collecting information from existing 'virtual' samples, such as survey panel members. They are typically low cost in setting up, and save time in the inputting and analysis of data. However they may not always be appropriate for collecting information on travel behaviours or associated issues, as they may require details of the participant's email address. They also exclude users who do not have internet access.

Semi-structured and unstructured interviews

5.25 Semi or unstructured interviews are usually conducted either face-to-face or on the phone and offer the opportunity to explore in greater depth issues which cannot be routinely collected through a structured survey. They can also be used as a method to explore responses to an earlier survey. Again they can be time consuming both in setting up, conducting, transcribing and analysing.

Telephone interviews

5.26 The telephone can be a useful method to administer surveys, as well as carry out in-depth interviews with rail users and stakeholders. Telephone interviews can be a targeted at householders in the immediate vicinity of a new line or station, and collect the views of users and non-users. The downside of contacting all households in the vicinity could be that insufficient number of users are sampled. As with on-line surveys, effective telephone interviews of rail users are best carried out with pre-existing sample who have opted to participate in research and have provided phone contact information. This could include Season Ticket Holders, or rail users who have participated in a promotion and have shared contact details.

Focus groups

5.27 As with semi and un-structured interviews, focus groups are a useful method in probing respondents on issues which may have been raised in earlier survey / interviews. Usually such groups comprise of between 6-12 participants who are guided by a facilitator in discussing related issues, opinions, beliefs, and attitudes. The information collected during focus groups would usually be less accessible without the interaction found in the group setting. Groups could comprise of rail passengers and other stakeholders. Due to the investment in time required by participants, incentives and travel costs would usually need to

be considered to ensure that an adequate number of respondents attend. Those attending the focus groups may not be representative of all users / stakeholder groups, and focus groups only collect information from a sample number of participants.

Count data

5.28 Manual or automatic traffic counts can be used to collect data on how respondents are travelling to and from the station (for instance by bike or by car). They can for instance be used on roads in the vicinity of a new line to see whether the new service has impacted on road traffic levels. However count data is limited as information on the purpose of a journey is not collected, and there are also issues around double counting.

New technology

5.29 New technology provides the opportunity to collect additional data on rail users. For instance mobile phone Bluetooth technology is used to monitor journey time on Scotland's roads, and crowd movement and behaviour at sporting events. Such an approach could be used on the trains to monitor train usage, as well as journeys to and from stations. As with count data there are limitations with this approach as information is not collected on the purpose of journeys.

GIS mapping

5.30 Geographical information systems (GIS) mapping can be used to track land-use changes where new stations have been built or services improved, to help assess whether there has been any impact on housing or business developments. Best practice would be to map relevant areas around the time of project announcement and to track changes periodically over time, until 10-15 years following commencement of services.

6. Data Analysis and Reporting

6.1 This final section of the guidance provides an overview of some of the steps to consider when analysing the data, and also on how to report it. The approach used for analysis should link back to the way the study was designed as outlined in the evaluation plan as well as the project objectives outlined in STAG. A key part of the data analysis is to assess how far changes can be attributed to the rail project itself. Below are some approaches which can be used to do this.

Analysis of attribution

6.2 The analysis plan for the evaluation should include the key comparisons to be made to order to determine the attribution of the rail project to the outcomes measured.

Before and after comparisons of key metrics

6.3 One way to understand the contribution of the rail project to a key outcome / metric is to compare the value observed at baseline (before) with the value observed after (at the point of the evaluation). The difference between the two observations is one measure of the impact of the project (although confounding factors may need to be taken into account – see below). This approach is dependent on baseline data having been collected.

6.4 Before and after comparisons should report on journey time savings, changes in patronage and changes in trip origin (home location) and trip destination (work location) before and after the project opening.

6.5 For studies that have used control groups, for each metric or outcome being measured, the average value of the control group and the average value of the study group should be compared. The difference between these two groups gives the contribution of the rail project to the key metric/outcome.

6.6 Alternative approaches to identifying attribution include making comparisons between views of groups who had moved into the area since the project opened and those who have lived in the area longer.

Case Study from Laurencekirk Station Evaluation on establishing a counterfactual:
The Laurencekirk Station Evaluation user survey asked residents if they

had always lived in Laurencekirk, or had moved to the town after the station had reopened. This provided a useful counterfactual and allowed for comparisons to be made between the two groups of residents. For instance long term residents were asked how they made their specified journeys prior to the reopening, and the latter were asked how they would have made the journeys if the station had not reopened.

6.7 Adjustments for abstraction from other railway lines may also be required to attribute changes in rail patronage to the railway project. Data from LENNON could be used to investigate if the new infrastructure has resulted in users switching from existing rail services.

Case Study from Airdrie to Bathgate evaluation on abstraction:
One of the original planning objectives of the Airdrie to Bathgate project was to provide an alternative to the Edinburgh – Glasgow main line, which would reduce congestion at peak times. To understand if abstraction had occurred the evaluation used LENNON data to look at annual Edinburgh and Glasgow patronage on the line compared to that on the Falkirk and Shotts lines which also offered direct services to the two cities.

6.8 Accessibility analysis should be carried out to ascertain if the project has made different opportunities and facilities more accessible. For example it can be used to calculate differences in journey times to employment and education centres before and after the project, or provide a percentage of the catchment population within the time threshold of a main destination.

Case Study from Laurencekirk Evaluation on accessibility analysis:
Accessibility analysis was undertaken using Accession Software to assess how public transport journey times from Laurencekirk to key employment and education centres had changed since the reopening of the station. The analysis did not consider car-based or mixed-mode (i.e. Park and Ride) journey times, but rather provided a before and after analysis of the times of journeys made by public transport only.

Postcodes were selected within a radius of three miles of Laurencekirk centre and the key destinations. Tests were undertaken over the time period 07:00-10:00am. The locations of each of the destinations in Aberdeen and Dundee and the three mile postcode catchment were then mapped.

6.9 In some cases, subject to the data being available, it may be possible to use regression analysis to attribute the differences between subject populations

and control group populations by using techniques such as Difference in Difference estimators. This could compare the average change before and after the rail project opening in the project location, against average change over the same time period in a control area.

Changes in the area profile and other confounding factors

6.10 The Magenta Book highlights that evaluation results should be set in the context of other knowledge about the project and/or the context in which it was delivered.

6.11 The evaluation should analyse and report on other changes (confounding factors) that have happened in the local area over the period since the rail project was announced. Examples of area based confounding factors are given in the box below.

Examples of area based confounding factors to analyse and report before and after comparisons of:

- The state of the economy (GVA).
- Population.
- Housing supply.
- House prices.
- Retail developments and other land use developments.
- Other transport modes (new services or bus services withdrawn, changes in the town centre).Opening or closing of major employers.
- Major events in the area (such as Commonwealth Games).

6.12 In some cases it may be possible to use regression analysis to adjust for these confounding factors, subject to the data being available.

6.13 Where it is difficult to fully quantify confounding factors, some narrative should be included in order to consider the overall impact of factors on key metrics / outcomes.

Analysis of geographic and distributional impacts

6.14 As well as considering wider changes to the area the analysis of the project should take into account the impact of the rail project on geographic areas, regions and communities. Geographic impacts should include s*pillover effects,*

where a policy implemented in one area has an impact on neighbouring areas (which may be positive or negative) and the *two-way road effect* where some negative impacts may be felt in one location due to the competitive impact at other connected points. For example, a previously poorly connected area may benefit from having greater employment opportunities outwith the area for residents, but retail businesses may suffer if residents can access cheaper shops with greater variety further afield.

6.15 The rail project may for instance result in a change in the location of economic activity which should be considered. Transport investment, other things being equal, can make the affected location more attractive destinations for investment (see TIEP study[11]). In some cases, a rail improvement may affect several locations in a positive or negative way.

> *Case study from Transport Investment and Economic Performance report:* The TIEP (2014) report highlights the complexity of transport improvements which affect a number of different locations. It highlights how reducing transport costs between centres will generally make both places more attractive locations for investment, but it could impact on one location than the other. Potentially drawing investment into one of the linked locations at the expense of the other. As is suggested the only way to demonstrate if this is happening is by having a robust counterfactual in an evaluation.

6.16 Consideration should also be given to the impact of the rail project on different groups in society (income and/or deprivation decile and other sub-groups such as age, sex, and disability groups) as it may allow for disadvantaged people to access opportunities that most people take for granted. The distributional impacts should explore *substitution and displacement effects.* For example there may be positive impacts on those directly affected by a policy or programme, but negative effects on others.

11

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

Case study from Larkhall – Milngavie Evaluation on Social Inclusion:
To assess whether this project had met its social inclusion objective of connecting people to employment opportunities, social networks, education and leisure activities, the research team used a number of different measures. Using data from the user survey respondents' journey purpose and destination were analysed to assess whether the project had provided a valuable link to employment, education and leisure opportunities. Data from the user survey was used to calculate how much nearer the new stations are 'as the crow flies' to a respondent's home address than the nearest station before 2005. Accessibility analysis was undertaken to assess the impact on journey times between the Larkhall and Hamilton areas and Glasgow Central Station.

The analysis showed that the project had contributed to the 'Social Inclusion' objective by promoting access to a range of opportunities and facilities in the surrounding area in particular employment and education opportunities with 64% of respondents using the line commuting purposes. It also improved the availability of public transport through the reopening of stations. For the majority of Larkhall residents the nearest station is now within walking distance – previous it been over 6km on average. For many the project had reduced public transport journey times to work, education and leisure.

Analysis of environmental impacts

6.17 Where applicable the analysis should analyse and report the environmental impact of the project, including the impact on emissions (which may have arisen from modal shift as well as impact from the construction), air quality, noise, landscape and habitats. The level of detail of the analysis should be proportionate and may depend on the original Transport Planning Objectives.

Case study from Larkhall Milngavie evaluation on environmental impact:
The environmental impacts of this scheme were considered using evidence collected by the User Survey on modal shift from car to rail, and also changes in car ownership. Drawing on this evidence the researchers suggest that the scheme has led to a decrease in the number of car trips and a reduction in greenhouse gas emissions.

As part of their analysis other considerations were examined such as emissions from extra rail services, and the negative impacts on the landscape and existing habitats as a result of constructing the new line and stations. These impacts were thought to be minimal, since the line followed an existing, abandoned route. Noise levels of the new line were also considered but it was thought these would be minimal.

Comparison between appraisal and outturn from the evaluation

Demand forecasts

6.18 In order to learn lessons for forecasting, the rail patronage resulting from the rail intervention should be reported and compared to the forecast demand forecasts from the appraisal.

6.19 The reasons why the forecasts differ from the outturn should be summarised. This may include confounding factors or design issues such as effective use of smarter measures including travel planning as well as inclusion of park and ride and cycle parking.

Case study from Laurencekirk on demand forecasting:
One of the key tasks of the Laurencekirk evaluation was to explore why the forecast number of passengers using Laurencekirk station from the original 2004 STAG report was significantly lower than the actual number of passengers. In reviewing the appraisal report it became apparent that it did not contain enough detailed information, and sufficient clarity on the assumptions made to forecast demand.

They recommended that to assist future evaluations that key underlying cause and effect assumptions must be recorded and clearly laid out in the STAG Report. Because the approach was high-level and not totally consistent with the expected methodological approach, it was not possible to simply re-run the demand analysis using the outturn data.

Recalculation of original Benefit Cost Ratio

6.20 As discussed in Section 4, where possible, the BCR should be recalculated. The recalculation of the BCR also provides an important insight and 'lessons learned' for future transport appraisals and will inform the recommendations for the Rail Evaluation Guidance.

Case study from Larkhall-Milngavie evaluation: Recalculation of BCR:
A key part of the specification for this project was to recalculate the rail project's BCR. Cost-benefit analysis undertaken at the economic appraisal stage of the Larkhall-Milngavie rail project in 2000 indicated a BCR of 0.66 for the project (excluding wider economic benefits). The majority of the benefits were expected to come from journey time savings for existing and new public transport users. However, it was expected that the project would bring other benefits that could not be easily monetised and captured within the formal cost-benefit appraisal.

Since the original economic appraisal of the Larkhall – Milngavie project

was carried out, there had been methodological developments to economic appraisal (e.g. extension of the appraisal period from 30 to 60 years and changes to the discount rate applied to future year costs and benefits).

To overcome this, two sensitivity tests were run. Despite this, inconsistencies remain between the two appraisal methodologies (e.g. appraisal assumptions such as values of time which will have been updated since 2000 as a result of new research).

Benefit Cost Ratio Sensitivity Tests

Sensitivity Test	Assumption	BCR
1	30 year appraisal period; 2010 price base year; 3.5%/3.0% discount rate	**2.43**
2	As per Test 1 but with 1994 price base year and 6% discount rate	**1.92**

The sensitivity tests show, however, that even with a shorter appraisal period and higher discounting of future year benefits, the outturn BCR is higher than the appraisal BCR.

6.21 Any subsidy required to run a new service will not feature explicitly in the BCR as it is already implicitly represented by the netting off of revenues, and operating and other non-construction costs. However, it will be of interest and should thus be reported in tabular form in any evaluation.

6.22 If assets such as land which have been purchased by Government or the project sponsor for the purposes of the project are subsequently disposed of at a profit (or loss) this should be recorded if possible, though presented separately, as a sensitivity, to the core BCR.

Presenting uncertainty

Sensitivity testing and statistical analysis

6.23 A range of analyses can be conducted to explore the uncertainty in the main results. Where it is not possible to carry out uncertainty analysis, narrative should be provided on whether or not there is sufficient evidence to draw a conclusion on impacts or instead whether there is insufficient evidence or alternative a zero effect.

Examples of uncertainty analysis include:

- Sensitivity testing around key assumptions.
- 'What if' analysis scenarios.
- Reporting of statistical significance by presenting standard deviations / confidence interval around differences.
- Use of ranges.
- Modelling longer term effects.

Generalisability of results

6.24 Reasons why the findings may not be generalisable to other rail projects should be considered. This may include reasons such as that the area/data are not representative of the wider population. It may also include reasons such as type of rail project: new station, new line, electrification of line etc.

Document methodology

6.25 The research methodology should be documented (see the Magenta book for further details on this), commonly as part of a separate technical report rather than in the main report. It is essential that the information remains available, even after all those working on a project have moved on. This should include research tools, such as questionnaires and topic guides used for qualitative/quantitative studies, as well as associated documentation, such as introductory letters and explanatory leaflets.

6.26 Details should also be reported on key steps taken to process and analyse the data.

Examples of steps taken to process and analyse the data include:

- data cleaning or imputation of missing values;
- weighting for non-response;
- how a final statistical model was selected; and
- how standard errors were calculated.

Dissemination of results

6.27 Planning on how the evaluation findings will be published should begin at the research design phase, and be a key part of an evaluation plan. Reporting an evaluation means more than writing a final report (see the Magenta Book[12] for further details on this).

6.28 The aim of the reporting process throughout a project is to ensure the evaluation commissioners, partners and stakeholders are consulted about research methods, progress and results on an agreed basis.

6.29 As discussed above, a useful first step is to report how the new evaluation findings compare with previous knowledge, particularly where there are clear consistencies or inconsistencies. New hypotheses may be required to explain the latter. It is useful to highlight research questions that emanate from the evaluation to inform future planners of research programmes and evaluations.

6.30 Scottish Government Social Research Guidance[13] for contractors also includes some useful tips on writing outputs from research projects. The guidance contains three steps (CAR) to consider when drafting research reports:

Example of steps to consider when drafting research reports from 'Scottish Government Social Research Guidance':

***Context*:**
outline the policy issue or managerial problem the research was seeking to address, as well as the aims and objectives of the project. Refer back to these throughout the report, and also highlight if these were redefined at any point. Highlight earlier research and the contribution current research may make.

Approach:
Outline your methods, including the design of the study, the sources of data and details on the sample, the response rate and analysis techniques. You should outline your approach as transparently as possible so that it can be scrutinised for quality / relevance / robustness. There should be clear documentation that the methods were implemented, along with a record of any changes. Describe how you worked with policymakers / decision makers on the project.

[12] https://www.gov.uk/government/publications/the-magenta-book

[13] http://www.gov.scot/Topics/Research/About/Social-Research/Guidance-for-Contractors/SR-ContractorHandbook. See Section 2

Results:
Summarise your results to show how they support the conclusions you have presented highlighting themes and messages. Conclusions should be drawn on the basis of the findings. However if they are inferred from external material / other sources then this should be made clear. Use graphs and tables if they will improve understanding.

Further copies of this document are available, on request, in audio and large print formats and in community languages (Urdu; Bengali; Gaelic; Hindi; Punjabi; Cantonese; Arabic; Polish).

اس دستاویز کی مزید کاپیاں آڈیو کیسیٹ پر اور بڑے حروف کی چھپائی میں اور کمیونٹی کی زبانوں میں طلب کیے جانے پر دستیاب ہیں، برائے مہربانی اس پتہ پر رابطہ کریں:

এই ডকুমেন্ট-এর (দলিল) অতিরিক্ত কপি, অডিও এবং বড়ো ছাপার অক্ষর আকারে এবং সম্প্রদায়ভিত্তিক ভাষায় অনুরোধের মাধ্যমে পাওয়া যাবে, অনুগ্রহ করে যোগাযোগ করুন:

Gheibhear lethbhreacan a bharrachd ann an cruth ris an èistear, ann an clò mòr agus ann an cànain coimhearsnachd. Cuir fios gu:

इस दस्तावेज़/काग़ज़ात की और प्रतियाँ, माँगे जाने पर, ऑडियो टैप पर और बड़े अक्षरों में तथा कम्यूनिटी भाषाओं में मिल सकती हैं, कृपया संपर्क करें:

ਇਸ ਦਸਤਾਵੇਜ਼/ਕਾਗ਼ਜ਼ਾਤ ਦੀਆਂ ਹੋਰ ਕਾਪੀਆਂ, ਮੰਗੇ ਜਾਣ 'ਤੇ, ਆੱਡਿਓ ਟੇਪ ਉੱਪਰ ਅਤੇ ਵੱਡੇ ਅੱਖਰਾਂ ਵਿਚ ਅਤੇ ਕੰਮਿਉਨਿਟੀ ਭਾਸ਼ਾਵਾਂ ਦੇ ਵਿਚ ਮਿਲ ਸਕਦੀਆਂ ਹਨ, ਕ੍ਰਿਪਾ ਕਰਕੇ ਸੰਪਰਕ ਕਰੋ:

此文件有更多備份，如果需要，語音版本和大字體版本及少數種族語言版本也可提供，請聯絡：

يمكن أن تطلب النسخ الأخرى من هذا المستند كالتسجيل الصوتي والخط المكبر ونسخ بلغات أخرى، يرجى الإتصال على:

Aby otrzymać niniejszy dokument w innej wersji językowej, na kasecie lub w wersji z powiększonym drukiem, prosimy o kontakt:

Transport Scotland, Buchanan House,
58 Port Dundas Road, Glasgow, G4 0HF
0141 272 7100
info@transportscotland.gsi.gov.uk
www.transportscotland.gov.uk

Any enquiries regarding this document / publication should be sent to us at info@transportscotland.gsi.gov.uk

This document is also available on the Transport Scotland website: www.transportscotland.gov.uk

Published by Transport Scotland, May 2015

An agency of Buidheann le

The Scottish Government
Riaghaltas na h-Alba